DNSSM: A Large Scale Passive DNS Security Monitoring Framework

Samuel Marchal, Jérôme François, Cynthia Wagner, Radu State, Alexandre Dulaunoy, Thomas Engel, Olivier Festor
Outline

1. Motivation
2. Solution
3. Experiments and Results
4. Conclusion
Outline

1 Motivation
2 Solution
3 Experiments and Results
4 Conclusion
Overview of DNS

- DNS (Domain Name System) is the service that maps a domain name to its associated IP addresses
  - www.example.com ⇒ 123.45.6.78
- DNS is the service that allows to find information about a domain:
  - A: IPv4 address
  - AAAA: IPv6 address
  - MX: Mail server
  - NS: Authoritative DNS server
  - TXT: any information
Why DNS monitoring?

- **DNS:**
  - critical Internet service
  - threats: cache poisoning, typosquatting, DNS tunnelling, fast/double-flux
    ⇒ enhance: phishing, botnet C&C communications, covered channel communications etc.
  ⇒ Patterns in DNS packet fields and DNS querying behavior

- Passive DNS monitoring to detect:
  - worm infected hosts
  - malicious **backdoor** communication
  - botnet participating hosts
  - phishing websites hosting
Existing solutions

- Mainly use supervised classification techniques
  - SVM, tree, rules, etc.
  - require malicious data for training
- Targeted identification of malicious domains
  - C&C communication involved domains
  - Phishing domains
  - Spamming domains
  - etc.
Automated clustering technique for online analysis

- No previous knowledge
- Group domains regarding their activity
- DNS information $\Rightarrow$ Domain activity
- Disclose the raise of new threats
- K-means clustering
- 10 relevant features
For each domain observed:

- Number of IP addresses
- IP scattering: entropy based and position weighted
- mean TTL
- Requests count
- Period of observation
- Requests per hour
- Name servers count
- Number of subdomains
- Blacklisted flag
DNSSM is an approach for automated analysis of DNS (passive traffic)

- Manual assistance in tracking anomalies:
  - Feed with cap file
  - All DNS packet fields extracted
  - MySQL database storage model
  - Web interface
  - Fast and efficient mining functions
  - Integrates with existing blacklist tools to assist in tagging data
  - Detection of fast/double flux domains, DNS tunnelling, etc.
  - Freely downloadable at: https://gforge.inria.fr/docman/view.php/3526/7602/kit_dns_anomalies.tar.gz
Architecture

DNS passive monitoring

End-user tool

Centralized Storage

MySql Database

Data access and process

Passive DNS Sensor

Recursive Queries / response

DNS Server

Recursive Queries / response

DNS Server

DNS Server

IP addresses of a domain? FastFlux domains? Most queried domains?

Web server (Apache + PHP)

Client

Client

Client

Query/response

Query/response

Client
Experiments

- 2 datasets (≠ location, ≠ type of network, ≠ users, ≠ quantity)
- Automatic results from k-means: 8 clusters exhibiting different properties

- Cluster 5: apple.com, amazon.fr, adobe.com (highly popular websites)
Cluster 6: google.com, skype.com, facebook.com (highly popular web sites)

Cluster 7: tradedoubler.com, doubleclick.net, quantcast.com (user tracking)

Cluster 3: akamai, cloudfront.net (CDN)
Cluster 0: small websites with low popularity
1 Motivation
2 Solution
3 Experiments and Results
4 Conclusion
Passive DNS monitoring solution
- Analysis of domain names activity
- Relevant data mining algorithm (unsupervised clustering techniques)
- Efficiency proved on two different datasets
- Freely downloadable interface

Applications:
- Investigate cyber security fraud
- Debug DNS deployment
- Penetration testing
DNSSM: A Large Scale Passive DNS Security Monitoring Framework

Samuel Marchal, Jérôme François, Cynthia Wagner, Radu State, Alexandre Dulaunoy, Thomas Engel, Olivier Festor