Machine Learning for Defensive Cybersecurity

Azqa Nadeem

PhD candidate

Cyber Analytics Lab

Lorentz workshop: Security & Privacy Day

TU Delft

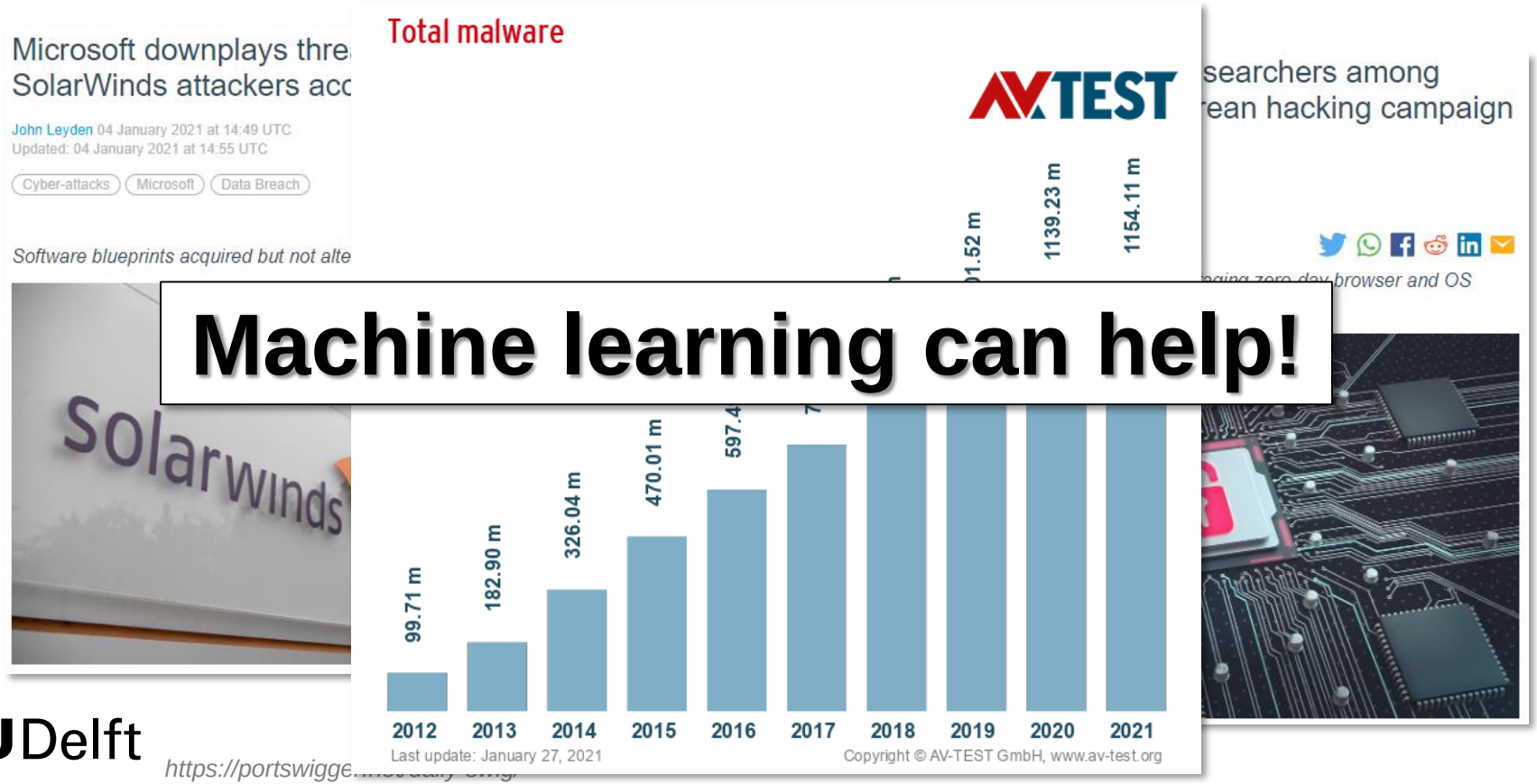
02 June 2021

Cyber Analytics Lab

> whoami

- 3rd year PhD candidate
  - Under supervision of Dr. Sicco Verwer
  - *Explainable sequential ML for network security*
- Security lecturer at TU Delft
- Co-organizer Cyber Security Next Generation (CSng) workshop
  - https://csng.nl/

Current state of security



Machine learning can help!

