CS 455 – Computer Security Fundamentals

Dr. Chen-Yeou (Charles) Yu

- Dive into Footprinting
  - Information collection using OSINT
  - Ride the spiderfoot to collect information!

Information collection using OSINT

- So far, I didn’t tell you how to use defensive tools or offensive in Kali
- At least, by the end of this Lecture1, we can do a better job in the

Footprinting by using the technology --- OSINT

- The OSINT is a way to collect information more efficiently.
- Open-Source Intelligence (OSINT) is the “collection” and “analysis (Maybe? sometimes you need to pay extra $)”, of data gathered from (many) open sources.
- It could be an online database or just a huge file.
- The good thing is. This kind of database updates quickly!
  - For example, if the backdoor is removed from a machine, this machine or its related ID will be removed from the blacklist.

Information collection using OSINT

- OSINT is like a kind of crowd-sourcing technology. This proved to be an effective technology to fight against malicious attacks.
- For example, there is one online database called
  - http://www.blocklist.de/
- They kept updating the database about blocked IP(s) in different categories: ssh, mail, apache, imap, sip, bot, strongips, ircbot, bruteforcelogin.
- You can download these database separately.
  - http://www.blocklist.de/en/export.html

Ride the spiderfoot to collect information!

- Unlike “nmap”, you need to scan the host by yourself ^_^
- spiderfoot is to check the online databases directly
- The good thing is, it has SO MANY built-in databases.
  - But some of them, you need to pay the $ and get the key.
  - However, some of the key are free. But some are not.
- Application → 01-Information Gathering → OSINT Analysis → spiderfoot
- spiderfoot-cli is about the python implementation. We can just bypass that this time.

Ride the spiderfoot to collect information!

- The version is now v4.0.
- This is the picture from the internet for v3.5

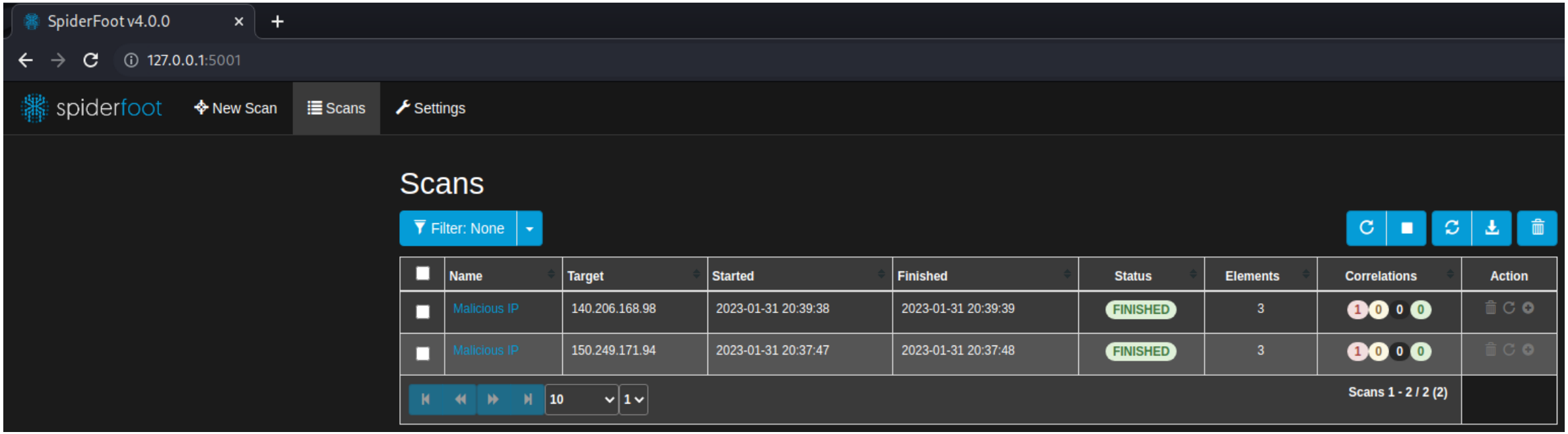
```
┌─[root@kali ~]
└──── www.bbskali.cn # spiderfoot -h
usage: sf.py [-h] [-d] [-l IP:port] [-m mod1,mod2,...] [-M] [-s TARGET]
             [-t type1,type2,...] [-T] [-o tab|csv|json] [-H] [-n] [-r]
             [-S LENGTH] [-D DELIMITER] [-f] [-F type1,type2,...] [-x] [-q] [-V]

SpiderFoot 3.5.0: Open Source Intelligence Automation.

optional arguments:
  -h, --help          show this help message and exit
  -d, --debug         Enable debug output.
  -l IP:port          IP and port to listen on.
  -m mod1,mod2,...    Modules to enable.
  -M, --modules       List available modules.
  -s TARGET           Target for the scan.
  -t type1,type2,...  Event types to collect (modules selected automatically).
  -T, --types         List available event types.
  -o tab|csv|json     Output format. Tab is default.
  -H                  Don't print field headers, just data.
  -n                  Strip newlines from data.
  -r                  Include the source data field in tab/csv output.
  -S LENGTH           Maximum data length to display. By default, all data is
                      shown.
  -D DELIMITER        Delimiter to use for CSV output. Default is ,.
  -f                  Filter out other event types that weren't requested with
                      -t.
  -F type1,type2,...  Show only a set of event types, comma-separated.
  -x                  STRICT MODE. Will only enable modules that can directly
                      consume your target, and if -t was specified only those
                      events will be consumed by modules. This overrides -t and
                      -m options.
  -q                  Disable logging. This will also hide errors!
  -V, --version       Display the version of SpiderFoot and exit.
```

