Universität Passau || Siemens CERT || Master's thesis - Exposé

Evaluation of domain reputation scoring algorithms in the field of IT-Security and development of a probabilistic hostile activities accounting algorithm.

Felix Steghofer
March 7, 2017
Advisor: Thomas Penteker
Supervisor: Prof. Dr. rer. nat. Joachim Posegga

1 Introduction

The domain name system (DNS) has been one of the corner stones of the internet for a long time. It acts as a hierarchical, bidirectional translation device between mnemonic domain names and network addresses. It also provides service lookup or enrichment capabilities for a range of application protocols like HTTP, SMTP, and SSH. In the context of defensive IT security, investigating aspects of the DNS can facilitate protection efforts tremendously. Estimating the reputation of domains can help in identifying hostile activities. Such a score can, for example, consider features like quickly changing network blocks for a given domain or clustering of already known malicious domains and newly observed ones.
The task of this work is to evaluate existing scoring mechanisms of domains in the special context of IT security, and also research the potential for combining different measurement approaches. It ultimately shall come up with an improved and evaluated algorithm for determining the probability of a domain being related to hostile activities.

2 Exposé

For the improved algorithm we want to investigate a couple of approaches. There has already been done some work in related topics so far, with an active research group residing at the Georgia Institute of Technology. Antonakakis et al. have developed two dynamic domain reputation systems based on machine learning. These are shortly introduced first as they can be referred to as the state of the art in the field of *DNS reputation score* with Notos (2010) being the

> first to create a comprehensive dynamic reputation system around domain names [2]

and Kopis (2011) following with a detection rate for malicious domains of ∼98.4%) and a false positive rate of ∼0.4%. Furthermore these are the most popular papers in this research section according to Google scholar citations [7] and Mendeley read counts [8].
Notos uses passive monitoring of DNS query data and its idea is described with:

> The premise of this system is that malicious, agile use of DNS has unique characteristics and can be distinguished from legitimate, professionally provisioned DNS services [2].

Kopis on the other hand is operating in the upper DNS hierarchy and makes use of global DNS query resolution patterns to detect malware related domains with features like the requester diversity, the requester profile or the reputation of involved IPs [3]. For a more detailed overview how Notos and Kopis accomplish this task, see the Related work section.
A third algorithm –Exposure– has been developed by Bilge et al. [4] and operates in the same DNS layer as Notos does (passive DNS monitoring of recursive resolvers) but uses a different feature set to evaluate domains.

To evaluate these existing algorithms, we are provided with data from DNS resolvers that are operated in a big company network with a size of of many gigabytes.
As this data set is collected in recursive DNS resolvers in a lower DNS level, it is only suitable for evaluating the Kopis and Exposure system. To evaluate Kopis, data obtained in an upper DNS layer (e.g. a country level TLD server) would be needed. Furthermore for data secrecy we are not given the full DNS client requests so this probably limits the features to be evaluated.
After evaluating the existing algorithms, a set of suitable features for the available data set has to be selected to develop a succeeding algorithm. Largely this consists of combining relevant features used by the evaluated algorithms and adding further features in order to result in a more accurate score. Additional parameters we have thought of that could be taken into account:

- the character distribution within the domain name: A lot of malware related domains include e.g. randomly generated domains as stated by Yadav et al. [11]

- the device class of the machine the DNS request is originating from (i.e. a PC or an embedded device as determined by passive OS fingerprinting): According to Fitz et al. embedded systems'

  > system architectures are based on highly insecure and error-prone foundations [6]

  and therefor are highly vulnerable for intrusions. Many incidents found in the last years included vulnerable embedded devices such as a DDoS in October 2016 with about 100.000 "compromised gadgets [that] knocked an Internet infrastructure provider partially offline" [9]

- Additionally we are provided with a manually maintained IP/Domains blacklist that can be used along with publicly available filter lists like the Alexa top 500 for whitelisting domains [1].

In the first step of this work (∼two or three months), all previous efforts for labeling domains with a reputation score have to be investigated and the developed algorithms are evaluated with the available data set. The next month or two will be used for the selection of suitable features as well as implementation of a succeeding algorithm and evaluating it on a suitable data set. In the last step (∼one months), the thesis will be finalized.