Facets of defensive cybersecurity

Spam/Malware detection

Automated code patching

Cyberattack detection
(IoT/Mobile/Kernel)

Access control

Attacker behavior profiling

Attacker modelling

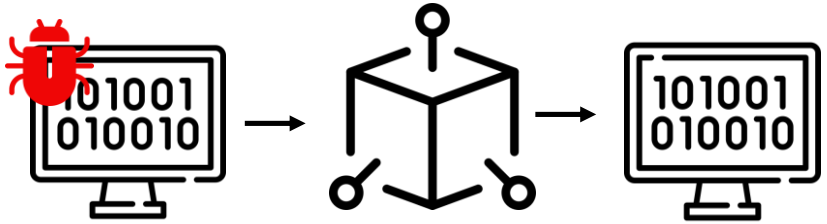
Forensic analysis

- *Offensive security applications*
  - *Crafting malware, hardware attacks, …*

A few use cases

[1] Automated bug-fix patches

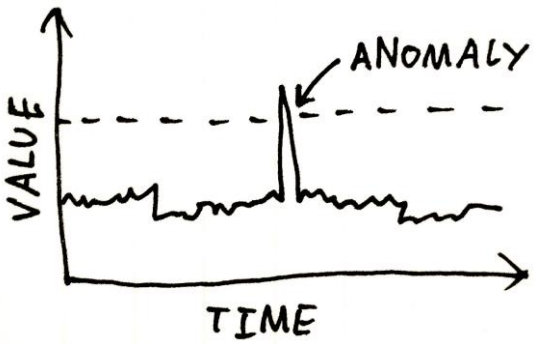
- *<u>Goal:</u> Automatically learn bug-free code variant*
- via Neural machine translation
- AST-level operations
  - Code structure, try-catch, casting, …



Tufano, Michele, et al. "An empirical study on learning bug-fixing patches in the wild via neural machine translation." ACM Transactions on Software Engineering and Methodology (TOSEM) 28.4, 2019.

[2] Industrial Control System defense

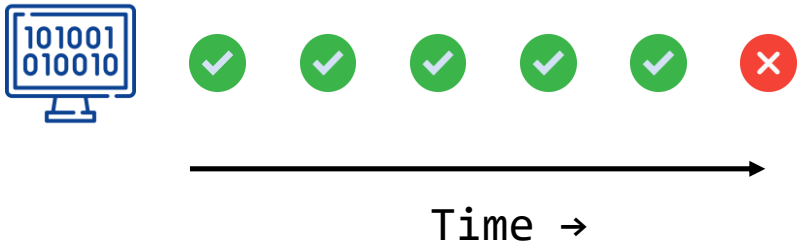
- *<u>Goal:</u> Detect irregular behavior in a water treatment plant*
- Learn ‘normal’ behavior
- Behavior localization via graphical models



Lin, Qin, et al. "TABOR: A graphical model-based approach for anomaly detection in industrial control systems." Proceedings of the 2018 on Asia Conference on Computer and Communications Security. 2018.
https://towardsdatascience.com/generating-critical-scenarios-using-anomaly-detection-f25e67e0553b

[3] Proactive malware detection

- *<u>Goal:</u> Alerting user of impending exposure to malicious content*
- Monitoring users' HTTP traffic
  - Mobile cellular network
- Predicting impending exposure
  - Via deep neural nets



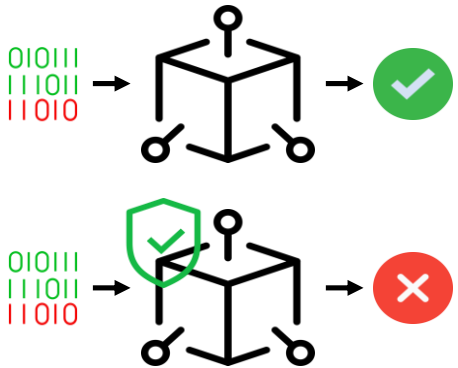
Sharif, Mahmood, et al. "Predicting impending exposure to malicious content from user behavior." Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security. 2018.