Ride the spiderfoot to collect information!

- You can directly type the command in the Kali, but it is not fun. Trust me!
- You can use the spiderfoot –h to show the help menu.
- Or you can use this command in the terminal to launch the web management interface first.
  - spiderfoot-l 127.0.0.1:5001
- Then? All you need to do is just to open your Chrome browser by typing this in the address line.
  - 127.0.0.1:5001

Ride the spiderfoot to collect information!

- One you get into the spiderfoot (of the web), you will see this screen



- Those are my scanned history. It could be totally empty when first time you are entering this page.

Ride the spiderfoot to collect information!

- The 2 x IP addresses are specifically picked up by me from the file, “blocklist.de.all.IP.listl.1.31.2023.txt”
- This file is actually from the “blocklist.de” on 1.31.2023
- The reason I want to pickup an IP address and choose the only one database is because --- I want the scanning to be as fast as possible!
- Performing a full scan, it might spend a lot of time in visiting different databases, but the report generated by the spiderfoot is fancy and is useful

Ride the spiderfoot to collect information!

- In the “New Scan” tab
  - The Scan Name is just the name of the scan you want to put in this time, it can be any. After the execution, the name will be stored into the “Scans” tab
  - Scan target could be many
    - Domain name
    - IP Address
    - Someone’s email
    - Someone’s name
    - Bitcoin Address!? I don’t know what is this!??
  - (check the next page for detail)

Ride the spiderfoot to collect information!

New Scan

Scan Name

The name of this scan.

Scan Target

The target of your scan.

Your scan target may be one of the following. SpiderFoot will automatically detect the target type based on the format of your input:

Domain Name: e.g. *example.com*
IPv4 Address: e.g. *1.2.3.4*
IPv6 Address: e.g. *2606:4700:4700::1111*
Hostname/Sub-domain: e.g. *abc.example.com*
Subnet: e.g. *1.2.3.0/24*
Bitcoin Address: e.g. 1HesYJSP1QqcyPEjnQ9vzBL1wujruNGe7R
E-mail address: e.g. *bob@example.com*
Phone Number: e.g. *+12345678901* (E.164 format)
Human Name: e.g. *"John Smith"* (must be in quotes)
Username: e.g. *"jsmith2000"* (must be in quotes)
Network ASN: e.g. *1234*

By Use Case | By Required Data | By Module

All — **Get anything and everything about the target.**

All SpiderFoot modules will be enabled (slow) but every possible piece of information about the target will be obtained and analysed.

Footprint — **Understand what information this target exposes to the Internet.**

Gain an understanding about the target's network perimeter, associated identities and other information that is obtained through a lot of web crawling and search engine use.

Investigate — **Best for when you suspect the target to be malicious but need more information.**

Some basic footprinting will be performed in addition to querying of blacklists and other sources that may have information about your target's maliciousness.

Passive — **When you don't want the target to even suspect they are being investigated.**

As much information will be gathered without touching the target or their affiliates, therefore only modules that do not touch the target will be enabled.

Run Scan Now

Ride the spiderfoot to collect information!

- Scan By Use Case
  - As the description in the sentence, easy to understand
  - But ALL of them are slow!
  - We better to know what we really want
- Scan By Required Data
  - A very lengthy list but is more human-readable
  - Check the next page
- Scan By Module
  - If you are very “familiar” with these databases, you can do that
  - Databases are from A to Z

Ride the spiderfoot to collect information!

