BLUEHAT SHANGHAI 2019

Advances in Machine Learning at Microsoft Threat Protection

5/29/2019

Christian Seifert, Principal Researcher

Security Research Superheroes

Security Research Experts

+

Machine Learning Systems

=

Security Research Superheroes

Microsoft Defender ATP

Built-in. Cloud-powered.

PRE-BREACH

SMARTSCREEN

Protect against malicious URLs and downloads

ENDPOINT PROTECTION

Protect against all types of emerging threats

GOAL	FP TOLERANCE	SCOPE
Block at first sight	**Low**	**www, files, "fileless"**

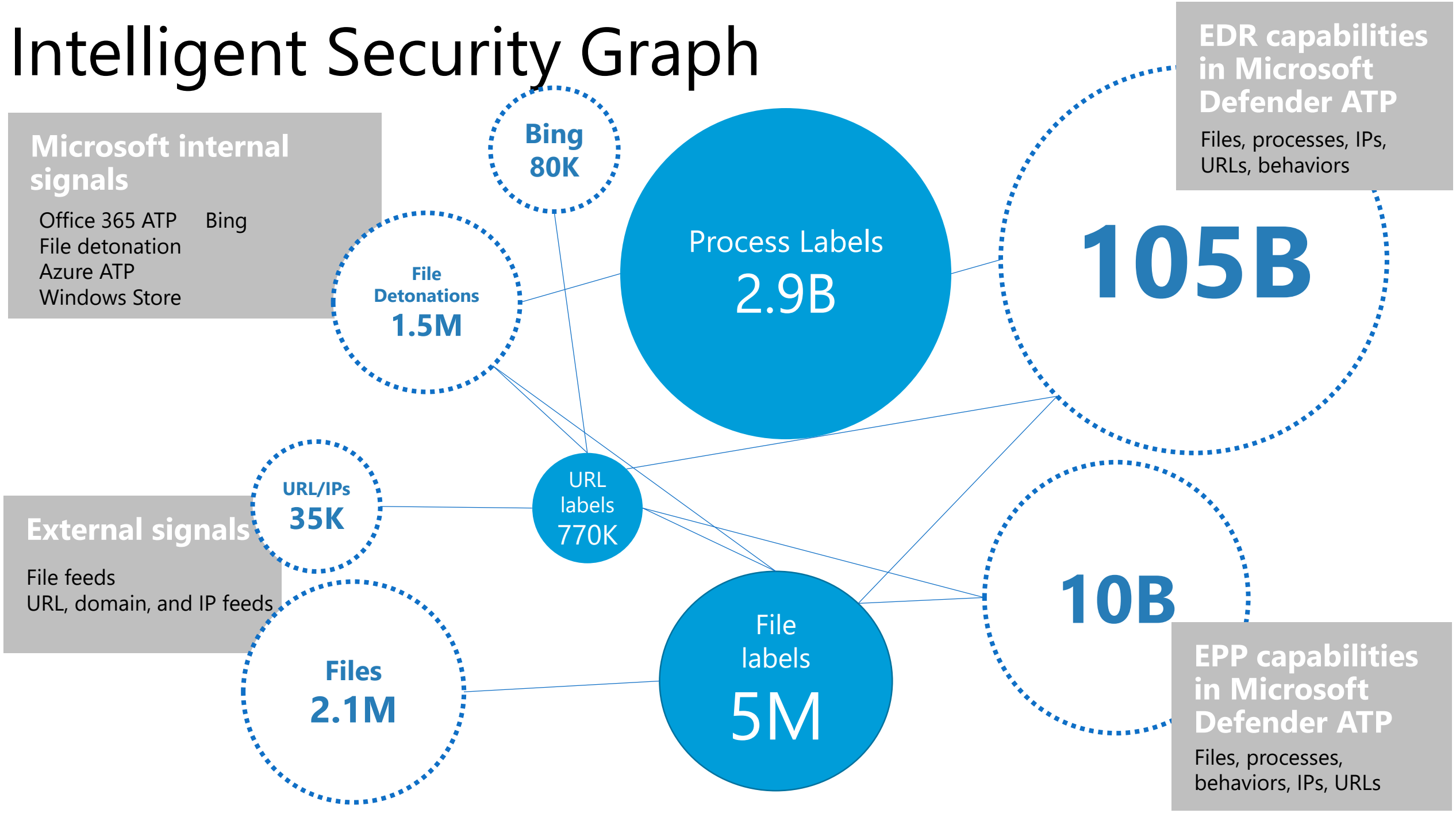
POST-BREACH

ENDPOINT DETECTION & RESPONSE

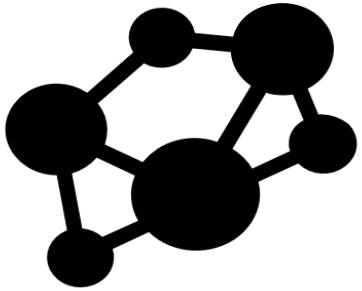