[4] Proactive malware detection

- *<u>Goal:</u> Adversarially robust malware detectors*

- Craft malware and learn from it
  - Greedy random multi-bit search approach

- Malware detection competitions
  - `Attackers` craft malware
  - `Defenders` learn robust detectors



Verwer, Sicco, et al. "The Robust Malware Detection Challenge and Greedy Random Accelerated Multi-Bit Search." Proceedings of the 13th ACM Workshop on Artificial Intelligence and Security. 2020.

[5] Malware behavior profiling

- *<u>Goal:</u> Automated behavior discovery of malware*

- Find behavior groups in network traffic
  - Clustering approach

- Malware profiles via cluster membership

Behavior profile

- Connects with C&C
- Opens backdoors
- Persistent

vs.

Label

ZeuS

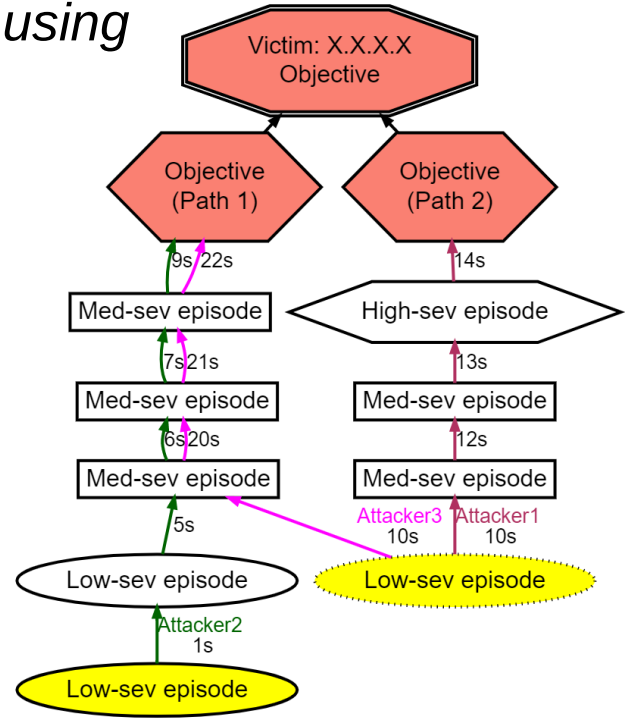
Nadeem, Azqa, et al. "Beyond Labeling: Using Clustering to Build Network Behavioral Profiles of Malware Families." Malware Analysis Using Artificial Intelligence and Deep Learning. Springer, 2021.

[6] Attacker behavior profiling

- *<u>Goal:</u> Automatically discover attacker strategies using intrusion alerts*

- Discover attacker strategies
  - Suffix-based probabilistic deterministic finite automaton

- Extract alert-driven attack graphs
  - Per-objective, per-victim basis



Nadeem, Azqa, et al. "Alert-driven Attack Graph Generation using S-PDFA." IEEE Transactions on Dependable and Secure Computing, (Submitted)

Words of Caution & Open questions

(Caution!) Machine learning is not a silver bullet [1/5]

- Cannot blindly apply ML to cybersecurity
  - Address unique problems

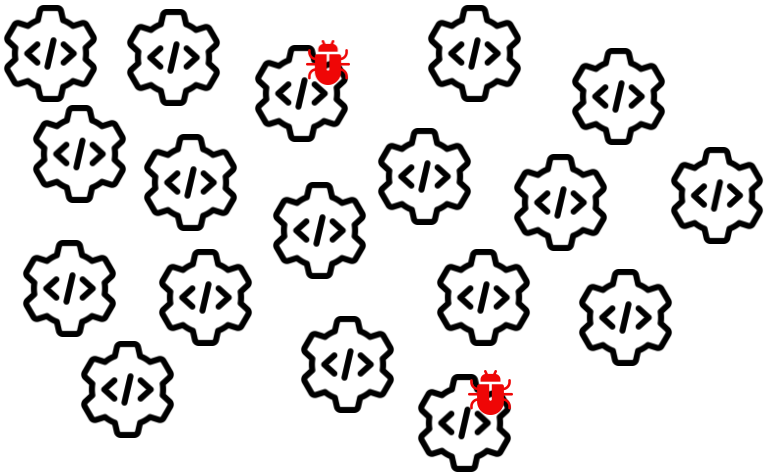
- Do not throw data in black-box
  - Ethical considerations

