Are Darknets All The Same? On Darknet Visibility for Security Monitoring

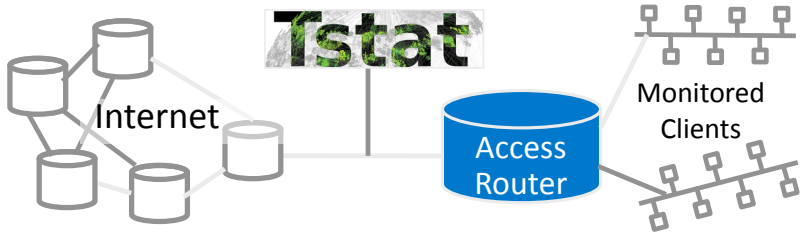
Francesca Soro, Idilio Drago, **Martino Trevisan**, Marco Mellia,
Joao Ceron, Jair J.Santanna

Conferenza GARR 2019

4 Giugno 2019

Data Science for Network Monitoring

The TNG Group and the SmartData@PoliTo center work on Data Science applied to networking

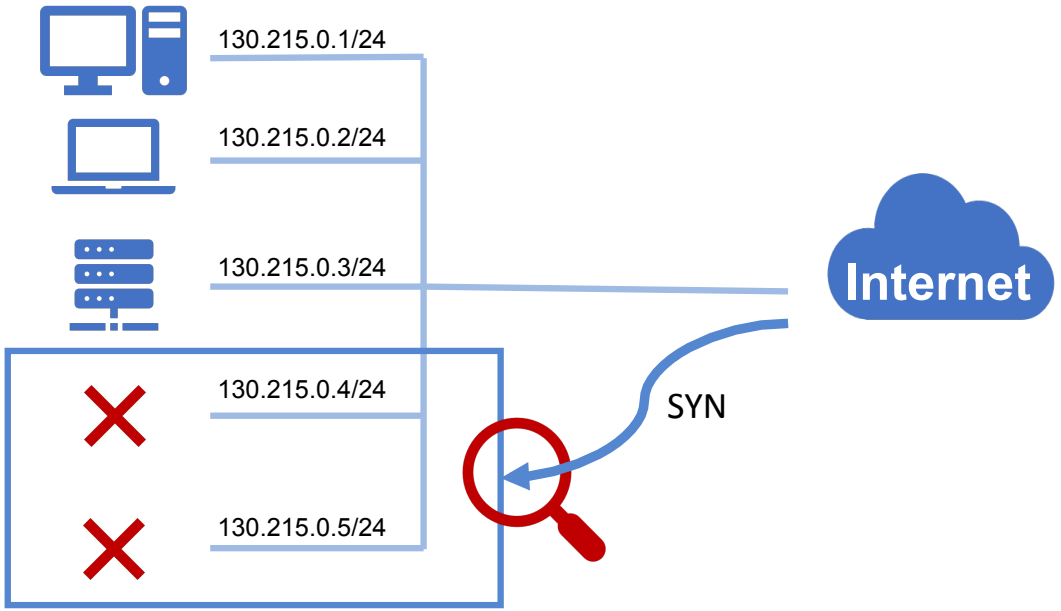


- Passive monitoring using custom software
  - TCP/IP level measurements
- Network measurement lab:
  - **Darknet traffic**, honeypots, social network data

Data Science to extract knowledge from measurements
- Characterization
- Classification
- Synthetic Data generation

What is a darknet?

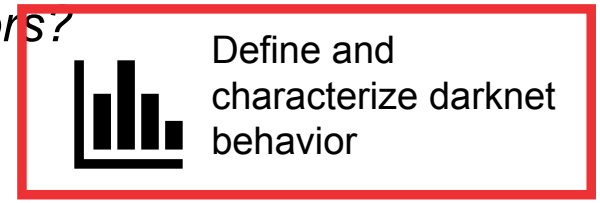
Darknets are sets of IP addresses that are advertised without answering any traffic. *They passively record the incoming packets aiming to assist on network monitoring activities.*

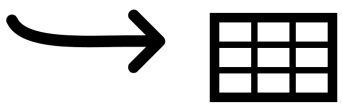


Objective

Darknets have proven to be a precious instrument when it comes to **network traffic monitoring scenarios**, prompt detection of **zero-day cyberattacks**, and analysis of the spread of a **botnet** infection.

