KNOCK? KNOCK? WHOIS THERE?
APT ATTRIBUTION AND DNS PROFILING

Frankie Li

ranerwang@gmail.com

Twitter: @espionageware

AGENDA

- APT Attribution: Who wrote these codes?
- Tactics, Techniques and Procedures (TTP)
- Behavior of APT adversary
- HUMINT extracted from DNS
- Gather intelligence from open source (aka OSINT)
- Dynamically monitoring of PassiveDNS ➔ PassiveWhois
- Analysis by visualization tool (Maltego)
- Tools and demo

WHO AM I?

- From a place in China, but not so China ;)
- Sunday researcher in malware analysis and digital forensics
- Part time lecturer
- A Lazy blogger (espionageware.blogspot.com)
- NOT associated with PLA 61398 or Mandiant
- NOT associated with PLA 61486 or CrowdStrike or Taia Global or ThreatConnect

APT ATTRIBUTION

APT ATTRIBUTION

- Disclaimer: Not going to provide any opinion on the latest indictment or 奶黄包 or 楼主 or 上海钟楼
- http://espionageware.blogspot.com or Twitter: @espionageware
- Not a concern for private sector, but for LE or intelligence agencies
- Not difficult, if you have source code
- Not hard, if you focus only on strings & human readable data within a malware program
- But, to attribute responsibility with “Certainty” is almost impossible, unless they make a mistake

WHO WROTE THESE CODES?

- Source code attribution
- Attributes of Windows binaries
- Attribution malware
- Attribution of APT by digital DNA

SOURCE CODE ATTRIBUTION

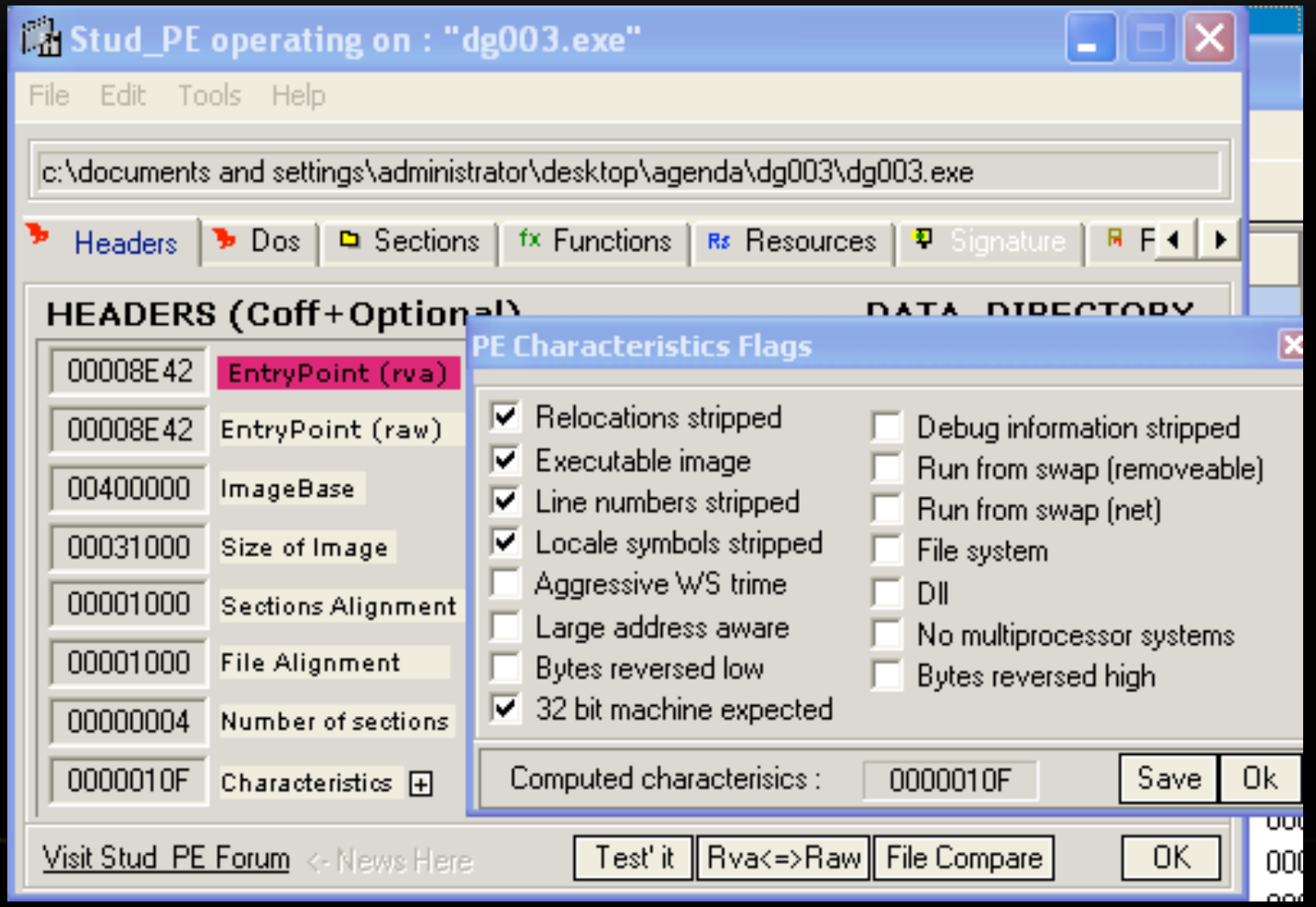
- Stylometry, the application of attribute the authorship by coding style
- Kind of profiling by writing style
- Comments and coding crumbs
- JStylo: By comparing unknowns documents with a known candidate author's document*
- Not a solution because most APT samples collected are compiled binaries

*Islam, A. (2013). Poster: Source Code Authorship Attribution

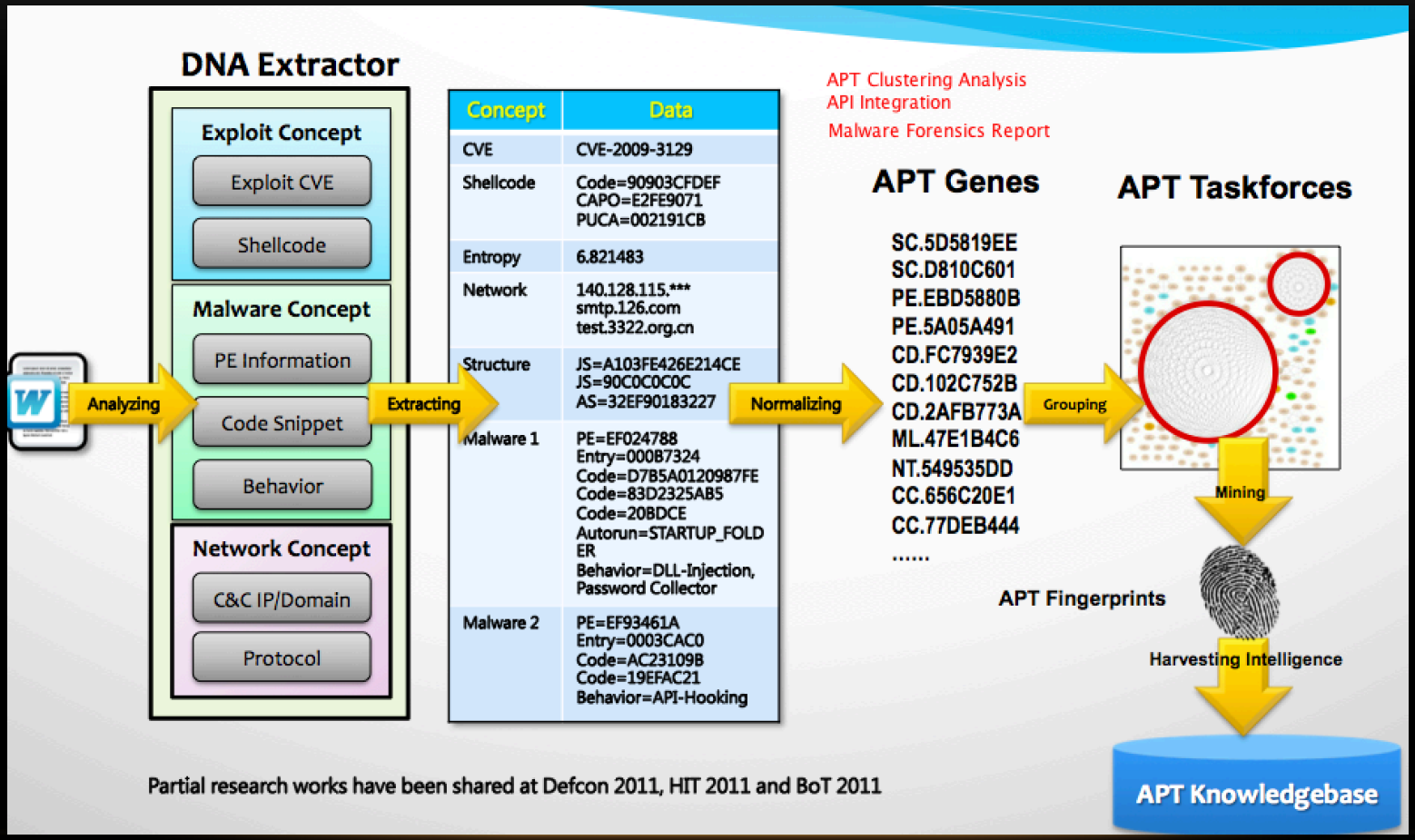
ATTRIBUTES OF WINDOWS MALWARE

- PE headers are des-constructed and metadata (artifacts) are categorized (Yonts, 2012)
- Extract the technical and contextual attributes or “genes” from different “layers” to group the malware (Xecure-Lab, 2012 and Pfeffer, 2012)
- By a proprietary reverse engineering and behavioral analysis technology (Digital DNA, 2014)

PE DECONSTRUCTION



ATTRIBUTION USING GENETIC INFORMATION



From: Xecure-Lab, 2012

IDENTIFIED APT GROUPS?