(Caution!) More benign than malicious [2/5]

- Security data has class imbalance
- Labels are often noisy

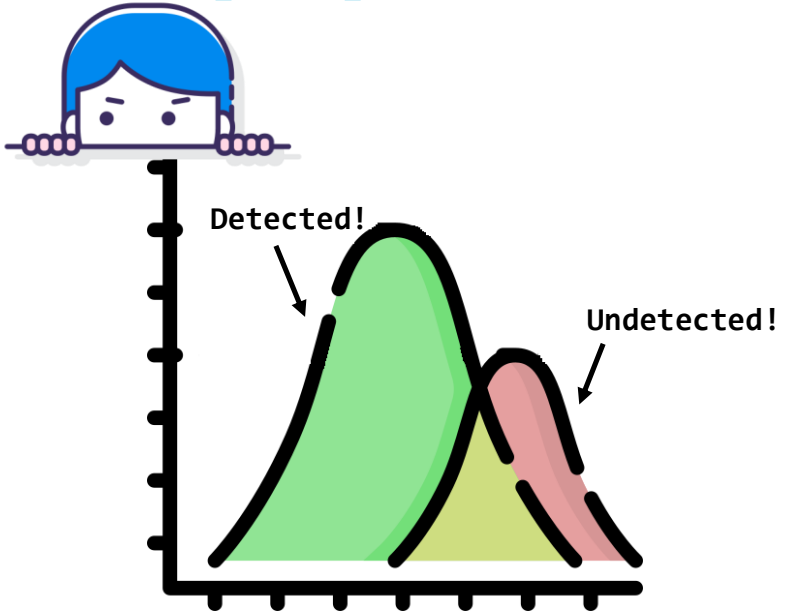
- Unrealistic class distribution
  - Bias in data → bias in models

- Real-world performance evaluation?



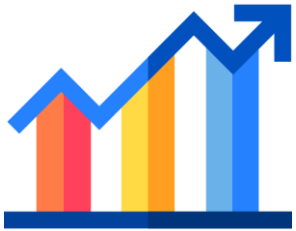
(Caution!) Landscape is adversarial [3/5]

- Attackers hide, malware evades detection
- ML cannot detect all evasion attempts!
- Representative dataset is required
- How to incorporate continual learning?
- How to design robust systems?



(Caution!) Know what to evaluate [4/5]

- Be mindful of evaluation metrics
  - Precision, Recall, AUC, F1 score …
  - Accuracy in imbalanced datasets



- Performance metrics ≠ improved security

- How to operationalize security evaluation?
- Value explainability over accuracy?

(Caution!) Know the limitations of ML [5/5]

- Can find patterns faster than humans
  - But is also really stupid
- Cannot replace human intelligence
  - Trade-off between automation and explainability

- Incorporate human intelligence in a smarter way?
- Build trust in ML systems?

https://hackernoon.com/human-intelligence-or-artificial-intelligence-we-need-both-if7w32b2

Summary: Open questions

- How to design robust systems that evolve with the threat landscape?
- What role does explainability play in designing and evaluating ML systems?
- How to build trust in ML systems?
- How to fill the gap between academia and industry to allow real-world performance evaluation?
- How to operationalize security evaluation?

Thank you!

- How to design robust systems that evolve with the threat landscape?

- What role does explainability play in designing and evaluating ML systems?

- How to build trust in ML systems?

- How to fill the gap between academia and industry to allow real-world performance evaluation?

- How to operationalize security evaluation?

azqa.nadeem@tudelft.nl @azqa_nadeem https://cyber-analytics.nl/

TU Delft