Detect, investigate, and respond to advanced attacks

GOAL	FP TOLERANCE	SCOPE
Alert on all possible breaches	**Moderate**	**Same as pre-breach + cross-service + bad actor behaviors**

Intelligent Security Graph



AI diversity



Supervised learning

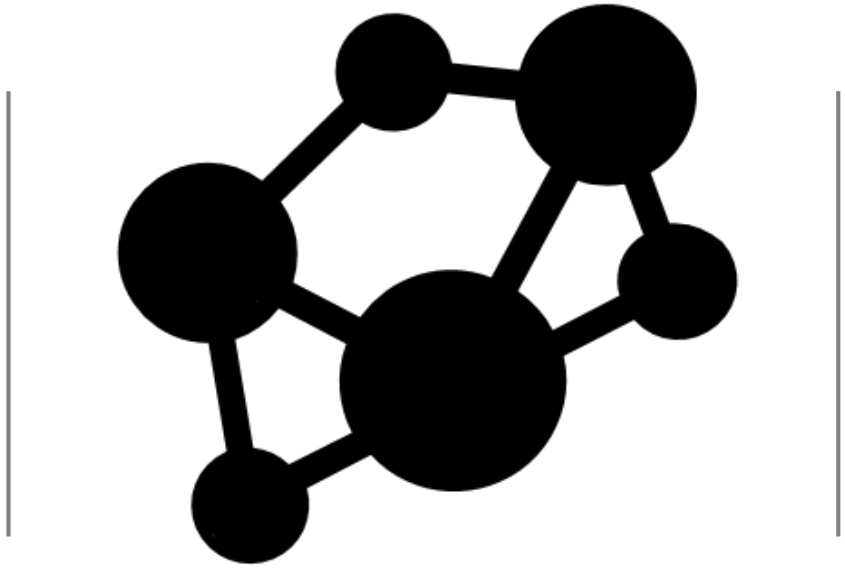
- Fast learners
- Deep learning

Unsupervised learning

- Anomaly detection
- Embeddings

New learners/ approaches

- Active learning
- Homomorphic encryption



Deep learning in MDATP

Case Study 1: Deep learning for malicious PowerShell detection

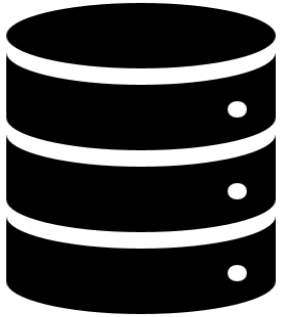
Why PowerShell?

Why Deep learning?

