Antonia Affinito

Using DNS to understand user behavior over the Internet

Tutor: Alessio Botta

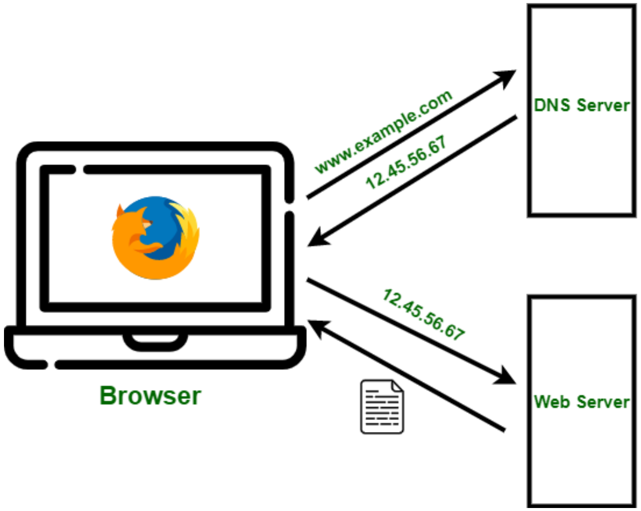
Cycle: XXXV

Year: Second

My background

- MSc degree in Computer Engineering, University of Naples “Federico II”
- Research group: COMICS with Prof. Alessio Botta
- PhD start date: Academic Year 2019-2020
- Scholarship type: MIUR research grant

The Domain Name System



- Its role is to convert human-readable names (ex. example.com) in their corresponding IP addresses (93.184.21.34)

- It is considered the phonebook of the Internet

- New domain names are registered every day, but 70% of them are malicious, sospicious or not safe to work [1]

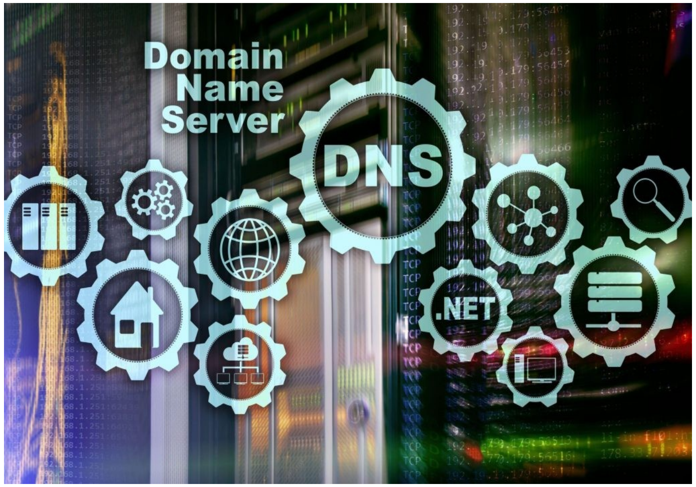
- The DNS is typically used to establish a link between IoT botnets and their Command-and-Control server.

[1]: Z. Chen, J. Javier Wang, K. Kwan; Newly Registered Domains: Malicious Abuse by Bad Actors. Palo Alto Company

Starting Ideas

- The DNS is considered a valid tool to analyse a lower percentage of traffic and to extract interesting information about the network operations.

- It is an important observation point in order to study the main issues of current networks, including performance and security.



Research Questions:

- Is it possible to look at the trend changes of the most popular apps with the DNS data?

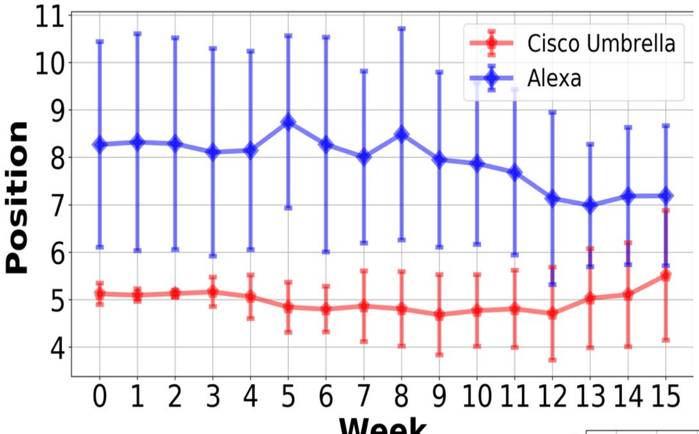
- Which type of DNS resolver - local or public - has the best security/response time ratio?