*But can we use them to create **general** models of such behaviors?*





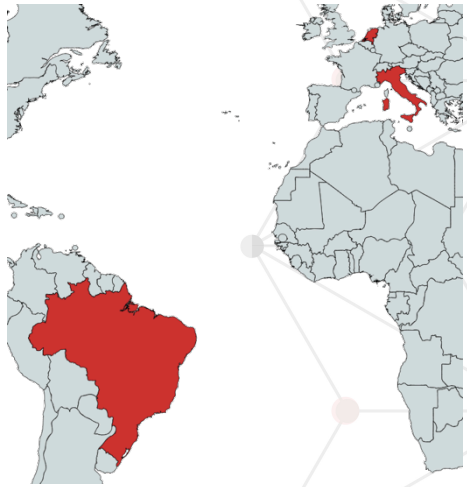
Methodology

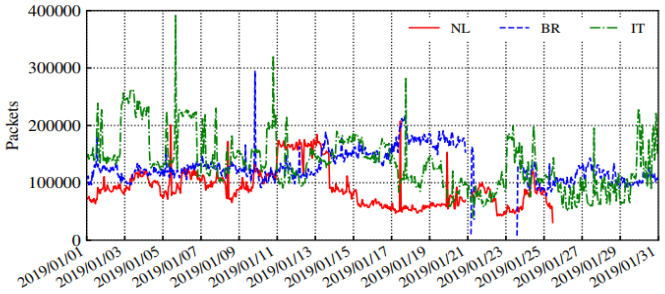
Comparison and characterization of the traffic hitting two darknets:

- **/19** located in Brazil → **8,192** IPs
- **/15** located in the Netherlands → **131,072** IPs
- **3 different /24** from GARR network → **768** IPs

In terms of:

- **Traffic volume**
- **Traffic type (TCP scan, UDP, …)**
- **Traffic origins** (AS and Country of sources)



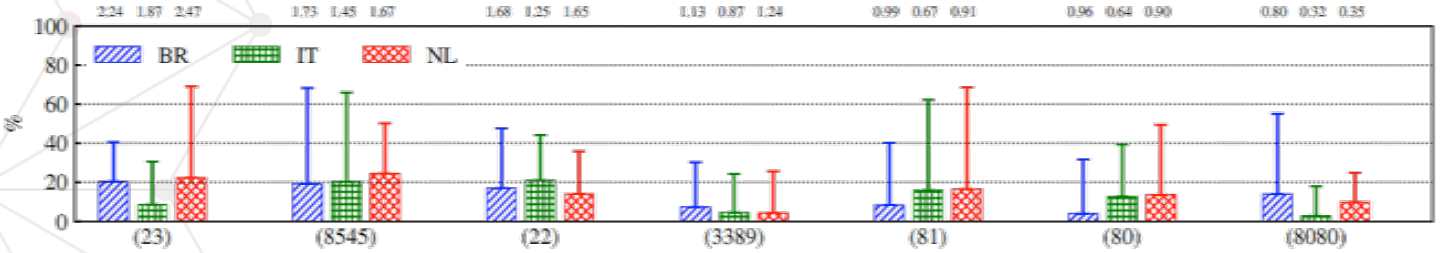


Type	*NL*/15		*BR*/19		*IT* 3 × /24	
	Pkts	IP addr.	Pkts	IP addr.	Pkts	IP addr.
Scan	85.1%	12.5%	84.8%	4.6%	86.9%	3.2%
Back.	3.7%	0.8%	2.3%	0.6%	0.2%	0.2%
UDP	5.7%	10.8%	4.3%	2.3%	3.8%	1.8%
ICMP	0.5%	1.6%	0.5%	0.8%	0.3%	0.6%
Other	4.8%	74.1%	7.8%	91.4%	8.6%	93.9%

We find that…

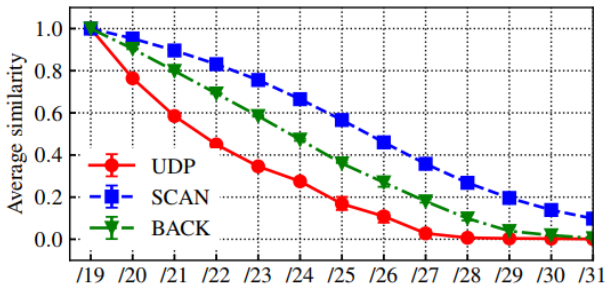
Darknets are (quite) similar!