- Sensational names created for APT actors:
  - (09) GhostNet
  - (10) Operation Aurora
  - (11) Lurid, Nitro, Night Dragon, 1.php, Shady RAT
  - (13) Comment Crew/APT1, Soysauce, Deep Panda, Red October, Net Traveler, SAFE …
  - (14) PutterPanda, PittyTiger (probably not a state-sponsored group)

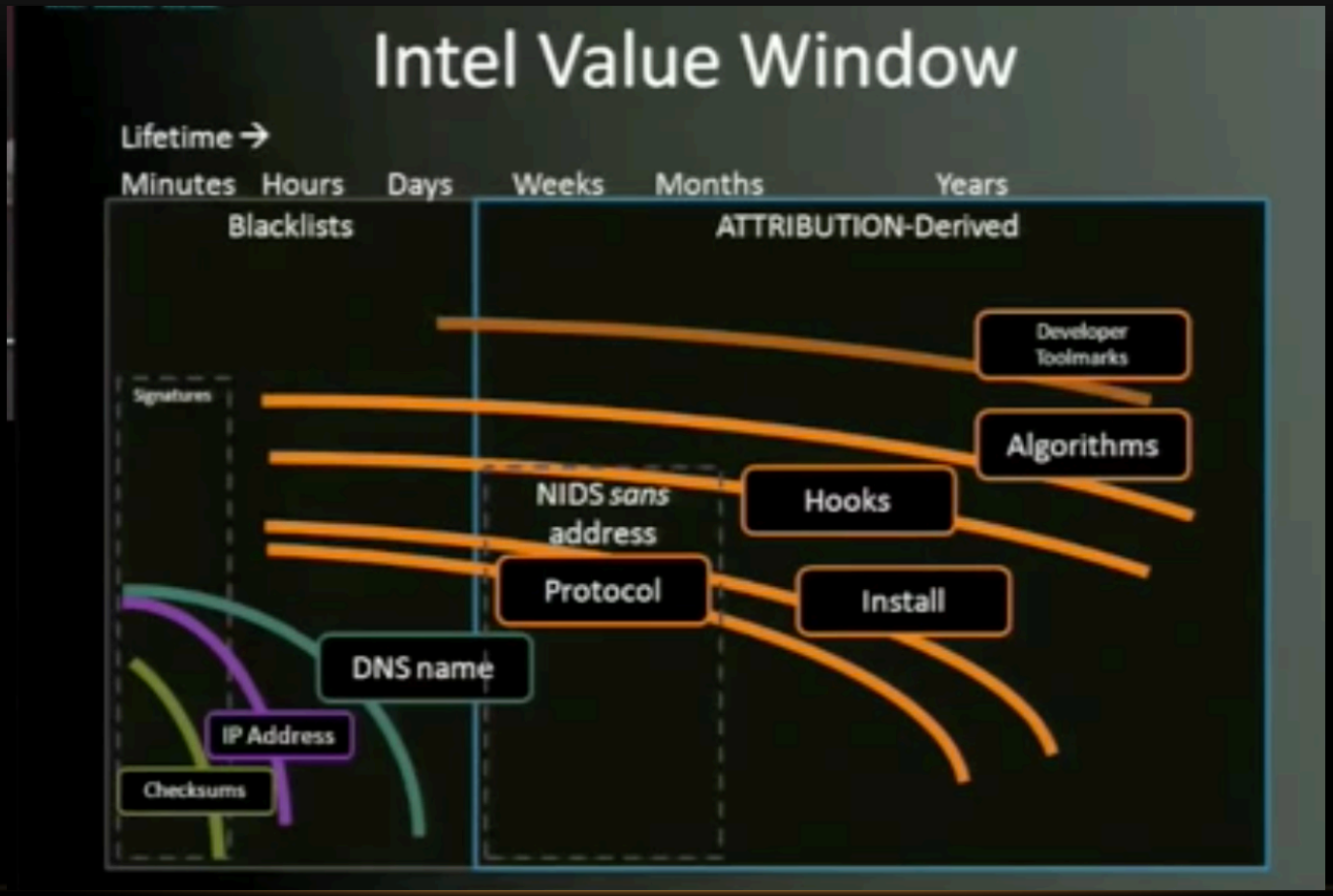
TACTICS, TECHNIQUES AND PROCEDURES (TTP)

HUMAN IS THE KEY

- Attribution: Tracking Cyber Spies & Digital Criminals (Hoglund, 2010)
- Forensics marks that could be extracted from raw data in three intelligence layers
  - Net Recon
  - Developer Fingerprints
  - Tactics, Techniques, and Procedures (TTP)
- Among these three layers, TTP should carry the highest intelligence value for identifying human attackers
- But, near impossibility of finding the human actors with definitive intelligence
  - Social Cyberspace (i.e., DIGINT)
  - Physical Surveillance (i.e., HUMINT)

http://www.youtube.com/watch?v=k4Ry1trQhDk

HOGLUND'S MALWARE INTEL LIFE TIME



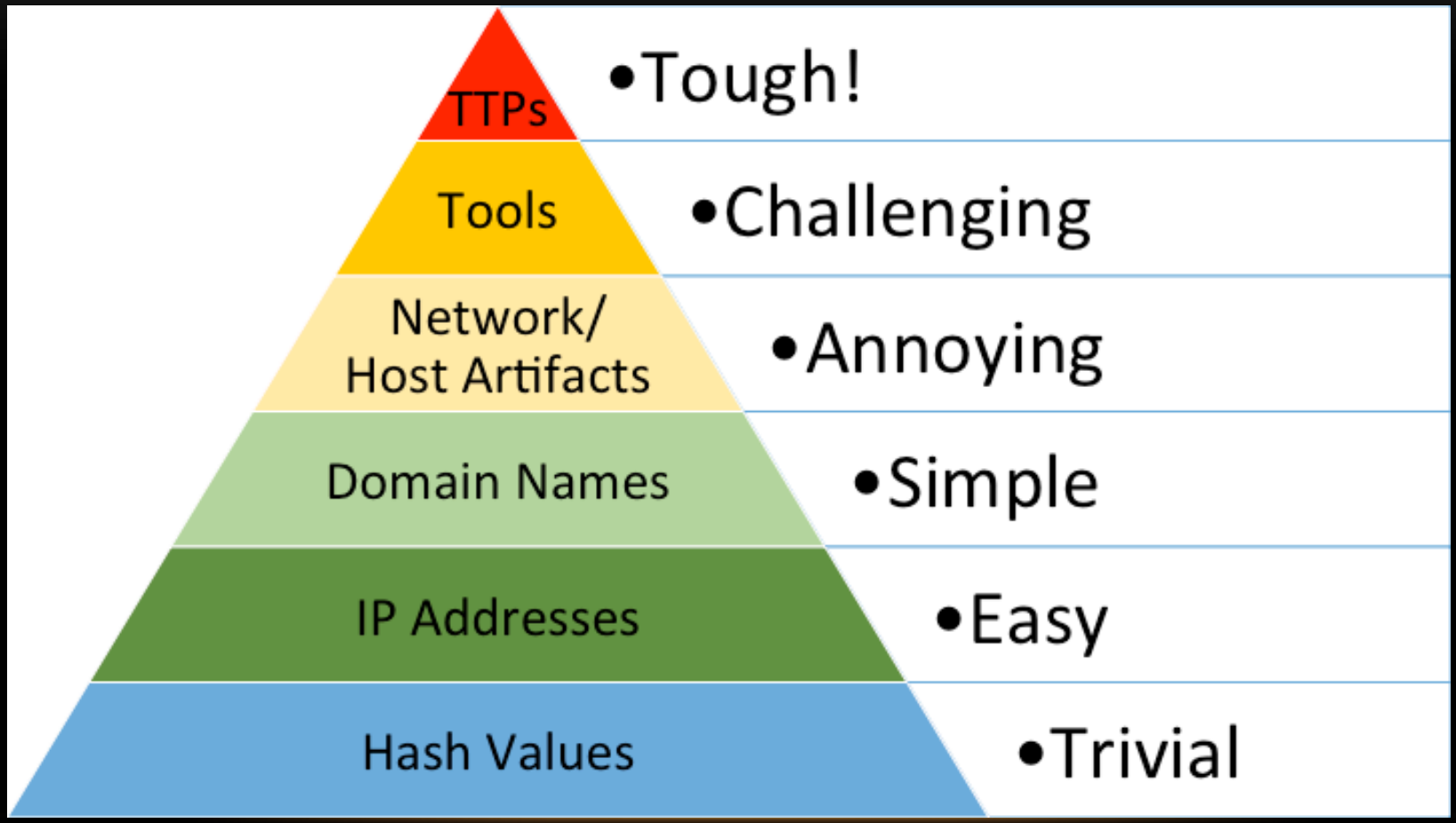
BOMAN'S VXCAGE

- Boman extracts technical metadata from a large collection of binaries
- Store the identified artifacts in a relational database for further analysis
- But still based on technological contexts from malware binaries instead of the behavior of the human working behind