Case Study 1: Deep learning for malicious PowerShell detection

1
Learn the language
PowerShell

2
Classification task
Clean/Malicious



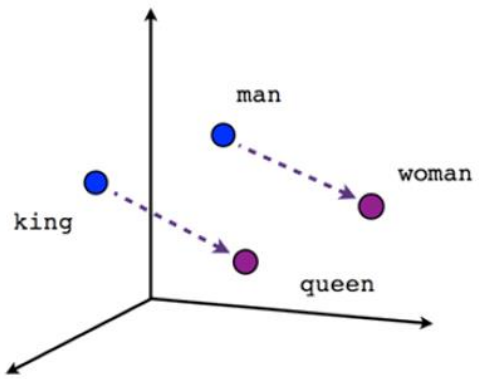
Big **unlabeled** dataset

Embedding

Small **labeled** dataset

Contextual Embedding - **English** (NLP)

- Tokens $\longrightarrow$ Vectors ($\mathbb{R}^n$)
- king – man + woman = queen

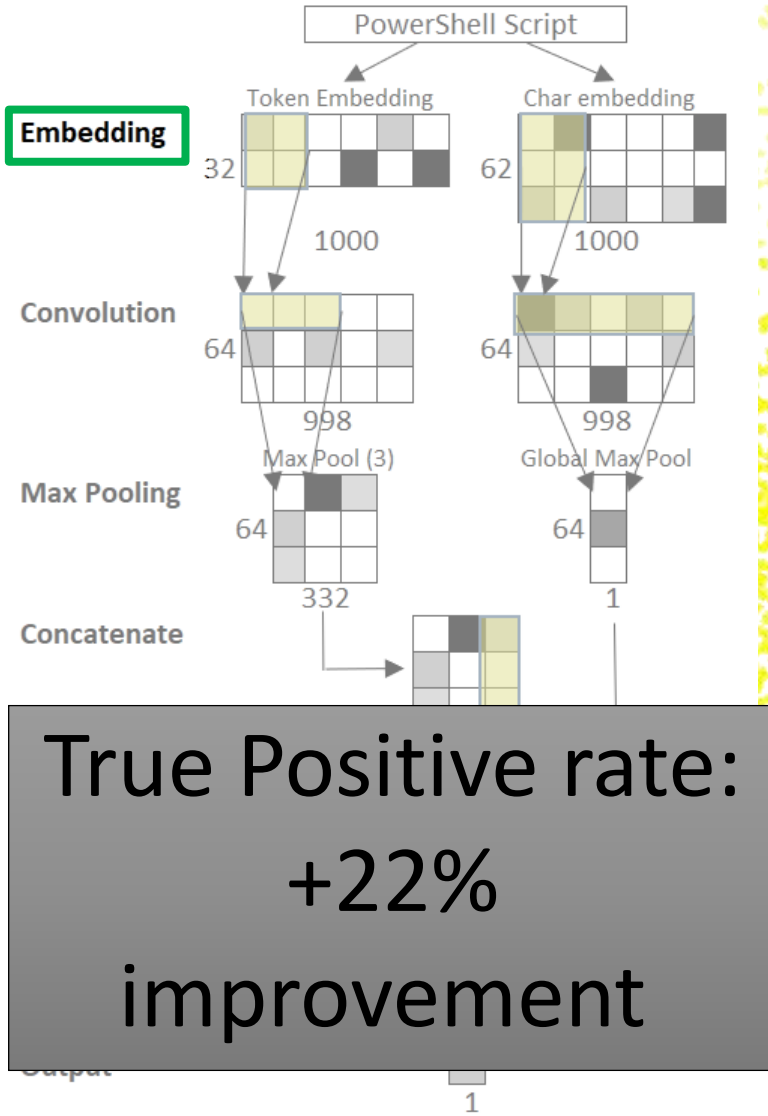


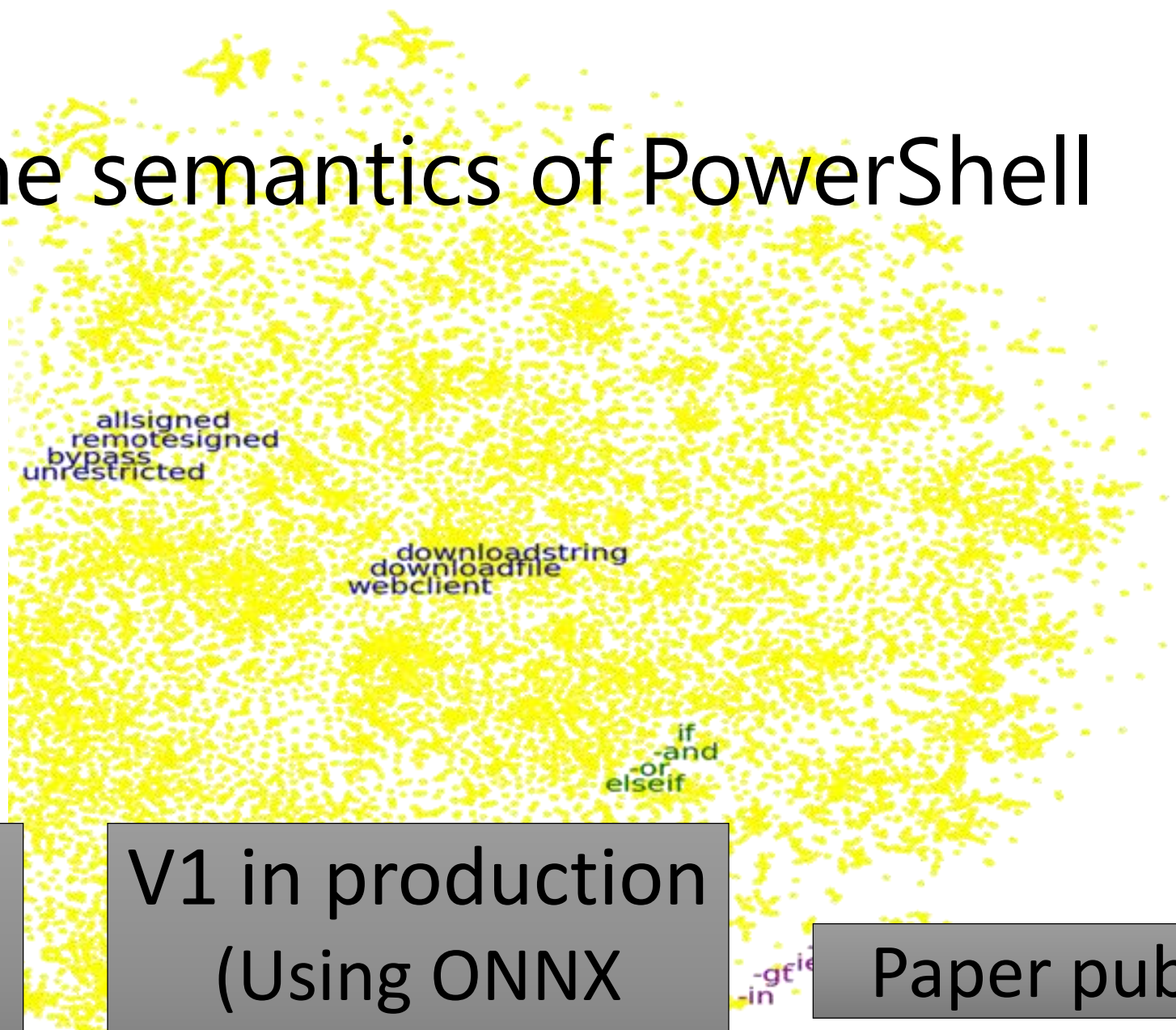
Male-Female

Contextual Embedding - PowerShell

- High - $false + $true = Low
- ‘Export-CSV’ - $csv + $html = ‘ConvertTo-Html’
- ‘Get-Process’ - $processes + $services = ‘Get-service’

Learning the semantics of PowerShell





True Positive rate: +22% improvement

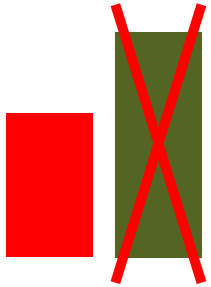
V1 in production (Using ONNX and ML.Net)

Paper published

Supervised learners

Case Study 2: Monotonic model

Problem: Adversaries!