- Is it possible to discriminate benign and malicious domains by looking at their lifetime and/or other features provided by the DNS?

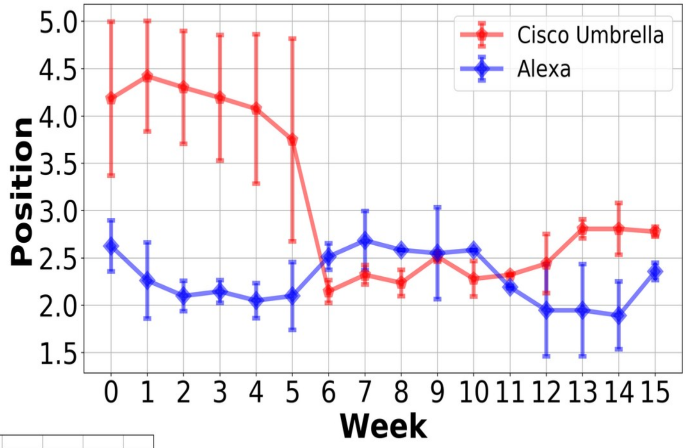
Top 1 Million Lists: Pandemic Period

- We started from two lists of the most popular **Top One Million** domain names
  - Collected every day by **Cisco Umbrella** and **Alexa**
- Looking at the trends of the most popular applications, divided by categories, we derived how their usage changed during the **COVID** pandemic period