TTP

- A military term?
- A term to describe the behavior of adversary?
- A modern term to replace modus operandi?
  - the method of operation
  - The habits of working
- TTP are human-influenced factors

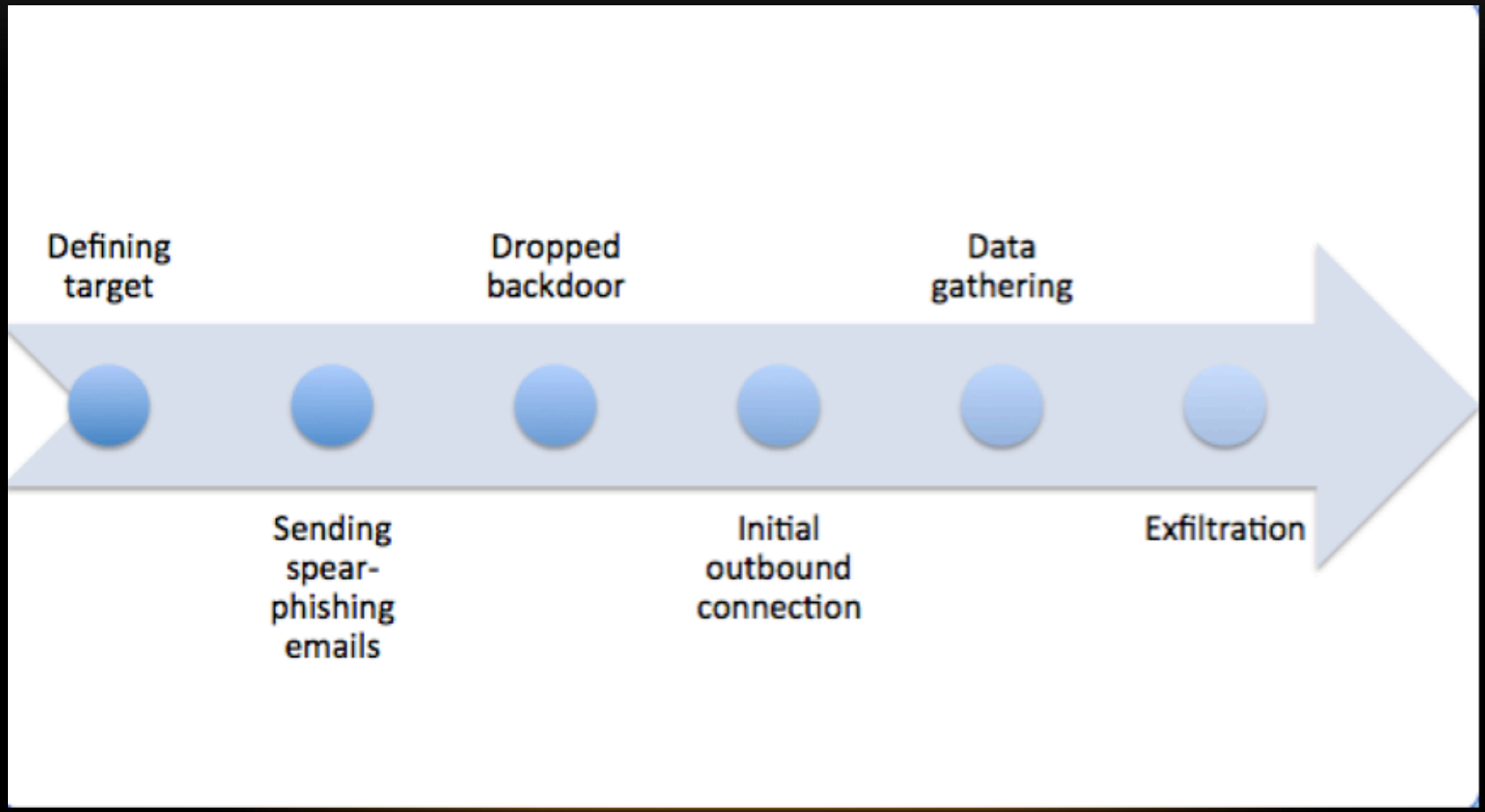
PYRAMID OF PAIN



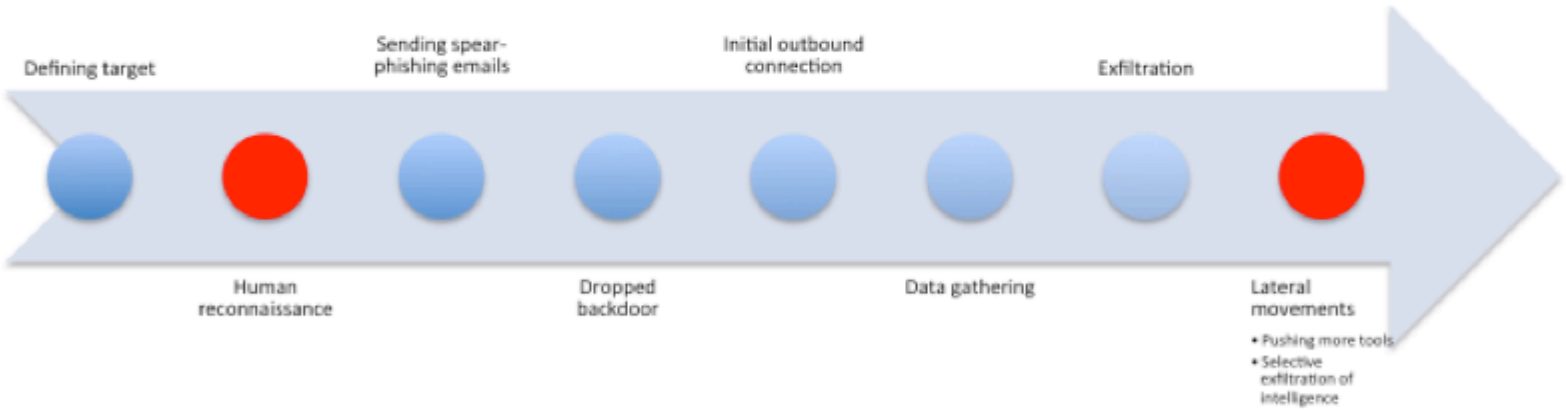
From David Bianco's Blog
http://detect-respond.blogspot.hk/2013/03/the-pyramid-of-pain.html

BEHAVIOR OF APT ADVERSARY

APT LIFE CYCLE (KILL CHAIN)



EXTENDED APT LIFE CYCLE



ASSUMED APT INFRASTRUCTURE TACTICS

- Domain registration
- Naming convention is not typo squatting, but follows a pattern of meaningful Chinese PingYing（拼音）
- Creation DNS-IP address pairs
- Engaging a “friendly ISP” to use a portion of their C-class subnet of IP addresses situated at the domicile of the targeted victims
- DNS names and IP addresses may be cycled for reuse (a.k.a. campaigns), which may provide indications or links to the attacker groups
- Embedding multiple DNS A-records in exploits
- Preparing spear-phishing email content after reconnaissance of the targeted victims
- Launching malicious attachments through spear-phishing emails

ASSUMED APT INFRASTRUCTURE TACTICS-2

- The exploits drop binaries that extract the DNS records and begin communicating with the C2 by resolving the IP addresses from DNS servers.
- The C2 servers or C2 proxies register the infections on the C2 database
- The intelligence analysts of the attacker groups review the preliminary collected information of the targeted victims through C2 portals.
- The infected machines are further instructed to perform exfiltration of collect further intelligence from the infected machines.
- The infrastructure technical persons of the attacker group apply changes (domain manipulation) to the DNS-IP address pair, domain name registration information (Whois information), and the “parked domains” from time to time or when a specific incident occurs
- In contrast with the Fast-Flux Services Networks mentioned by the HoneyNet Project, the information does not change with high frequency

