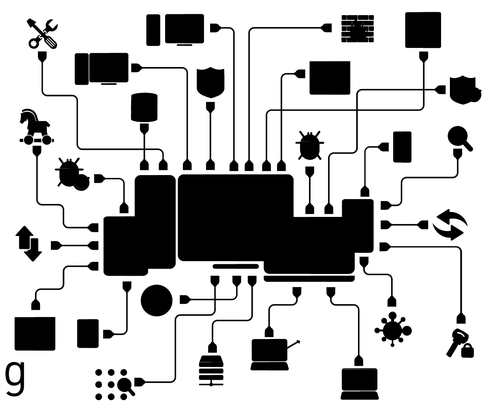
Generative adversarial networks for cybersecurity applications

Francesca Soro (francesca.soro@polito.it)
Supervisor: Prof. Marco Mellia
Department of Electronics and Telecommunications Engineering

Motivation and background

- Nowadays **network traffic** scenario is experiencing a continuous growth in both **volume** and **complexity**
- New **threats** and **anomalies** are generated everyday, and the design of **efficient cybersecurity systems** is a problematic task
- **Automatic detection** of zero-day attacks is almost impossible given the **lack of training data**
- The current state of being calls for a **big-data approach to** extract relevant features and behaviours



Addressed problem

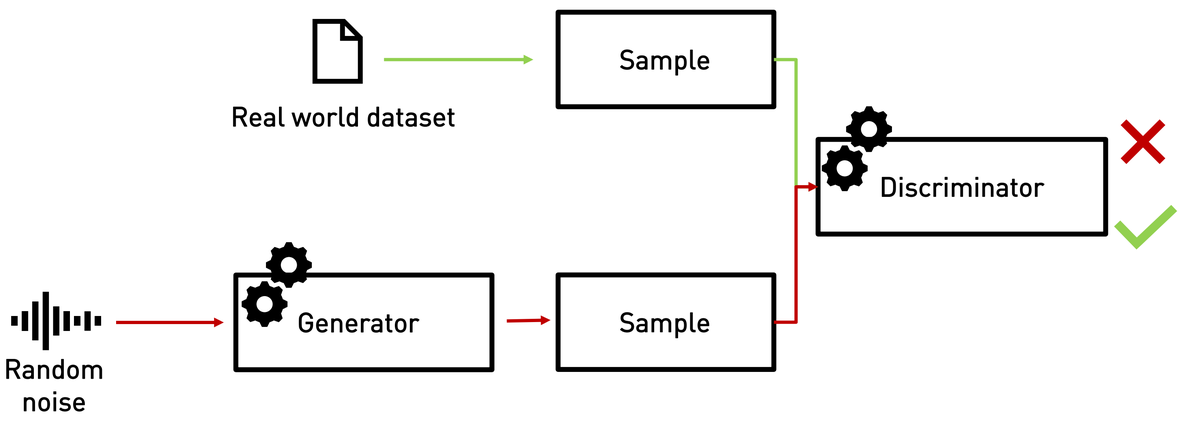
- **Machine learning algorithms** are able to correctly classify well known attacks for which labeled datasets are available
- Existing solutions are **knowledge-based** systems adopting **signature-based** or **novelty detection**[2]
- **Heuristics** for recognizing network anomalies and threats are implemented[3]

No **ground truth** available for newly created threats and unknown attacks

Adopted methodology

Model

The adoption of **Generative Adversarial Network**[1] models can help in the generalization of a sample labeled dataset to an artificially richer one.



Dataset

- Basic features extracted from **darknet**[4,5] packet captures
- Attack signatures provided by **Bro**[6] and **Snort**[7]
- IP geolocation database from **Maxmind GeoIP2**[8]

PROTOCOL	FEATURE
IP	VERSION
	SRC ADDRESS
	DST ADDRESS
	PROTOCOL
UDP	SRC PORT
	DST PORT
TCP	SRC PORT
	DST PORT
	FLAGS

Results

Protocol distribution in darknet traffic (3 weeks):

TCP	UDP	ICMP	Others
94,3%	5,23%	0,34%	0,13%

Nature of darknet TCP traffic:

SYN (Scan)	Backscattering	Misconfigurations
92,38%	1,61%	0,34%

Top-3 countries per origin flow:

Russia	Seychelles	Netherlands
48,54%	12,88%	9,58%

Conclusions and future work

- A simple yet systematic **traffic profiling** allows a deeper understanding of the nature of attacks
- Such output will lead to a deeper **threat signature analysis**, detecting correlations and co-occurrences among threats
- Deriving an **association rule** among threats allows a first labelling of the traffic dataset, to be further generalized with GANs

References

1. Goodfellow, Ian, et al. "Generative adversarial nets." *Advances in neural information processing systems*. 2014.
2. A. Sperotto, G. Schaffrath, R. Sadre, C. Morariu, A. Pras, B. Stiller, "An Overview of IP Flow-Based Intrusion Detection," in IEEE Communications Surveys Tutorials.
3. R. Fontugne, et al., "MAWILab: Combining Diverse Anomaly Detectors for Automated Anomaly Labeling and Performance Benchmarking", in ACM CoNEXT, 2010
4. Fachkha, Claude, et al. "Investigating the dark cyberspace: Profiling, threat-based analysis and correlation." *Risk and Security of Internet and Systems (CRiSIS), 2012* .
5. Moore, David, et al. *Network telescopes: Technical report*. Cooperative Association for Internet Data Analysis (CAIDA), 2004.
6. https://www.snort.org
7. https://www.bro.org/
8. https://maxmind.com