CHAPTER

Case Studies and Examples

14

INFORMATION IN THIS CHAPTER

- Introduction
- Case studies
- Example scenarios
- Maltego machines

INTRODUCTION

After working with so many tools and techniques and going through so many processes of information gathering and analysis, now it's time to see some scenarios and examples where all this comes together for practical usage. In this chapter we will include some real scenarios in which we or people we know have used OSINT (open source intelligence) to collect the required information from very limited information. So without wasting any time let's directly jump into case study 1.

CASE STUDIES

CASE STUDY 1: THE BLACKHAT MASHUP

One of our friends returned from Black Hat US conference and he was very happy about the meetings and all. Our friend works for a leading security company and takes care of US sales. He was very excited about a particular lead he got there. The person he met there was in senior position of a gaming company and interested in the services offered by our friend's company. They had a very good networking session in the lounge while having drinks and in excitement he forgot to exchange the cards. So he had to find the person and send him the proposal that he committed.

- Problem No. 1: He forgot his full name but remembers his company name and location.
- Problem No. 2: While discussion the other person said that as many people approach him for such proposals, he uses a different name on LinkedIn.
- Problem No. 3: We know the position of the other person, but it is not a unique position such as CEO or CTO.

Hacking Web Intelligence. http://dx.doi.org/10.1016/B978-0-12-801867-5.00014-8

When he came to us with this case we had a gut feeling that we can find him. We have some sort of information about the person though that does not include any primary information such as e-mail address or full name.

These are the steps we followed.

The first thing we asked him was that whether he can recognize the person's picture or he forgot that also, to our good luck he said "yes." So the major point in our case was that if we will find some profiles of persons he can validate and confirm the right person.

Step 1:

As usual we first started with a simple Google query with the first name, his position, and company name. Let's say his position is senior manager and company name is abc.inc. The query we used is

Senior manager abc.inc

Step 2:

We tried the same in Facebook to get a profile equivalent to that but never the less no leads.

Step 3:

We went to LinkedIn tried that simply and failed.

Step 4:

We went to that company profile page and tried to visit all the employees' profiles but we found that there are more than 7000 registered employees therein LinkedIn and it's really a tough task to find anyone from there.

Step 5:

As we covered in Chapter 2 LinkedIn provides an advanced search feature and we have some of the direct data for the fields. So we have decided to use that.

https://www.linkedin.com/vsearch/p?trk=advsrch&adv=true

We filled data in the fields such as title, company, and location as our friend has these information. In result we got many equivalent profiles but this time we got far less results which we went through one by one and finally found the person in twenty-first result as he had shared that he has recently been to the conference. After visiting his profile we got a bit more details about that person and our friend confirmed that we got what we were looking for.

What could have been done after this?

We might get his primary e-mail id, company e-mail id, and other details using different sources such as Maltego or simple Google. Using the image we might go for a reverse image search to get related image and the sources. We might get the blogs or websites created by that person and many more. There were endless possibilities but we stopped there because for the time being that was out of our scope. We sent the link to our friend, who then sent him connection request and later got the deal and so we got a small treat.

CASE STUDY 2: A DEMO THAT CHANGED AUDIENCE VIEW

We can't forget that demo because it changed many audiences' view on security. We were working on a project with a client who wanted to become partner with our organization. We were told that we need to provide some quality reports and all and then have to provide a report walkthrough to our prospective client. Later we were informed that some delegates including the operational manager and a senior consultant will visit us. We requested our management to let them sit in our internal knowledge sharing program where we were giving a demo on OSINT and how it can help in penetration testing and after that we can have report discussion. Our management agreed on that so we prepared a special OSINT demo.

We had the names of the delegates so the first thing we did was that quickly visited their profile to understand their background to know more about them. LinkedIn helped us a lot in this. We got bit understanding on their likes, dislikes, and technical background. We found that both were having a very good technical background in security so we planned the demo accordingly.

We started the demo with basics such as what is OSINT, why it is more important now a days, and so on. We moved on to talk about its importance in pentesting. Then we moved to how to collect different data about an organization or an application (web application without sending a single packet).

- Problem: We cannot demonstrate anything that requires access to their environment for which we don't have authorization.

We decided to start with the website of our client and then will jump to something interesting.