HUMINT EXTRACTED FROM DNS & WHOIS

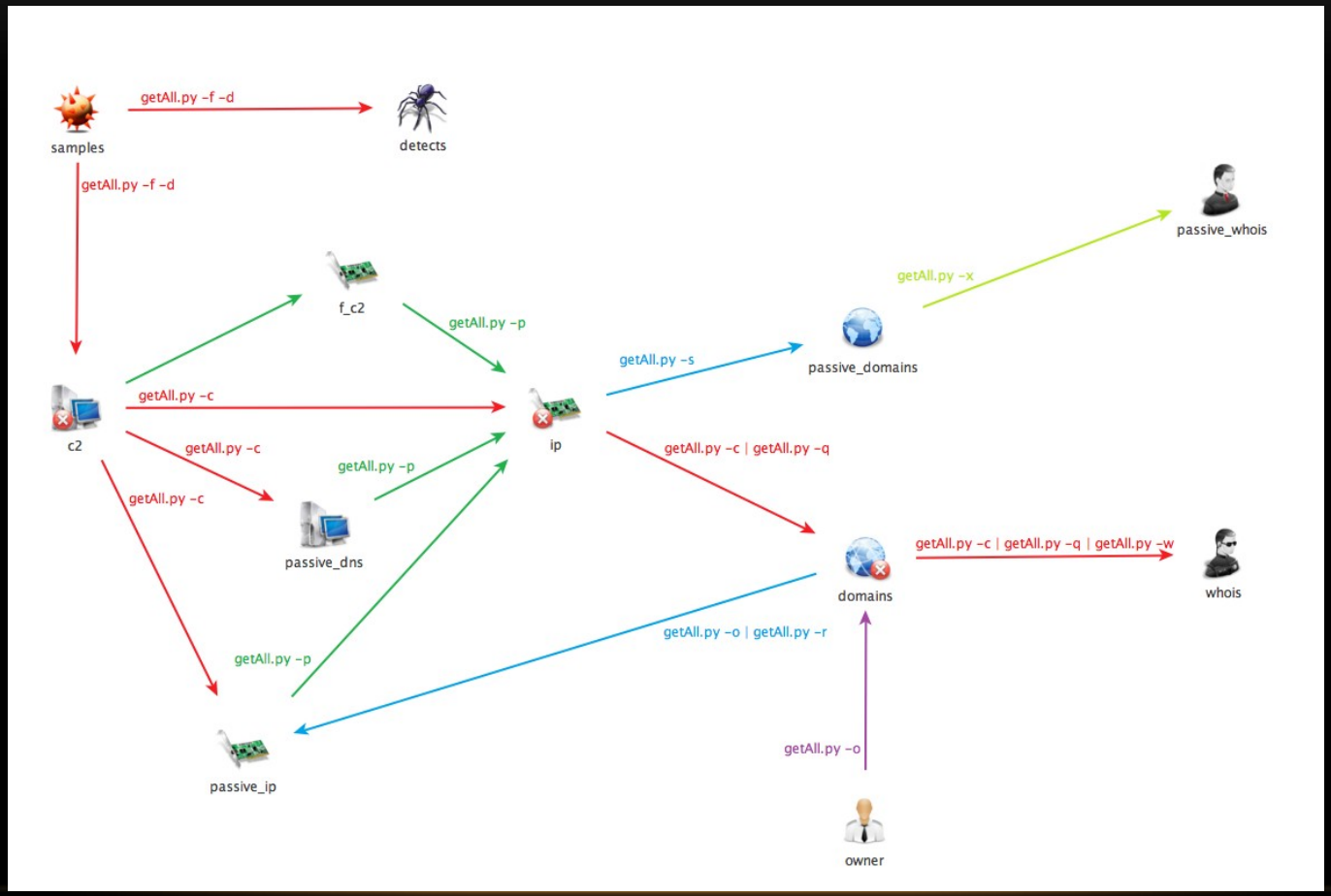
WHAT IS KEPT IN DNS & WHOIS

- Domain names: A Record, Cname, NS record
- Whois records: valid email address (at least once), name, street address, name servers
- Parked-domains: temporary IP address assigned creation of first DNS record on the name server (newly created domains are kept under 1 IP address for future use)

HUMINT INTEL TO BE COLLECTED

- Extract DNS from the malicious code (sandbox)
- Lookup the currently assigned IP address
- Retrieve all parked-domains from the identified IP address
- Retrieve whois information from the identified domains
- Update identified record to a relational database for future analysis
- Repeat the process and record all changes in the database

INTEL COLLECTION PROCESS



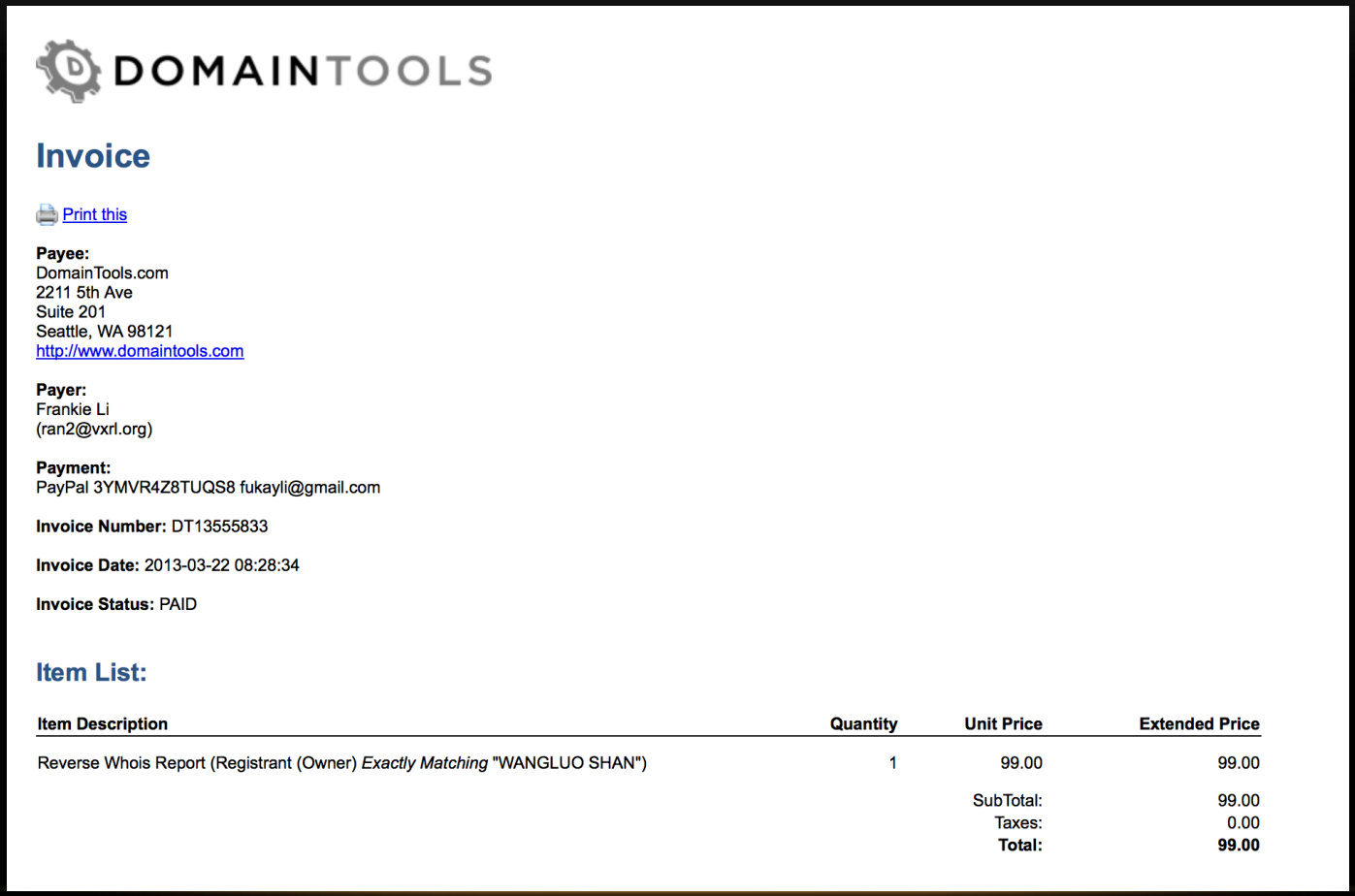
QUERIES FROM OPEN SOURCE OSINT

the only available weapon we have

OSINT

- Nslookup
- Whois
- Domain tools: reverse DNS and reverse whois
- http://bgp.he.net
- http://virustotal.com
- http://passivedns.mnemonic.no
- https://www.farsightsecurity.com
- https://www.passivetotal.org

DOMAINTOOLS – OUCH!



DOMAINTOOLS

Invoice

Print this

Payee:
DomainTools.com
2211 5th Ave
Suite 201
Seattle, WA 98121
http://www.domaintools.com

Payer:
Frankie Li
(ran2@vxrl.org)