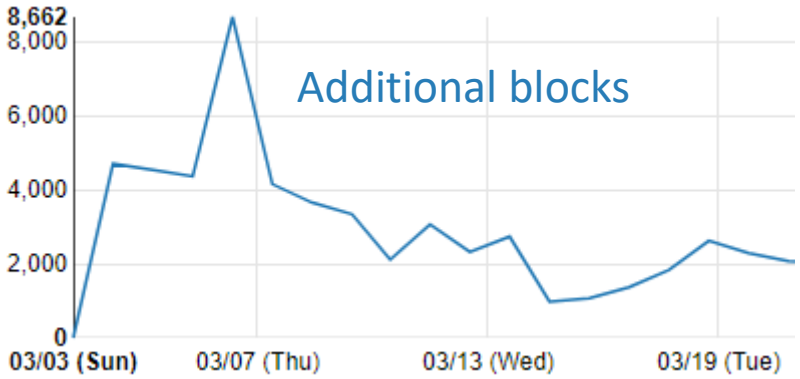
malicious features ≯ clean features

Solution: Monotonic approach only weights malicious features

Incer, Inigo, et al. "Adversarially robust malware detection using monotonic classification." *Proceedings of the Fourth ACM International Workshop on Security and Privacy Analytics*. ACM, 2018.

Over the past month…

0.25M additional blocks



Monotonic model

REUTERS

Business Markets World Politics TV Mor

TECHNOLOGY NEWS MARCH 26, 2019 / 8:20 AM / UPDATED 6 HOURS AGO

Norsk Hydro's initial loss from cyber attack may exceed $40 million

Nerijus Adomaitis 3 MIN READ

OSLO (Reuters) - Norwegian aluminum maker Norsk Hydro may have lost more than $40 million in the week that followed a cyber attack that paralyzed parts of its operations, and a full recovery of IT systems will take weeks or more, the company said.

Sha256	c97d9bbc80b573bdeeda3812f4d00e5183493dd0d5805e2508728f65977dda15
Determination	Malware
Signer	ALISA LTD
Age	0
Prevalence	0
FilePredictV4_Malware	34.4%
FilePredictV4_Clean	27.7%
FilePredictV4_Malware_Monotonic	78.8%
WinningRuleName	FilePredictV4_Malware_Monotonic

Monotonic model ignores clean features (certificates with positive rep)

Signature Info

Signature Verification

A certificate was explicitly revoked by its issuer.

File Version Information

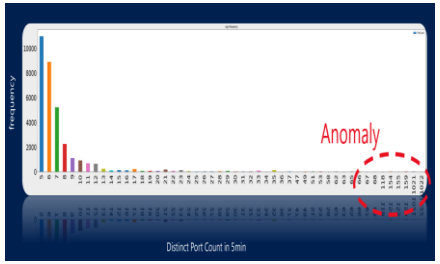
Copyright	Copyright (C) ALISA LTD 2019
Product	Service tgytutrc
Description	Background Tasks Host
Original Name	tgytutrc
Internal Name	tgytutrc
File Version	1.5.1.0
Date Signed	2:44 PM 3/21/2019

Signers

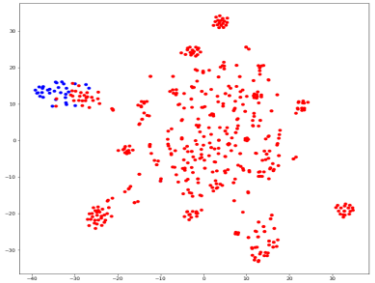
- ALISA LTD
- Sectigo RSA Code Signing CA
- USERTrust Secure™