Following are the steps we followed:

Step 1:

We opened Maltego and added the domain name of our client. There is an option in Maltego to add a domain as an entity and we did the same and then ran some transforms to get different data such as Name server records and many more.

Step 2:

Running the buildwith transform on some of their domains displayed the technologies being used. They were happy seeing it in a graphical and organized way. Other transforms helped us to discover some other related domains/subdomains.

Step 3:

In one of the domains, we found that a very older version of php is being used. So what we did was simply opened the application in a browser where Shodan and Punk Spider plugins were enabled and running. The moment we opened that application in browser Shodan showed that application is vulnerable to Heartbleed and some sensitive ports opened while the punk spider provided that there were two blind SQL injections and some cross-site scripting vulnerabilities present previously.

Running an advanced Google search for the same parameters (vulnerable to SQLI and XSS) on their other domains (discovered in previous step) provided some URLs. This suggests that they might also be vulnerable to the same vulnerabilities.

```
site:example.com inurl:vulnpar
```

They were amazed by the results. They told us that it was a very older site they used to use and due to some technicality they forgot to bring that down. But we found some interesting facts about that application and we were happy.

Step 4:

Then we again opened Maltego where we ran the transform "Domain to Email address" other similar ones, where we collected some of the e-mail addresses.

Step 5:

Then we showed them the local transform written by us, HaveIBeenPwned. It's based on the API provided by Troy Hunt. In the previous chapter, we just covered about its working but just for the information it checks in the database of the popular sites where there is an account breach. If the e-mail id of any person has been breached there then it would have popped up an alert. After running this transformation in all the e-mail ids that we got from the domain name luckily we found two of their colleagues' accounts have been breached in a popular product-based company site.

It was something new for our clients and they immediately informed their colleagues about this. Though we got their attention, yet we wanted to demonstrate the impact of such information. We could have simply explained that there were sites to get particular passwords related to these e-mail ids such as pastebin, but we tried to show the bigger picture.

Step 6:

Then we selected those two e-mail ids and ran another transform written by us that is "Email-Rapportive." It is based on the service Rapportive: http://www.rapportive.com/ (based on the code of Jordan Wright). What it does exactly is that finds information about the person based on his e-mail address and provides us result. The result contains LinkedIn, Facebook, twitter profile links along with the name and job title of the person.

So basically it helps us to get the person's social existence and some important primary details such as full name.

Step 7:

After that we ran a Maltego machine on the e-mail generated in a previous step. A machine is nothing but a collection of one or more transforms to run in a collective fashion. We discussed a bit about this in Chapter 6 and will discuss more later in this chapter. So this machine discovers other e-mail addresses of similar pattern on e-mail services such as Gmail, Yahoo, Outlook etc. We got many results and explained them

that there are people who use same password for different e-mail ids. So if someone can collect a password associated with a breached e-mail id and can find all other e-mail ids and registered website details, then there is a possibility that attacker might compromise some or all the accounts of the victim based on the information he collected.

We found our delegates understanding the risk associated with this scenario and they were happy that we presented it in a live demo to them. The only question from their side was that though IDS/IPS and Firewalls installed in their infrastructure if someone tries to do the same, will they get the IP address from the log?

The answer to which is that we used all openly available information to show how we can break into the security. The tools used such do not send any malicious packet to the original infrastructure, which makes it almost impossible to identify who is performing such enumeration. Apart from this there are various anonymity techniques such as Tor and proxy which we can hop through to conceal our identity.

It changed their perception about security. Security is not just to secure your network or infrastructure. It's also associated with what we were sharing on the internet. We won't discuss the after effects of the demo but as expected it resulted in a positive outcome.

CASE STUDY 3: AN EPIC INTERVIEW

One of our friends was looking for a change in security domain. One fine morning he got a call from a resource consulting firm about an opening in one of the leading security companies. The first discussion went well as everything was in his favor. The salary expectation, location, and the notice period everything but the only problem was that the profile was a bit different. Our friend was into pentesting earlier but this profile consists a bit of pentesting and more on security monitoring. He got some personal experience based on his own interest on this domain but was not fully into it or in other words he was not having any hands on experience in this. So he was a bit worried about the interview. He wanted to crack this one badly. After two days he again got a call from the same firm stating that you were selected for the technical round and could expect a call from Mr. John Doe within a day or so.