Payment:
PayPal 3YMVR4Z8TUQS8 fukayli@gmail.com

Invoice Number: DT13555833

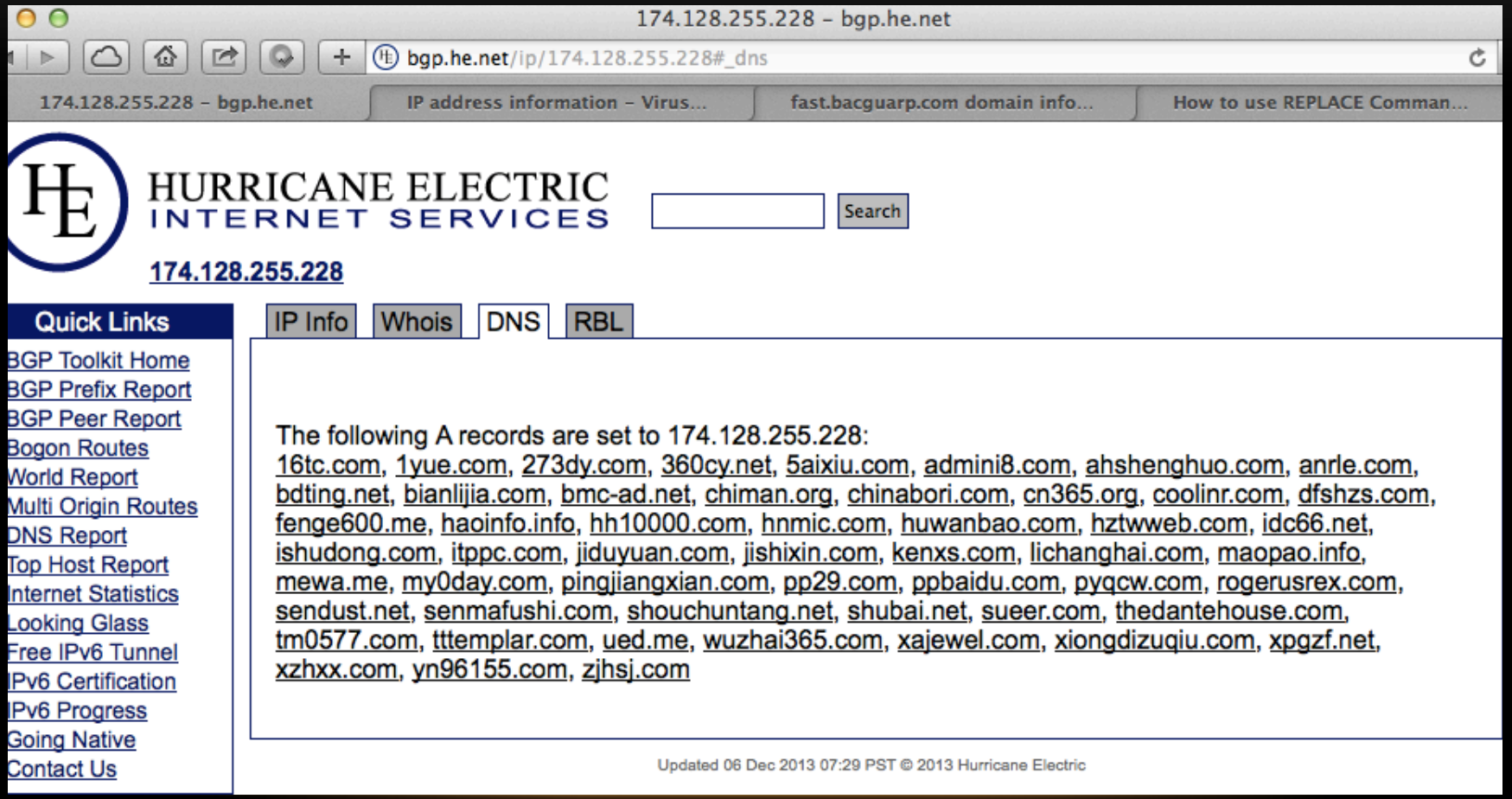
Invoice Date: 2013-03-22 08:28:34

Invoice Status: PAID

Item List:

Item Description	Quantity	Unit Price	Extended Price
Reverse Whois Report (Registrant (Owner) *Exactly Matching* "WANGLUO SHAN")	1	99.00	99.00
		SubTotal:	99.00
		Taxes:	0.00
		Total:	**99.00**

HTTP://BGP.HE.NET

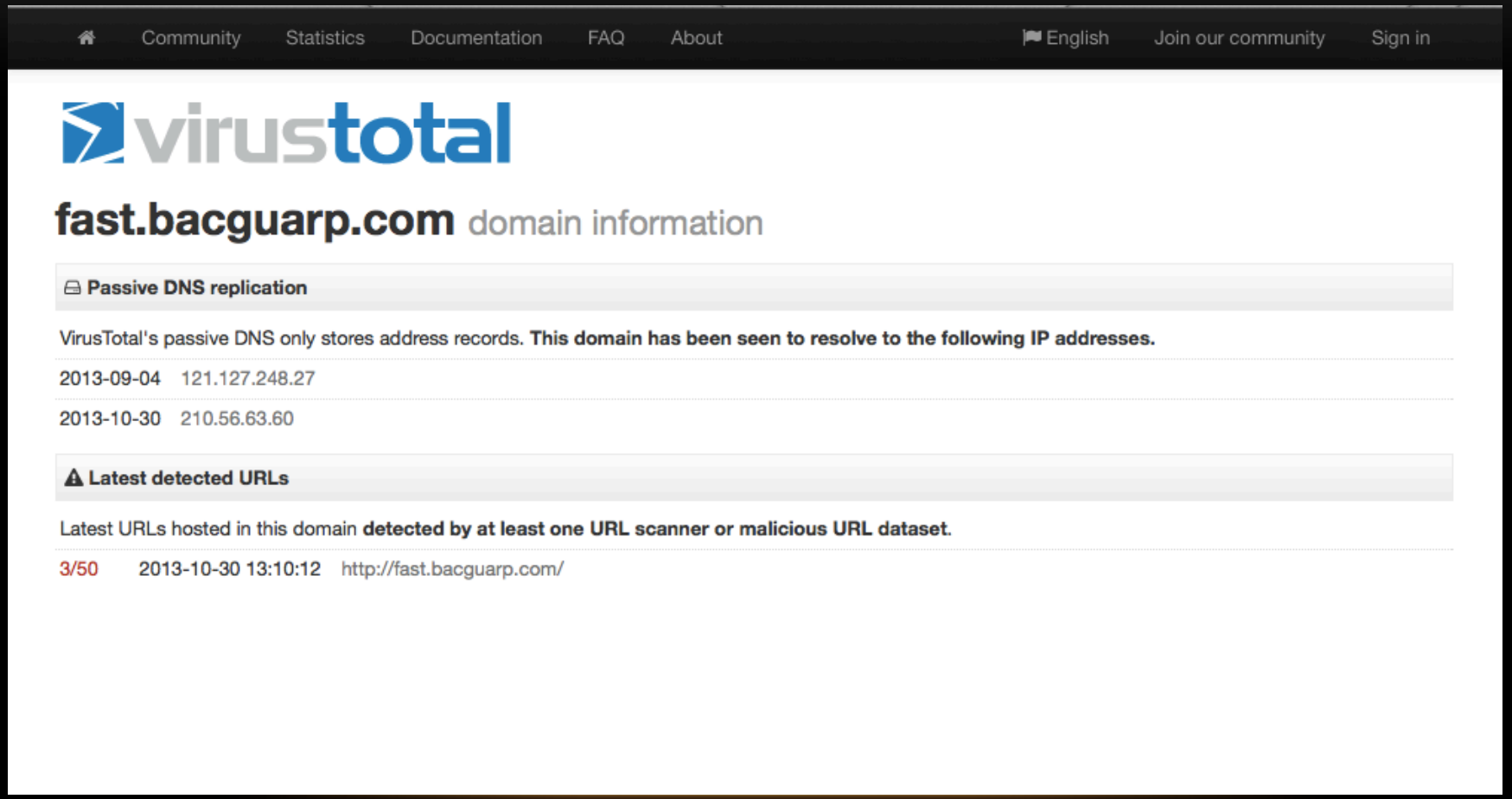


PASSIVE DNS TO PASSIVE WHOIS

PASSIVE DNS

- Passive DNS is a technology that constructs zone replicas without cooperation from zone administrators, and is based on captured name server response
- Passive DNS is a highly scalable network design that stores and indexes both historical DNS data that can help answer questions such as:
  - where did this domain name point to in the past
  - which domain name points to a given IP network
- VirusTotal kept passive DNS records collected from malicious samples
- Higher chance to find malicious historical DNS-IP records

VIRUSTOTAL - PASSIVEDNS



PASSIVE WHOIS

- There are no open source keeping those whois changes, like VirusTotal Passive DNS project (or whois history at who.is)
- By stepping through the IP lookup, retrieval of parked-domains and whois lookup, any changes will then be updated to a relational database

PASSIVE WHOIS

```
select t3.date, t3.name, t1.scan_date, t1.dns, t1.ip_addr, t2.domain, t2.Cname from c2 as t1, domains as t2, samples as t3 where t1.id = t2.sid and t3.id = t1.sid
```

Execute query

Error message from database engine:

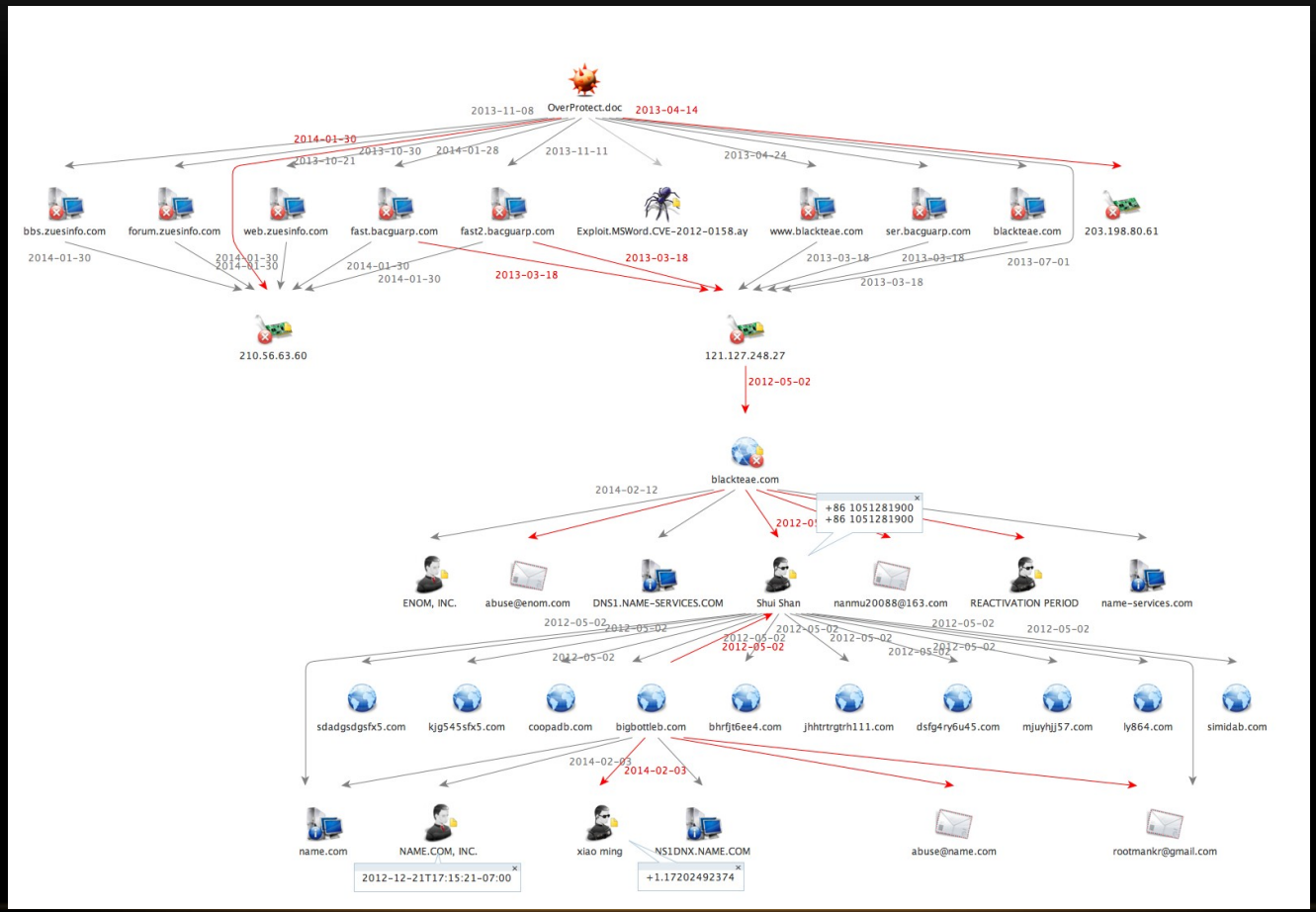
No error

Data returned:

date	name	scan_date	dns	ip_addr	domain	Cname
2013-04-12	Insurance	2013-10-24	wznewbook.gicp.net	174.128.255.228	ued.me	mytension.gicp.net
2013-04-12	Insurance	2013-10-24	wznewbook.gicp.net	174.128.255.228	wuzhai365.com	shiyuekai.gicp.net
2013-04-12	Insurance	2013-10-24	wznewbook.gicp.net	174.128.255.228	xajewel.com	xianidc.gicp.net
2013-04-12	Insurance	2013-10-24	wznewbook.gicp.net	174.128.255.228	xiongdizuqiu.com	syq10086.gicp.net
2013-04-12	Insurance	2013-10-24	wznewbook.gicp.net	174.128.255.228	xpgzf.net	tangjiands.gicp.net
2013-04-12	Insurance	2013-10-24	wznewbook.gicp.net	174.128.255.228	xzhxx.com	xzhxx.gicp.net
2013-03-02	Japan	2013-11-02	webmonder.gicp.net	174.128.255.228	050sf.com	chaocha.gicp.net
2013-03-02	Japan	2013-11-02	webmonder.gicp.net	174.128.255.228	2bbaike.com	116.112.7
2013-03-02	Japan	2013-11-02	webmonder.gicp.net	174.128.255.228	chilia-info.com	qq329684750.gicp.net
2013-03-02	Japan	2013-11-02	webmonder.gicp.net	174.128.255.228	chinabori.com	zoweeoffice.gicp.net
2013-03-02	Japan	2013-11-02	webmonder.gicp.net	174.128.255.228	design-zy.com	qq329684750.gicp.net
2013-03-02	Japan	2013-11-02	webmonder.gicp.net	174.128.255.228	goodnoon.com	todayliu.gicp.net

ANALYSIS BY VISUALIZATION
MALTEGO

SAMPLE CALLED OVERPROTECT



CONCLUSION

INTUITIVE VIEWS ON THE ATTRIBUTION OF APT ATTACKERS

- Continuously monitoring “whois servers” and DNS–IP address pairs
- Intelligence may be lost if they change their TTP in the future, particularly after the publication of this paper
- TTP are determined by the cultural background of the attacker groups
- The intelligence collection process should thus be adjusted toward these changes and analysts should have the same cultural mindset

IS ATTRIBUTION WITH CERTAINTY POSSIBLE?

- All discussed methods may generate some value to the attribution
- But, TTP should carry the highest intelligence value for identifying human attackers
- Any artifacts that support the highest human link should be allocated with highest value to the attribution
- If APT Attribution with Certainty is line starting from 0 to 100%, any artifacts extracted from malware may have some value in this line. No only well funded threat intelligence companies can perform a objective and conclusive attribution
- However, the increasing sharing of TTP and tools by various actors may reduce the reliability to associate with them. (I even read a paper promoting a framework called OpenAPT)
- As a result, the actor groups boundaries are blurred and Espoionage-As-A-Service will be expected
- Another challenging factor is attribution intelligence are not shared enough and intelligence community are not fully understood

THE TOOLS

https://code.google.com/p/malicious-domain-profiling/

MALPROFILE AND MALTEGO TRANSFORM

- The tools consists of 2 parts:
  - MalProfile script to grabbing intelligence from the Internet
  - Maltego Local Transforms to help analysis process

MALPROFILE.PY

```
Ran2:myscripts fukayli$ getAll.py -h
Usage: getAll.py [options]

Options:
  -h, --help   show this help message and exit
  -i           initialize c2 database [c2_dev.db]
  -f FILENAME  Provide a FILENAME to check
  -d DNS       Provide a DNSNAME to check
  -c           rescanning c2 to update all subsequent tables
  -o           rescanning owner table to update all subsequent tables
  -p           rescanning passive tables to update ip table
  -q           rescanning ip table to update domains & whois tables
  -r           rescanning domains table to update passive_ip table
  -s           rescanning ip table to update passive_domains & passive_whois
               tables
  -t           rescanning and update tmp table
  -w           rescanning and update domains table to update whois
  -x           rescanning and update whois table from passive_whois
Ran2:myscripts fukayli$
```

FURTHER RESEARCH
PLUG-INS

MALPROFILE.PY

- The script is modified as a class to allow plugins be added
- To allow more intelligence can be added when new TTP be identified
- Or, combined the technical context be included as a supplement when performing intelligent analysis

GOOGLE PROJECT

- Special thanks go to Kenneth Tse, Eric Yuen who is upgrading my messy code into a class and Frank Ng help me to manage the project
- You can find the code at: https://code.google.com/p/malicious-domain-profiling/
- Any interested are welcome to contribute to this project. Please contact ranerwang@gmail.com or kennetht@gmail.com

MALICIOUS-DOMAIN-PROFILING

Introduction

MalProfile? is a set of tools to:

1. Fetch useful data from different sources include malware samples, suspicious IP/Domain being used, passive DNS records, md5 hash and save to a database at different time slot for behaviour and/or timeline analysis
2. Present in Maltego the relationship of malware, current and passive domain/IP/Email/Telephone etc to get the origin of the source. And elaborate the relationship to get suspected IP/Domain for proactive prevention and detection.

History

Please refer to CHANGELOG?

Requirements

1. Kali Linux 1.0.7 or later (for illustration purpose only, for advance users, just use the tool per your preference, in my case, I install it on my Mac)
2. Maltego Edition 3.4.0 or later (If community version is used, only 12 records will be randomly displayed)
3. Virustotal registration and API key
4. Maltego Basic Python Library - https://www.paterva.com/web6/documentation/developer-local.php

(Other system with Python 2.7 and Maltego may work but never tried :))