Scan By Required Data

By Use Case | By Required Data | By Module

Select All | De-Select All

Account on External Site	Affiliate - Company Name
Affiliate - Domain Name	Affiliate - Domain Name Unregistered
Affiliate - Domain Whois	Affiliate - Email Address
Affiliate - IP Address	Affiliate - IPv6 Address
Affiliate - Internet Name	Affiliate - Internet Name - Unresolved
Affiliate - Internet Name Hijackable	Affiliate - Web Content
Affiliate Description - Abstract	Affiliate Description - Category
App Store Entry	BGP AS Membership
BGP AS Ownership	Base64-encoded Data
Bitcoin Address	Bitcoin Balance
Blacklisted Affiliate IP Address	Blacklisted Affiliate Internet Name
Blacklisted Co-Hosted Site	Blacklisted IP Address
Blacklisted IP on Owned Netblock	Blacklisted IP on Same Subnet
Blacklisted Internet Name	Cloud Storage Bucket
Cloud Storage Bucket Open	Co-Hosted Site
Co-Hosted Site - Domain Name	Co-Hosted Site - Domain Whois
Company Name	Compromised Password
Compromised Password Hash	Cookies
Country Name	Credit Card Number
DNS SPF Record	DNS SRV Record
DNS TXT Record	Darknet Mention URL
Darknet Mention Web Content	Date of Birth

Ride the spiderfoot to collect information!

Scan By Module:

If you see a “**lock**” in front of the database, that means, it needs a key

Sometimes, it needs $, but sometimes, it just a registration for free.

By Use Case | By Required Data | By Module | Select All | De-Select All

AbstractAPI	Look up domain, phone and IP address information from AbstractAPI.
abuse.ch	Check if a host/domain, IP address or netblock is malicious according to Abuse.ch.
AbuseIPDB	Check if an IP address is malicious according to AbuseIPDB.com blacklist.
Abusix Mail Intelligence	Check if a netblock or IP address is in the Abusix Mail Intelligence blacklist.
Account Finder	Look for possible associated accounts on nearly 200 websites like Ebay, Slashdot, reddit, etc.
AdBlock Check	Check if linked pages would be blocked by AdBlock Plus.
AdGuard DNS	Check if a host would be blocked by AdGuard DNS.
Ahmia	Search Tor 'Ahmia' search engine for mentions of the target.
AlienVault IP Reputation	Check if an IP or netblock is malicious according to the AlienVault IP Reputation database.
AlienVault OTX	Obtain information from AlienVault Open Threat Exchange (OTX)
Amazon S3 Bucket Finder	Search for potential Amazon S3 buckets associated with the target and attempt to list their contents.
Apple iTunes	Search Apple iTunes for mobile apps.
Archive.org	Identifies historic versions of interesting files/pages from the Wayback Machine.
ARIN	Queries ARIN registry for contact information.

Ride the spiderfoot to collect information!

- For example, the AbuseIPDB in the “Settings”

Settings

Settings

Save Changes | Import API Keys | Export API Keys | Reset to Factory Default

Global
Storage
AbstractAPI
abuse.ch
AbuseIPDB
Abusix Mail Intelligence
Account Finder
AdBlock Check
Ahmia
AlienVault OTX
AlienVault IP Reputation
Archive.org
Azure Blob Finder
Bad Packets
Base64 Decoder
BinaryEdge

AbuseIPDB (sfp_abuseipdb)

Summary	Check if an IP address is malicious according to AbuseIPDB.com blacklist. AbuseIPDB is a project dedicated to helping combat the spread of hackers,spammers, and abusive activity on the internet. Our mission is to help make Web safer by providing a central blacklist forwebmasters, system administrators, and other interested parties toreport and find IP addresses that have been associated with malicious activity online.
Categories:	Reputation Systems
Tags:	apikey
Website:	https://www.abuseipdb.com

Settings

Option	Value
AbuseIPDB.com API key.	
Apply checks to affiliates?	True
The minimium AbuseIPDB confidence level to require.	90
Maximum number of results to retrieve.	10000

Ride the spiderfoot to collect information!

- I had attached this file “blocklist.de.all.IP.listl.1.31.2023.txt” onto the blackboard but is only for “blocklist.de”
- Sometimes, try to search your personal information in this page!
  - Email, name, even phone number
  - In case your info. is collected and is misused by someone!

Ride the spiderfoot to collect information!

- If you choose the “All scan”, it will give you a list of something found in different categories (data types). An example from internet.

jack RUNNING

Summary | Browse | Graph | Scan Settings | Log

Search...