The first thing he discussed with us was, "I just want to get a bit detail about this person so that it will help me understand what kind of questions he may ask, understand his experience and hence expectation." We three decided to jump into OSINT and collect as much as information about Mr. John Doe.

- Problem 1: We just have his name but not his position details.
- Problem 2: Time frame is not fixed, we don't know exactly when to expect the call.

So we suggested him to visit the website of the company to understand the services offered and go through other details such as vision and all. Meanwhile we started digging more information about the person Mr. John Doe.

Step 1:

As we had the name of the person and company name we directly searched the person in LinkedIn. We found his profile and the profile consists of many information such as his current and previous work experience. We found that the person is one of the technical leads of that company. The LinkedIn profile also consists of some of his articles and latest achievements. The person recently got OSCP (Offensive Security Certified Professional). We also found the GitHub account link from the LinkedIn profile. We visited each of his articles, and most of the articles were about how he found some of the bugs in many major sites which consist of some of the zero-days in popular CMS system.

Step 2:

After getting these we visited his GitHub account. He wrote all his scripts to automate the testing process in Python.

Step 3:

We did a simple Google search on his name and got many links along with a slideshare account. We visited that slideshare account. There were some presentations on how to write your own IDS rules. In one of the older posts we found a comment link to his older blogs.

Step 4:

We visited that old blog of him which consists of different road trips he did with his bullet motorcycle.

Step 5:

We searched his Twitter account and found an interesting post that he was recently attending one of the popular security conferences in Goa, India.

Step 6:

We visited the conference site and found that the person Mr. John Doe had given a talk on network monitoring.

Step 7:

We searched him in Facebook and got information about his hometown, current location, educational details, and all.

Step 8:

A quick people search on Yasni provided us a link to another website of a local security community where we found his phone number as he was the chapter leader. We verified this phone number through Truecaller and it checked out right.

It almost took 25–30 min to which we stopped digging more. In the meanwhile our friend was ready with all the information he got from the company website. Based on the information we collected in different sources we concluded this.

- Save the phone number and greet him with his name in case he calls.
- First thing you need to ask him is that how his talk went in Goa.
- Read his talk abstract and tell him it was great and you are regretting that you missed the talk.
- Expect questions on IDS rule writing and can refer to the slideshare presentations for answer.
- Expect some questions on tool automation and that too in Python. So a quick Python revision was required.
- Expect some questions on network penetration testing as he recently did OSCP.
- Expect some questions on web application security, bug bounties, and zero-days as he got listed in many.
- If he asks you about hobbies, tell him your road trips and how you wanted to have a bullet motorcycle but haven't got it yet.
- If you get any question related to your vision and all tell him something related to the company's existing vision aligned with your personal thoughts.
- If he asks you where you see yourself after some years or future plans, tell him you want to go for OSCP certification and be a Red team leader. This was true anyway.

Our friend had a very good knowledge and experience in pentesting and because of his expertise and with a little homework on the company and on the background of the person, he got selected. Mr. John Doe was not only happy with his technical skills but also for the reason that he and our friend got many things in common.

So these were some interesting case studies, we have certainly added as well as subtracted some points here and there as required but all in all these are the kind of situations everyone faces. Let's learn about some basic types of information related to commonly encountered entities and how to deal with them.

It's quite easy to start with primary information such as name, e-mail id to collect all other information but there are cases where we might not have primary information but that does not mean we cannot get these primary information from secondary information. The process might be a bit difficult but it's possible. So now we will discuss in particular about a person's details. What can be collected about a person? Where and how?

Below are some of the information we might be interested to collect about a person.

PERSON:

- First Name
- Last Name
- Company Name
- E-mail Address (Personal)
- E-mail Address (Company)
- Phone Number (Personal)
- Phone Number (Company)
- Address (Home)
- Address (Company)
- Facebook Account URL
- LinkedIn Account URL
- Twitter Account URL
- Flickr Account URL
- Personal Blog/Website URL
- Keywords
- Miscellaneous

From above list we can start with any point and can gather most of the rest. The steps may differ from what we got as a source and how to get to others one by one but we will be using the same tools/techniques just in different order.

Whatever we take as a source, basically we need to start with simple Google search or any other popular traditional search engine search such as Yandex. If we get any relative information, use the same information to collect any of the other related information by treating it as the source.