3 Related work

Malware related dynamic domain reputation systems (passive DNS request/response monitoring Machine Learning approaches):

- Notos (passive monitoring of recursive DNS traffic) [2]
- Exposure (like Notos, but different feature set) [4]
- Kopis (working in the upper DNS hierarchy) [3]

See Figure 1 for an example of possible features. (Extracted by Exposure to do a sentiment analysis)

Feature Set	#	Feature Name
Time-Based Features	1	Short life
	2	Daily similarity
	3	Repeating patterns
	4	Access ratio
DNS Answer-Based Features	5	Number of distinct IP addresses
	6	Number of distinct countries
	7	Number of domains share the IP with
	8	Reverse DNS query results
TTL Value-Based Features	9	Average TTL
	10	Standard Deviation of TTL
	11	Number of distinct TTL values
	12	Number of TTL change
	13	Percentage usage of specific TTL ranges
Domain Name-Based Features	14	% of numerical characters
	15	% of the length of the LMS

Table 1: Features.(LMS = Longest Meaningful Substring)

Figure 1: Features used in Exposure [4]

In comparison, the features of Kopis:
At first, the following data is extracted out of each DNS request/response pair.
$Qj(d) = (Tj, Rj, d, IPsj)$ where

- Tj is the epoch (time of the request/response [e.g. on a daily basis])
- Rj is the IP of the requests initiator
- d the queried domain and
- $IPsj$ is the set of resolved IPs for this domain as responded

Using this information, the following features are used to build the reputation score:

- Requester Diversity: Where do request originate (overall)
- Requester Profile: Is the requester a single computer or does it itself handle/serve many client (RDNS server of a large ISP)? Different profiles can therefor be weighted accordingly.
- Resolved-IPs Reputation (IPR): This basically checks a database for the reputation of all resolved IPs. In detail the following aspects are audited:
  - *Malware Evidence*: Average number of know malware related domains that have pointed to that IP in the last month (with respect to the epoch)
  - *SBL Evidence* very much like the Malware Evidence but with a external IP spam list (Spamhaus Block List [10])
  - *Whitelist Evidence*: Number of IP addresses pointed by known good domains (DNSWL [5] and top 30 domains according to Alexa [1])

Comparing those three systems, Kopis succeeds for a dynamic, independent and global domain reputation scoring algorithm so far. It uses a supervised machine learning approach where within the training mode, a set of sentimentally annotated *malware-related* and *known legitimate* domain names is used to build a model based on query/response patterns that can later be used to statistically classify in operational mode. In total numbers it features a high detection rate (~98.4%) as well as a low false positive rate (~0.4%)

References

[1] Amazon. Alexa. The web information company. http://www.alexa.com/, Feb. 2017.

[2] M. Antonakakis, R. Perdisci, D. Dagon, W. Lee, and N. Feamster. Building a dynamic reputation system for dns. In *USENIX security symposium*, pages 273–290, 2010.

[3] M. Antonakakis, R. Perdisci, W. Lee, N. Vasiloglou II, and D. Dagon. Detecting malware domains at the upper dns hierarchy.

[4] L. Bilge, E. Kirda, C. Kruegel, and M. Balduzzi. Exposure: Finding malicious domains using passive dns analysis. In *Ndss*, 2011.

[5] DNSWL. DNS Whitelist Protect against false positives. DNSWL. https://www.dnswl.org/, Feb. 2017.

[6] R. Fitz, W. A. Halang, and L. Zhang. Malware-proof embedded systems. In *Future Information Technology-II*, pages 145–153. Springer, 2015.

[7] Google. Google Scholar. https://scholar.google.de/scholar?q=dns+reputation+score&btnG=&hl=de&as_sdt=0%2C5, Feb. 2017.

[8] Mendeley. Mendeley. https://www.mendeley.com/research-papers/search/?query=dns+reputation+score, Feb. 2017.

[9] B. Schneier. Schneier on Security – Botnets. https://www.schneier.com/blog/archives/2017/03/botnets.html, Mar. 2017.

[10] Spamhaus. SBL. The Spamhaus Project Block List. https://www.spamhaus.org/sbl/, Feb. 2017.

[11] S. Yadav, A. K. K. Reddy, A. Reddy, and S. Ranjan. Detecting algorithmically generated malicious domain names. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, pages 48–61. ACM, 2010.

List of Figures