Package Files

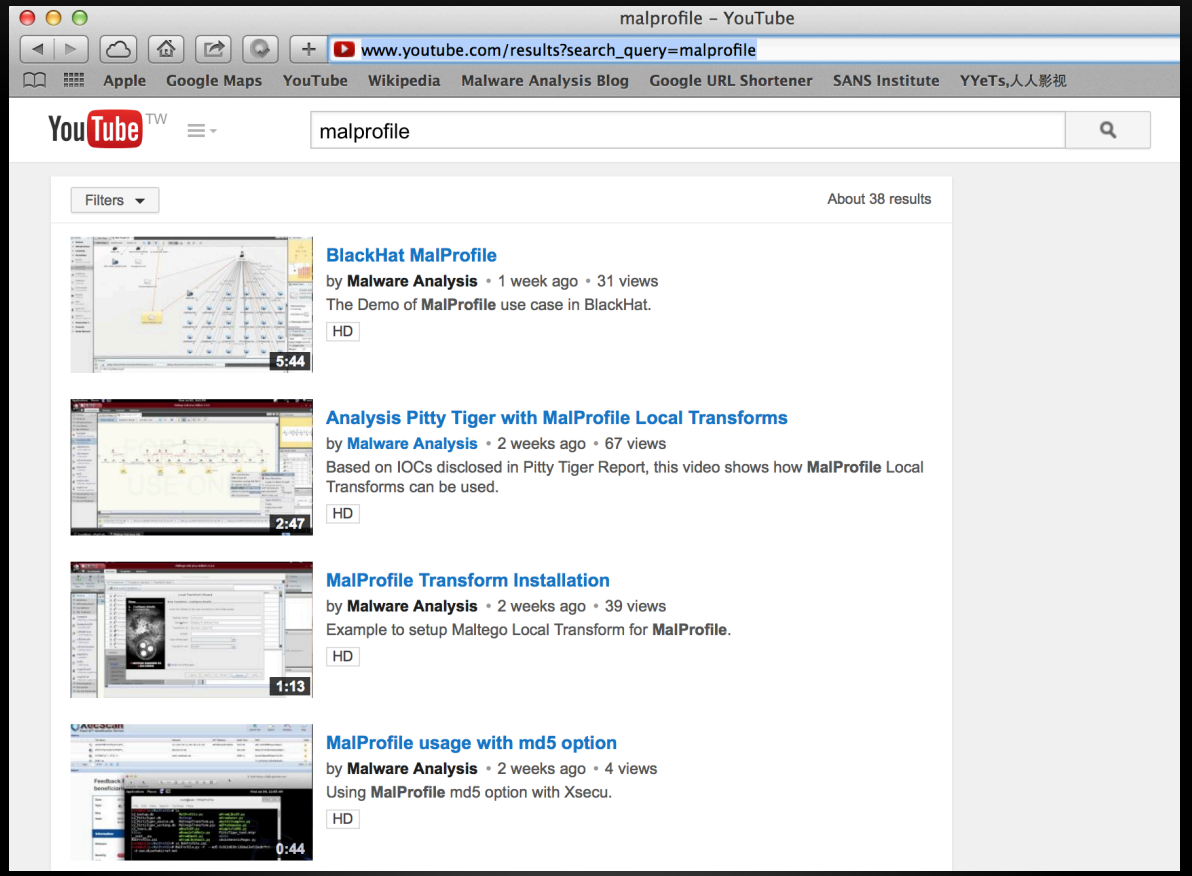
The following files are included in the MalProfile? package.

```
MalProfile/MalProfile.py                    # MalProfile main script
MalProfile/MalProfile.ini                   # MalProfile configuration file
MalProfile/README.txt                       # this file
MalProfile/c2_PittyTiger                    # Sample database file (not included in the code email ran2@vxrl.org)
MalProfile/c2_Xsecu                         # Sample database file (not included in the code email ran2@vxrl.org)
MalProfile/Maltego/MyEntities.mtz           # Maltego Input Entities
MalProfile/Maltego/*                        # Maltego Transform scripts, Refer to ReadMe/Transform_Readme for more info
MalProfile/Utils/*                          # Libraries and plugins for MalProfile
ReadMe/*                                    # Documentation of MalProfile design and usage
Samples/*                                   # Samples for demonstration (not included in the code email ran2@vxrl.org)
```

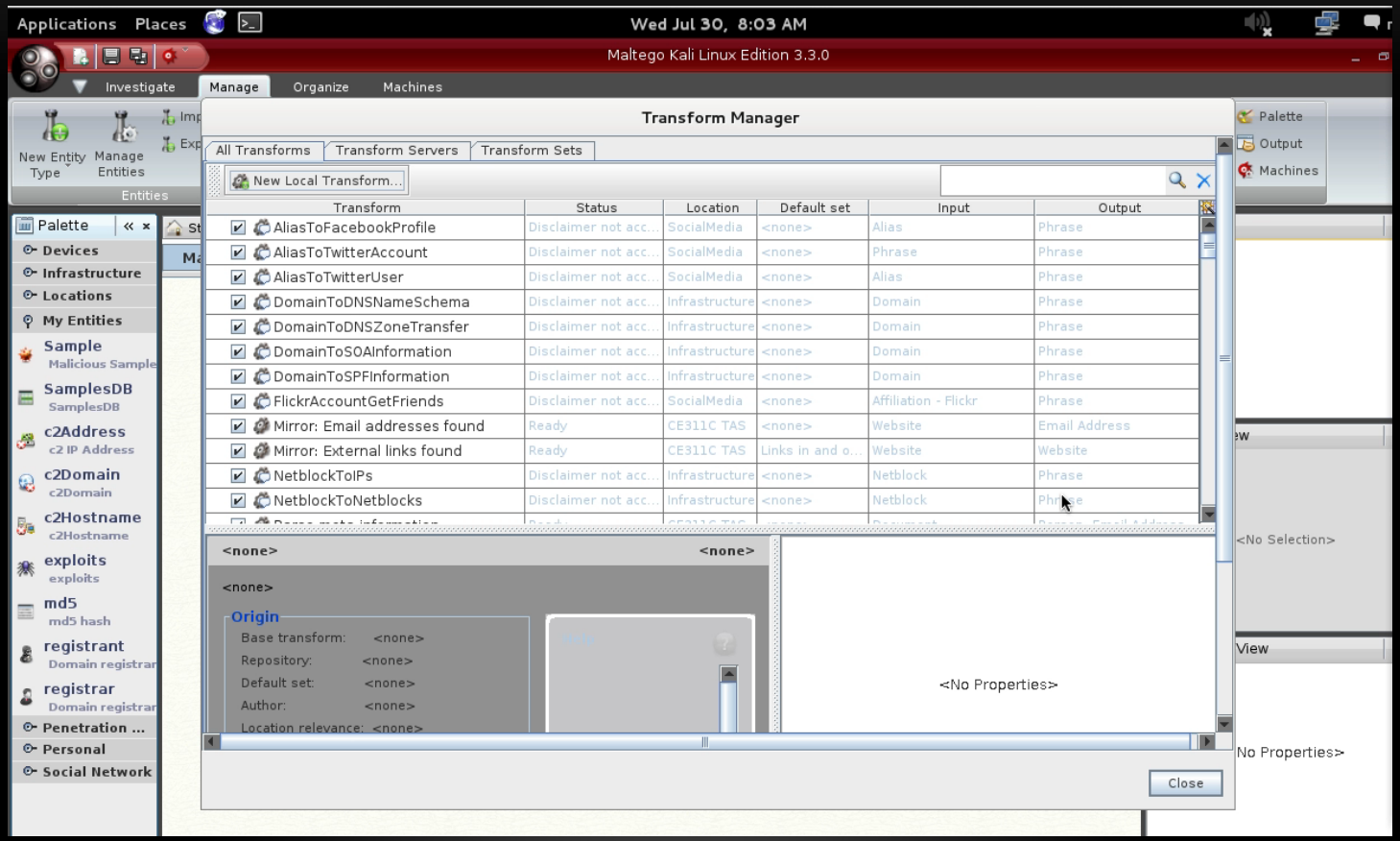
Installation

1. unzip the MalProfile.zip to /Root/MalProfile
2. apt-get install python-setuptools
3. easy_install pip
4. pip install python-whois
5. pip install hashlib