Let's say we got simple name that contains first name and last name then we can simply use a Google query to get results. Let's say using Google we were able to get the personal blog or website.

Visit that site to search related information such as any details about the person it may be the area of interest, age, date of birth, e-mail, hometown, educational details, or any such information that can be used to get other details.

Let's say we got the educational details. Open Facebook and try to search for the name with educational details. We may get the person's profile. In Facebook we will get lots of information such as company he/she is working on, friends, pictures of him/her, and sometime other profile links along with the personal e-mail address also.

Now using the company name and person's name we can get the LinkedIn profile quite easily and can craft an e-mail address. Generally companies use a typical pattern to create e-mail addresses. Let's say the company name is abc.inc and the site is www.abc.com, and it uses the pattern, first letter of the first name and the last name without any spaces. So from the person's name and company name we can easily craft the e-mail address or we can use tool like harvester to harvest e-mail address from the company domain name and after looking at all the e-mails we can easily pick the e-mail associated with the person. In this way it is possible to get or collect information through correlation.

Collecting company details is quite easy as compared to the personal information. Most of the company information is public. So we can get it in the website itself and we can start collecting company information just by knowing the company name. It can be used as one point of source to get all the information. So if we know the company name we can get rest of the details quite easily.

Below are the list of information we generally look out for in terms of a company detail.

COMPANY:

- Company Name
- Year of Establishment
- Directors
- Website
- Registrant Name
- Phone Number
- Address
- Keywords
- Number of Employees
- Employee E-mail Samples
- HR E-mail
- Openings
- Facebook Account URL
- LinkedIn Account URL
- Twitter Account URL
- Flickr Account URL
- Other Blog/Website URL/Subdomains
- Miscellaneous

If we know the company name we can easily go to any search engine we want to get its registered website URL. Once we get the registered website, we can get information such as year of establishment, directors, phone number, address, HR e-mail, openings details, and, in some case, links to different social networking profiles. A simple Whois query will also result in a lot of information.

There are other places where we can get all these information. We can see these information on Glassdoor, Zoominfo, and LinkedIn. For openings we can look in job portals as well. The number of employee and employee samples can be easily found using LinkedIn or else we can simply use tool such as harvester, Maltego to get e-mail patterns. The keywords and all can be easily found in the meta part of the website source, through various SEO and SEM services such as SEMRush.

This part is a bit technical and frequently needed by technical people such as administrators, IT consultants, pentesters etc. But if you have followed the book till here even after not being a tech person, then you won't find it much difficult. In this case, for carrying out a technical audit or assessment basically analysts are provided with the domain name or the IP address, so from OSINT point of view we can consider these two information as primary information. From these or any of it, it is quite possible to get the rest of the information mentioned below.

DOMAIN:

- Domain
- IP Address
- Name Server
- MX server
- Person
- Website
- Subdomains
- E-mail Samples
- Files
- Miscellaneous

So as we discussed earlier let's take the domain as a primary entity and from that we want to get all other information that is mentioned above. If we wanted to simply get the IP address of that domain, we just need to run a simple ping command in command prompt or terminal based on the operating system we use.

```
ping <domain name>
```

This command will execute and will provide the IP address of the domain.

For other domain-specific information, there are domain tools freely available in the internet and from the Whois record, we will get different information such as registered company name, name server details, registered e-mail ids, IP address, location, and many more. Resources like w3dt.net can be very helpful here. Directly using domain tool we can find lots of information about a domain, or else we can use different domain specific Maltego transformations for the same.

We can also use harvester to collect subdomains, e-mail address etc., from a domain name. From the e-mail addresses we can search for the profiles of the persons in different social networking sites.

And to get subdomains and a particular file from the domain we can use search-Diggity, Google, Knock (Python tool).

To get different subdomains we can use site operator or create a Python script which will take subdomain names from our list and enumerates it along with the domain provided.

```
site:domainname
```

To get a particular type of file from the domain with a keyword we can use filetype or ext operator and can run below query,

```
site:domainname keyword filetype:ppt
```

So in this way we can get all the domain-specific information from different sources.

So these were some case studies and examples in which OSINT can be collected and be helpful in our personal and professional life.

As promised earlier in our next topic we will be learning about Maltego machines.