- The contacted **ports** are the same
- **UDP** traffic more similar than **TCP** in terms of sources



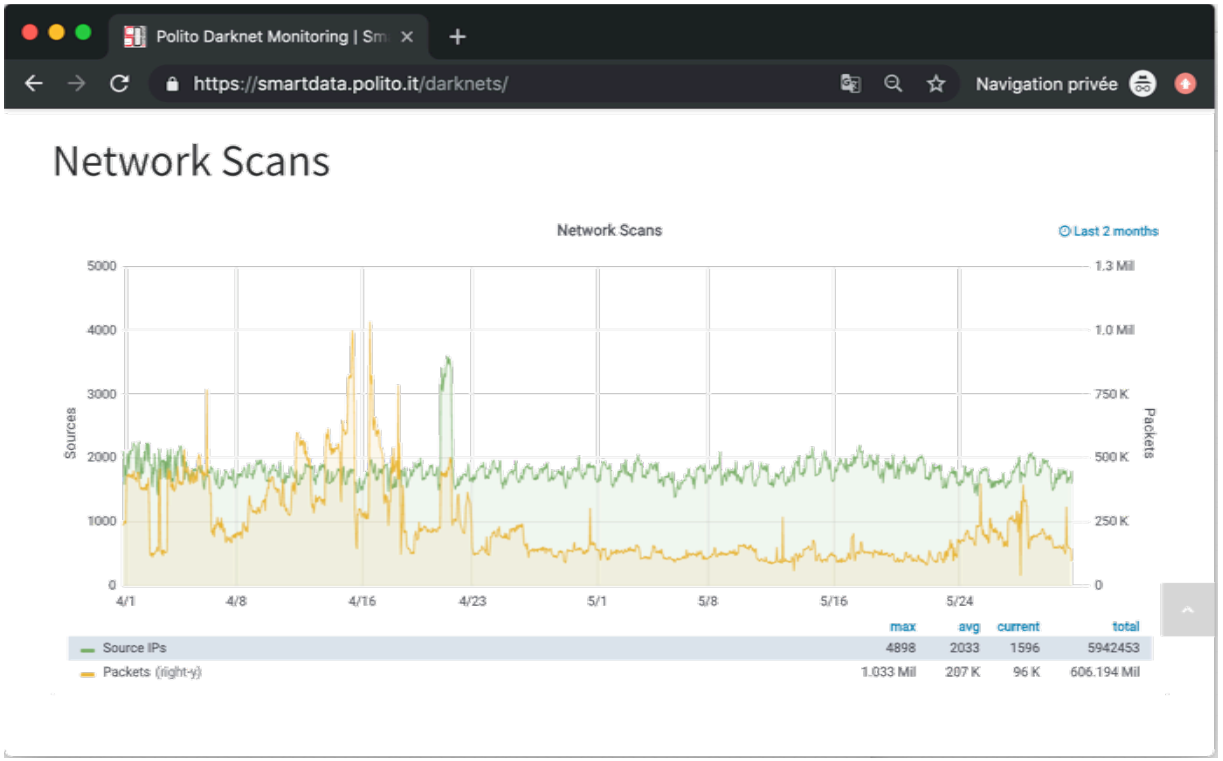
The size matters!

- TCP Scans can be found even with small darknets
- Specific events need large darknets to be understood
  - e.g., backscattering traffic resulting from spoofed source addresses



Data visualization

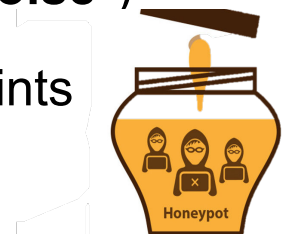
Real-time monitoring framework a:
https://smartdata.polito.it/darknets/



Next steps

- Extend the darknet space to better evaluate **Internet Background Radiation**
  - In PoliTo/GARR
  - Coupling it with the study of worldwide spread sensor greynets (such as **Greynoise**[1])

- Extract and better characterize anomaly fingerprints with the usage of **honeypots**



- Compare our traffic with data **passively collected on a production network**

  - Build models to **automatically characterize anomalous traffic** by means of Machine Learning techniques

[1] https://greynoise.io/

Thank you for your attention