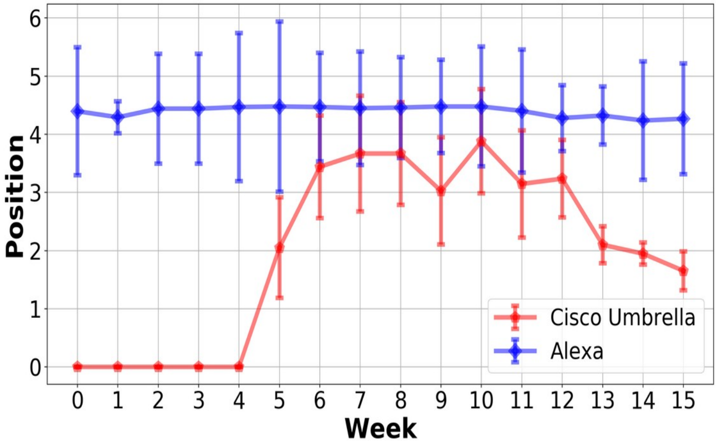
skype.com



facebook.com

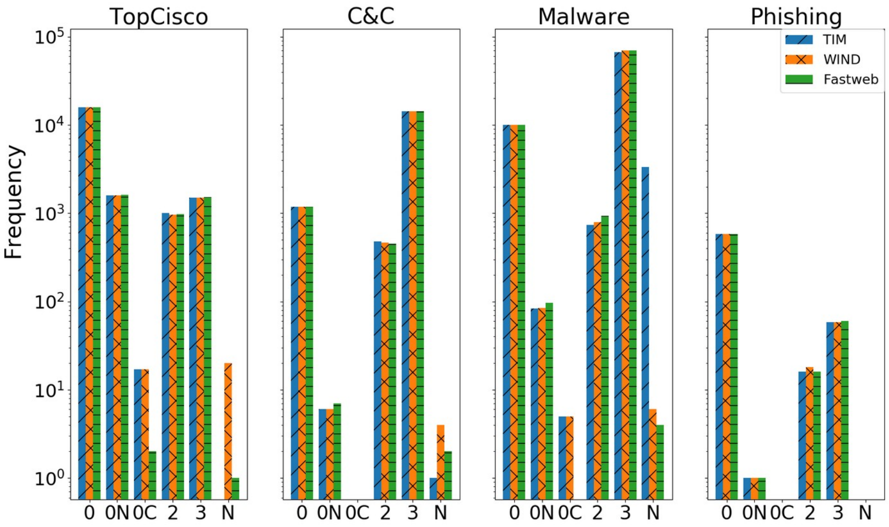


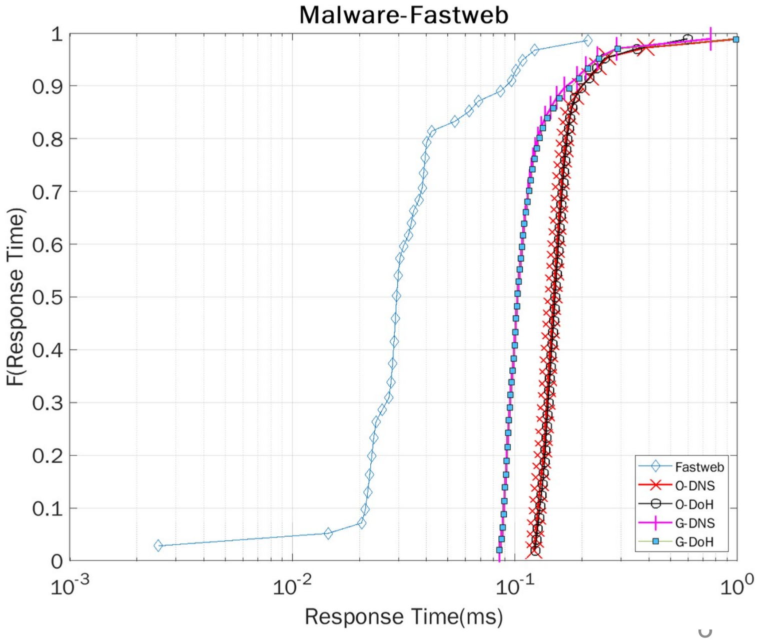
netflix.com



DNS Resolvers – Local vs Public

- We created a **dataset** performing queries on a large number of domain names to three Italian ISPs and two Public resolvers.
  - We relied on the domain names provided by Cisco analysts, divided in three **malicious categories**: malware, phishing, command-and-control.
  - We focused particularly on the **response time** and **response code**.
- We investigated the **performance** of the DNS resolvers - Local and Public- in terms of capability to recognize malicious domains and the response time.