MALTEGO MACHINES

We have covered various aspects of Maltego in previous chapters from understanding the interface to creating local transforms. As this chapter is more about combining the knowledge we have covered till now, so related to Maltego we will learn how to create Maltego machines. Although we have already defined what a Maltego machine is, yet for quick recall it programmatically connected a set of transforms. It allows us to take one entity type as input and move toward another type(s) which are not directly connected to it, but through a set/sequence of transforms. There are some inbuilt machines in Maltego such as Company Stalker which takes domain entity as input and runs various transform in sequential fashion to get different types of information from it such as e-mail address, files etc.

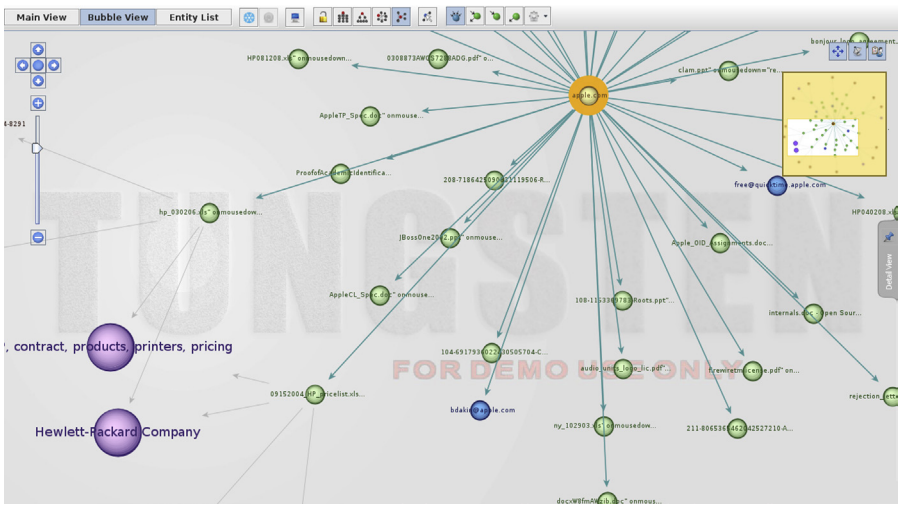

FIGURE 14.1

Maltego "Company Stalker" Machine

To create our own machine we need to use Maltego Scripting Language (MSL). The documentation for MSL is available as a PDF at http://www.paterva.com/MSL.pdf. The documentation is clear and simple, and anyone having basic programming skills can easily understand it. As all the terms and process are clearly described we do not need to cover them again, so straight away jump to create our own simple machine using local transforms we learned to create in a previous chapter.

Creating a Maltego machine is pretty simple, first we need to go to the Machines tab, under which we can find the New Machine option. Clicking on it will bring a window where we need to provide the name and other descriptive details related

to the machine we are going to create. In the next step we need to choose type of machine we are going to create. For this we have three options:

- Macro: runs once
- Timer: runs periodically until stopped
- Blank: a blank template