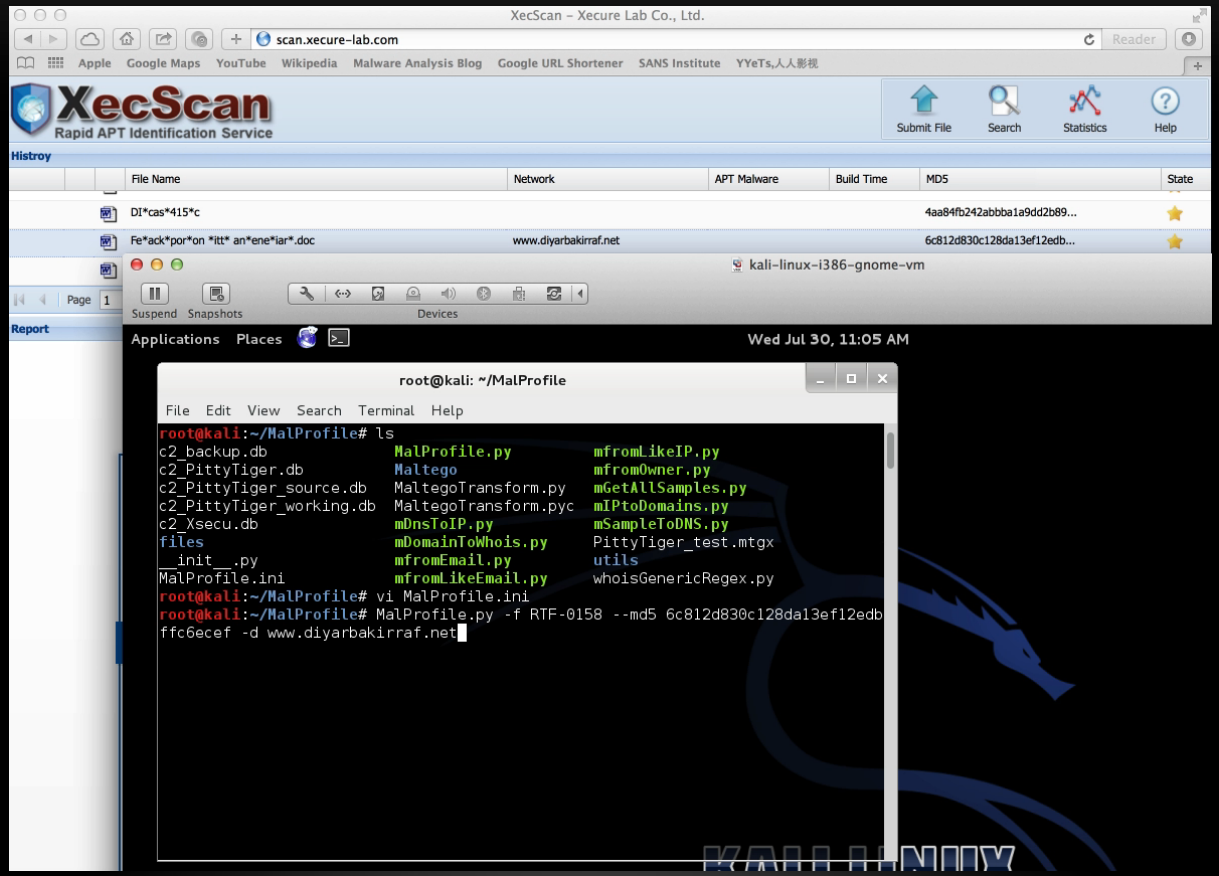
HTTP://WWW.YOUTUBE.COM/RESULTS?SEARCH_QUERY=MALPROFILE



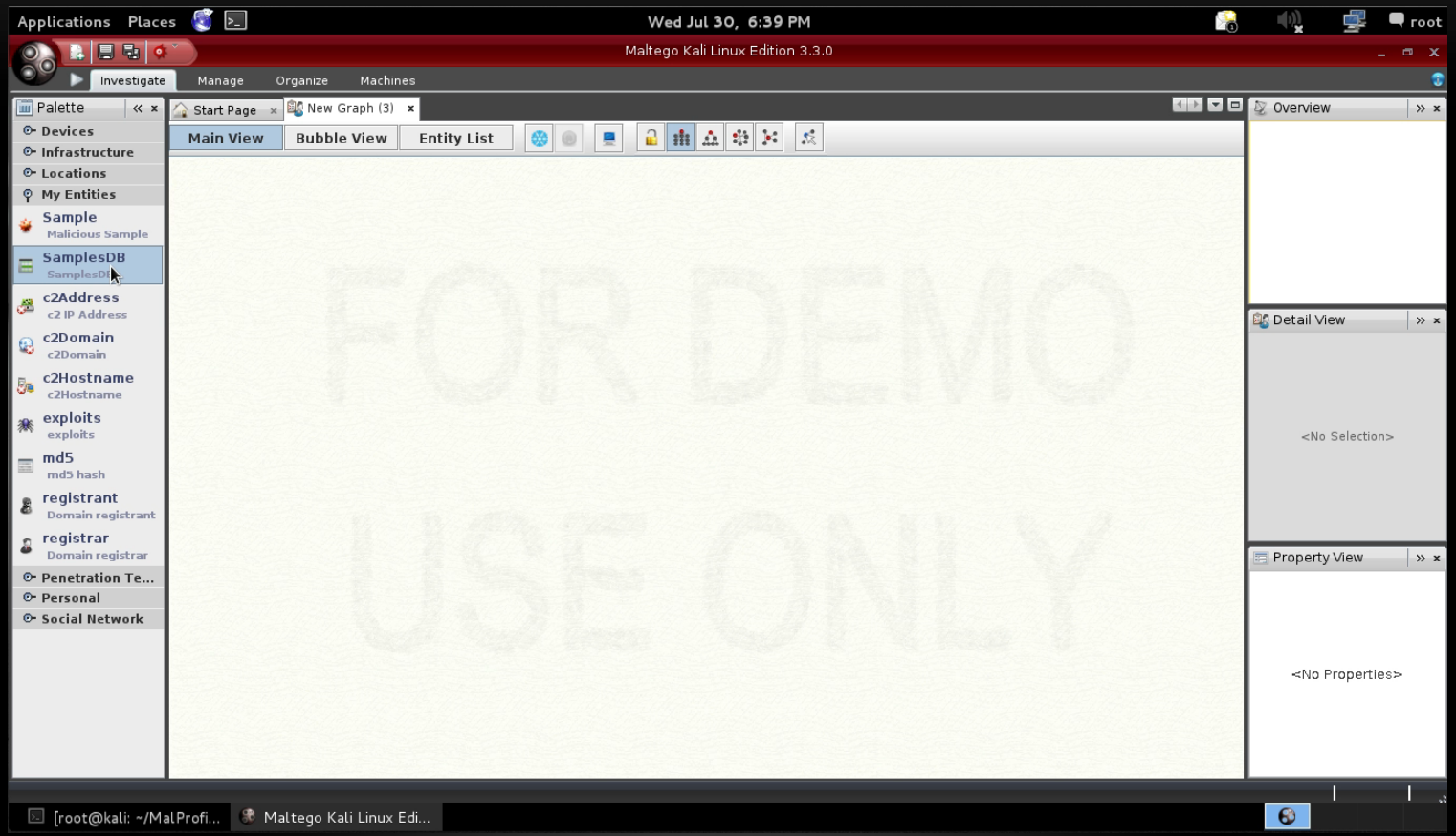
MALPROFILE TRANSFORM INSTALLATION



USING MD5 WITH OSINT FROM XECSCAN ☺

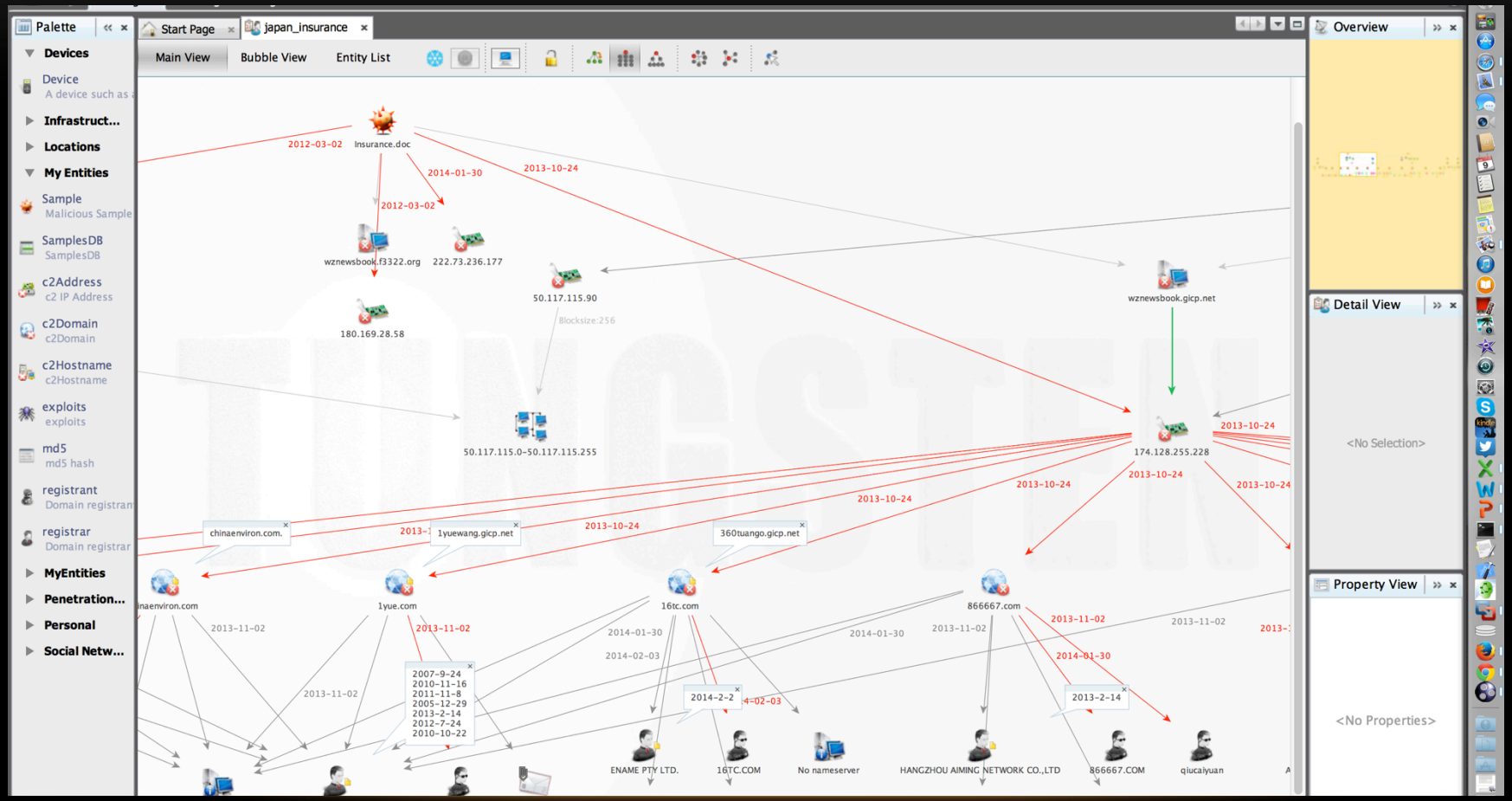


PITTY TIGER ANALYSIS



DEMO

SAMPLE CALLED INSURANCE & JAPAN



THANK YOU!
Q&A

Frankie Li

ranerwang@gmail.com

http://espionageware.blogspot.com

PLEASE COMPLETE THE SPEAKER FEEDBACK SURVEYS

Frankie Li

ranerwang@gmail.com

http://espionageware.blogspot.com