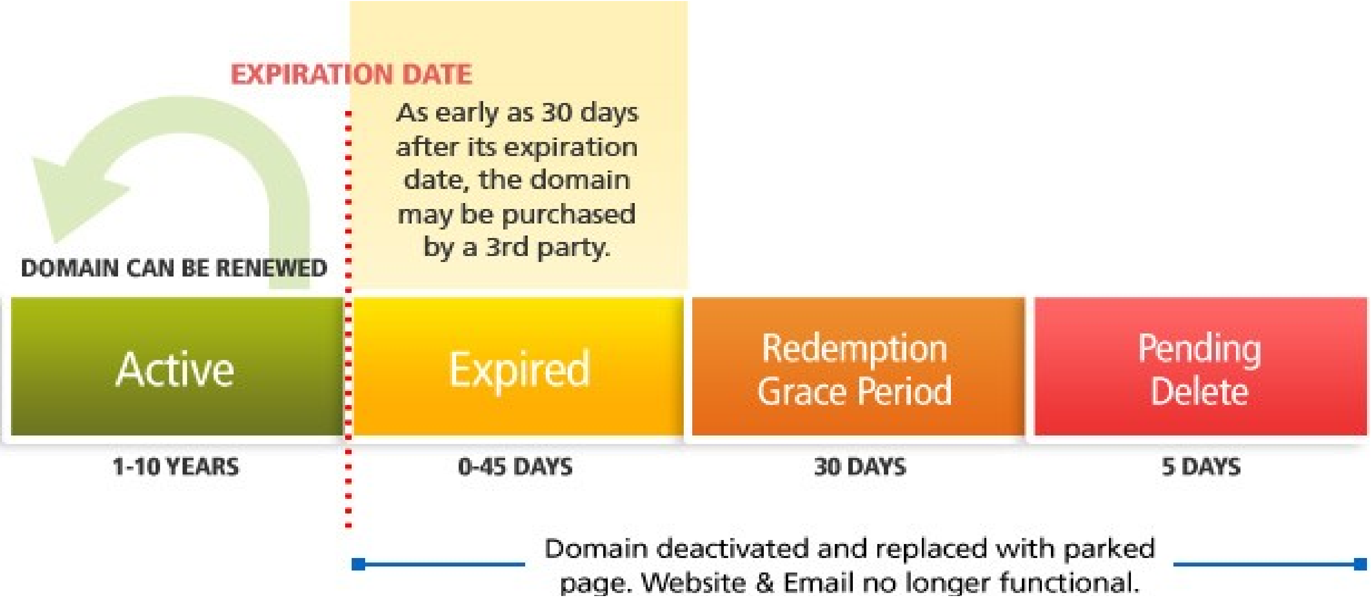


Lifetime of the Domain Names – Period Abroad

- I am currently spending my **period abroad** at the University of Twente (Netherlands) from April 2021 to March 2022, working on the lifetime of domain names.

- The **lifetime** of a domain is set to approximately 1-2 years for benign domains.
  - The lifetime of a **malicious** domain name is shorter than that of a benign domain [1].

- Detection of the malicious domain names through their lifetime retrieved from the information in the **zone files**.


EXPIRATION DATE
As early as 30 days after its expiration date, the domain may be purchased by a 3rd party.
DOMAIN CAN BE RENEWED
Active
Expired
Redemption Grace Period
Pending Delete
1-10 YEARS
0-45 DAYS
30 DAYS
5 DAYS
Domain deactivated and replaced with parked page. Website & Email no longer functional.

- [1]: N. Hason; A. Dvir, C. Hajaj; Robust Malicious Domain Detection; Cyber Security Cryptography and Machine Learning. CSCML 2020

Third Year – Next Ideas

- The DNS traffic contains a significant amount of meaningful features useful to identify domain names associated with malicious activities.

- The **main objective** of the research project is to detect malicious domain names:

    - Analysing response **times** and **codes** of local and public DNS resolvers;

    - Studying their **lifetime**;

    - Using features collected by OpenIntel and Certificate Transparency Logs applying supervised **machine learning** algorithms.

My Products

[P1]	*Antonia Affinito, Alessio Botta, Luigi Gallo, Mauro Garofalo, Giorgio Ventre; “Spark-based Port and Net Scan Detection”; The 35th ACM/SIGAPP Symposium on Applied Computing- ACM SAC; published; 2020.*
[P2]	*Antonia Affinito, Alessio Botta, Giorgio Ventre; “The impact of Covid on network utilization: an analysis on domain popularity”; IEEE International Workshop on Computer Aided Modeling and Design of Communication Links and Networks- CAMAD; online conference; published; 2020.*
[P3]	*Antonia Affinito, Alessio Botta, Giorgio Ventre; “Local and Public DNS Resolvers: should we trade off performance against security?”; IEEE/IFIP Network Operations and Management Symposium; submitted; 2021.*

	Courses	**Seminars**	**Research**	**Tutorship**	**Total**
Bimonth 1	**0**	**3.3**	**6**	**6**	**15.3**
Bimonth 2	**6**	**1.3**	**5**	**0**	**6.3**
Bimonth 3	**0**	**0.6**	**9**	**0**	**9.6**
Bimonth 4	**0**	**1.2**	**8**	**0**	**9.2**
Bimonth 5	**5**	**0**	**7.5**	**0**	**12.5**
Bimonth 6	**0**	**0.4**	**8**	**0**	**8.4**
Total	**11**	**6.8**	**43.5**	**6**	**61.3**
Expected	**30 - 70**	**10 - 30**	**80 - 140**	**0 – 4.8**	

Thanks for the attention!