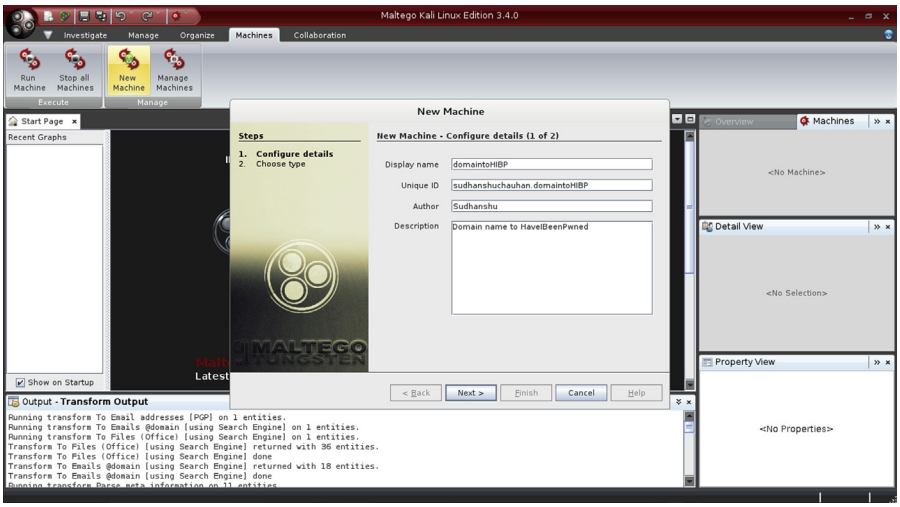

FIGURE 14.2

Create Maltego Machine

Once we have selected the machine type we can write the code for our machine and include transforms in it at the appropriate position, from the right-hand side panel by selecting the transform and double clicking on it. The "start" block contains the transforms and all other execution part. The "run" functions are used to execute a specific transform. To run functions in a parallel fashion we can include them inside the "paths." Inside "paths" we can create different "path" which will run in parallel with each other but the operations inside a path will run sequentially. Similarly we can provide different values, take user inputs, use filters etc.

Let's create a simple machine which extracts e-mail ids from a provided domain and further runs our HIBP local transform on these. For this we need to provide the machine name and select the macro type machine. Next we need to include the inbuilt transforms which can extract e-mails from domain such as domain to e-mail using search engine, Whois etc. Next we need to include our local HIBP transform. As we need to run these in parallel we need to create separate "path" for each e-mail extraction transform. Our final code looks like this:

```
machine("sudhanshuchauhan.domaintoHIBP",
  displayName:"domaintoHIBP",
  author:"Sudhanshu",
```

```
 description: "Domain name to HaveIBeenPwned") {

start {

 paths{
  path{
   run("paterva.v2.DomainToEmailAddress_AtDomain_SE")
   run("sudhanshuchauhan.emailhibp")
      }
  path{
   run("paterva.v2.DomainToEmailAddress_SE")
   run("sudhanshuchauhan.emailhibp")
      }
  path{
   run("paterva.v2.DomainToEmailAddress_Whois")
   run("sudhanshuchauhan.emailhibp")
      }
  path{
   run("paterva.v2.DomainToEmailAddress_PGP")
   run("sudhanshuchauhan.emailhibp")
      }
     }
 }
}
```

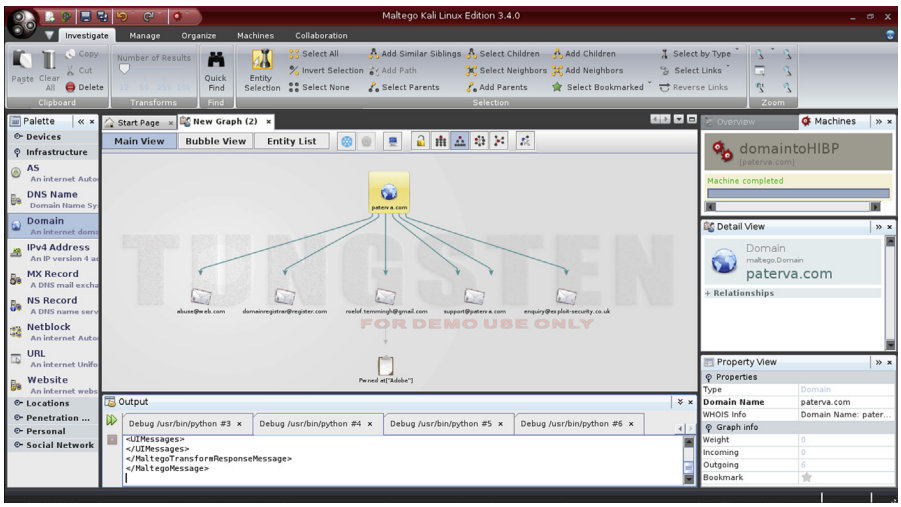

FIGURE 14.3

Our Maltego machine output.

Two important things that need to be kept in mind are that our local transform must be integrated into Maltego before creating the machine and the input and output data types need to be taken care of when creating a sequence.

So we learned to create Maltego machine. Though there is still much more to explore and learn related to Maltego, we have attempted to touch upon its every important aspect.

In this chapter we have learned about combining all the knowledge we gained till now and also saw some practical scenarios and examples. This is important as in real-life projects. It's not just about knowing things but also about implementing and utilizing them in an integrated manner according to the situation and generating a fruitful outcome.

In our next and last chapter we will be learning about certain general topics related to the internet which are often connected directly or indirectly to the information gathering. Having a basic understanding of these terms will be helpful for anyone utilizing internet for investigative purpose.