Type	Unique Data Elements	Total Data Elements	Last Data Element
Affiliate - Domain Name	3	9	2022-01-18 14:43:42
Affiliate - Email Address	137	143	2022-01-18 14:41:41
Affiliate - IP Address	7	7	2022-01-18 14:41:11
Affiliate - IPv6 Address	4	4	2022-01-18 14:41:11
Affiliate - Internet Name	9	16	2022-01-18 14:44:35
BGP AS Membership	1	1	2022-01-18 14:32:07
Blacklisted Affiliate IP Address	1	1	2022-01-18 14:40:14
Blacklisted Affiliate Internet Name	1	1	2022-01-18 14:38:51
Blacklisted Co-Hosted Site	1	1	2022-01-18 14:35:18
Blacklisted Internet Name	3	3	2022-01-18 14:31:57
Co-Hosted Site	3	3	2022-01-18 14:30:01
Co-Hosted Site - Domain Name	3	3	2022-01-18 14:38:33

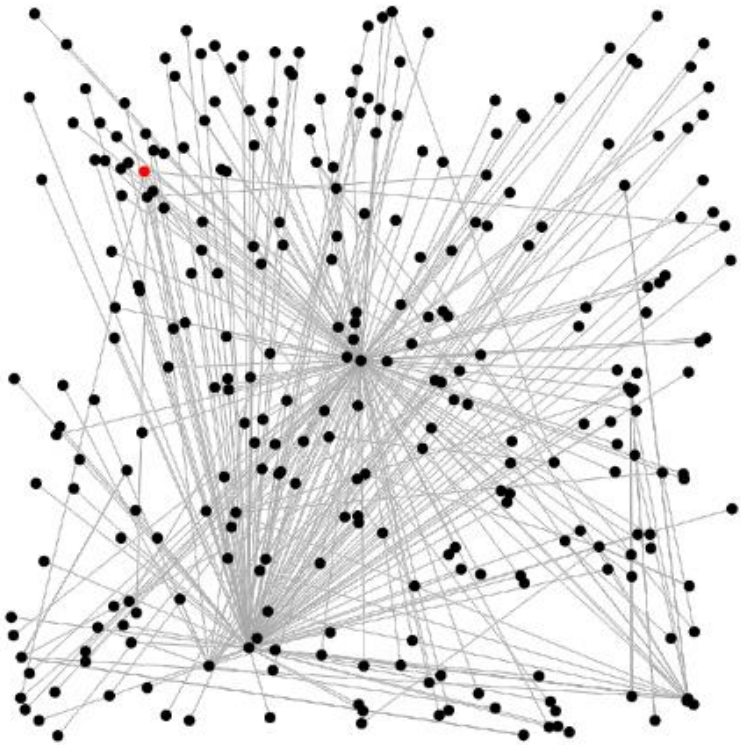
Ride the spiderfoot to collect information!

- Or, the spiderfoot can gererate the “spiderweb” --- the network topology. This graph is from the internet
- The red dot is the target

you are investigating

- Black dots and lines are

the relationships with the target



Ride the spiderfoot to collect information!

- If you know the spiderfoot very well, it is actually implemented by Python.
- Here is an example of the query about **threat info.** and the info about the **blacklist**
- “-s” means the specifying the target. And for all the databases begins with “sfp”, you can easily find the mapping databases in the Scan “By Module” page
- The outputs of the IP Addresses might not be consecutive because not all the IP Addresses are **blacklisted**
- (check the next page for detail)

Ride the spiderfoot to collect information!

```
steve@hxdev:~/spiderfoot$ python3 ./sf.py -m sfp_sorbs,sfp_spamcop,sfp_abusech,sfp_alienvault,sfp_malwarepatrol,sfp_isc -s 93.189.42.146
 -q
Source                          Type                                      Data
SpiderFoot UI                   IP Address                                93.189.42.146
sfp_sorbs                       Blacklisted IP Address                    SORBS - Spammer (93.189.42.146)
steve@hxdev:~/spiderfoot$
steve@hxdev:~/spiderfoot$ # OK, so it's considered malicious by SORBS. We can also do
steve@hxdev:~/spiderfoot$ # the same with a whole subnet, just expect it to take a little longer..
steve@hxdev:~/spiderfoot$
steve@hxdev:~/spiderfoot$ python3 ./sf.py -m sfp_sorbs,sfp_spamcop,sfp_abusech,sfp_alienvault,sfp_malwarepatrol,sfp_isc -s 93.189.42.0/2
4 -q
Source                          Type                                      Data
SpiderFoot UI                   Netblock Ownership                        93.189.42.0/24
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.9)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.10)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.11)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.13)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.22)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Recent Spammer (93.189.42.22)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.31)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.34)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.38)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Recent Spammer (93.189.42.38)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.40)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.41)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.43)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.45)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.54)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.55)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.58)
sfp_sorbs                       Blacklisted IP on Owned Netblock          SORBS - Spammer (93.189.42.76)
```

Ride the spiderfoot to collect information!

- Finally, the info. collected by the spiderfoot might be “wrong”. You need to be very careful to verify everything, if there are anomalies in the scanned outputs.
- Be patient as a cybersecurity detective ☺