Anomaly detection

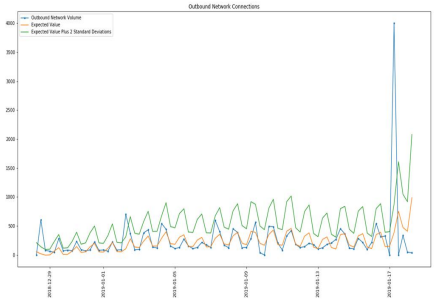
Statistical anomaly detection



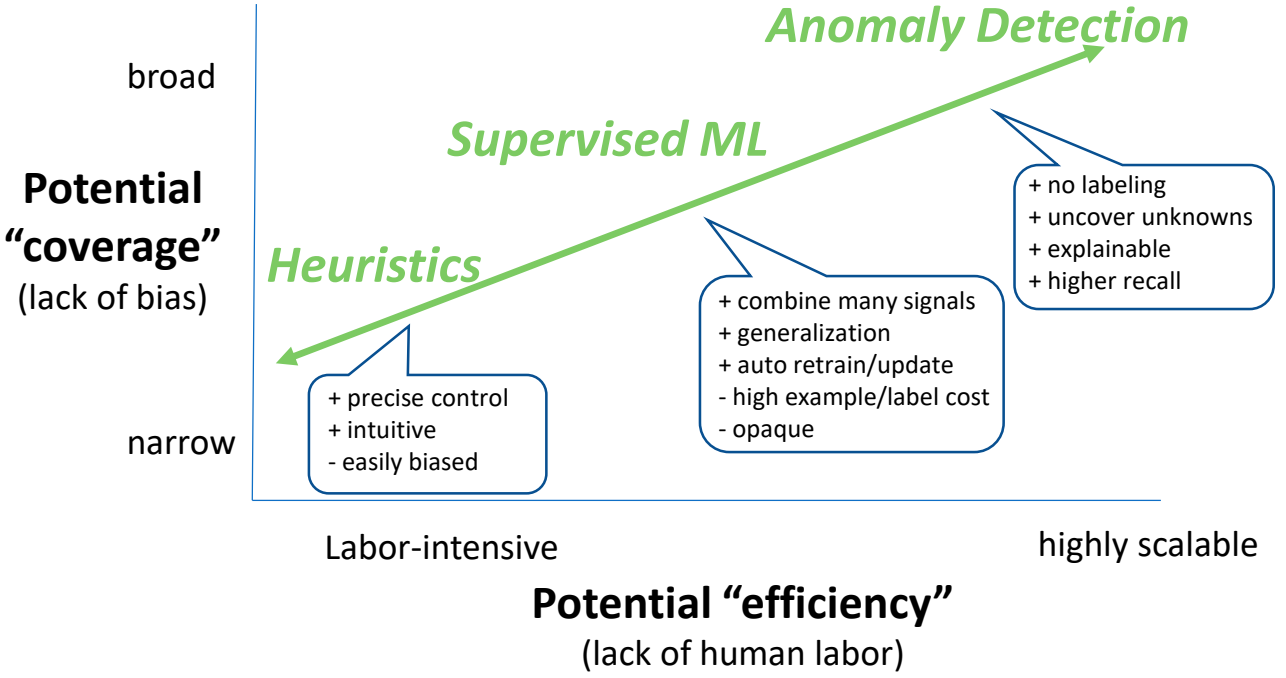
Categorical



Population



Time-Series



Case Study 3: Port scan anomaly detection

Identify internal attack reconnaissance

- Both vertical and horizontal port scanning detector
- 90%+ precision

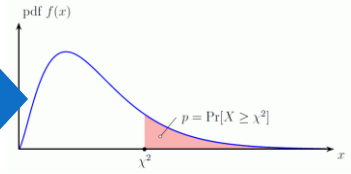
Feature
Distinct Port Count
Distinct IP Count
Connection Count
Failure connection rate
Port Importance
Is scanned by common scan process
Is scanned by sys process
Is process valid signed

P(Observed port count) = X_1

⋮

P(Is process valid signed) = X_8

$$X = -2\sum_{i=1}^{8}\log(X_i)$$



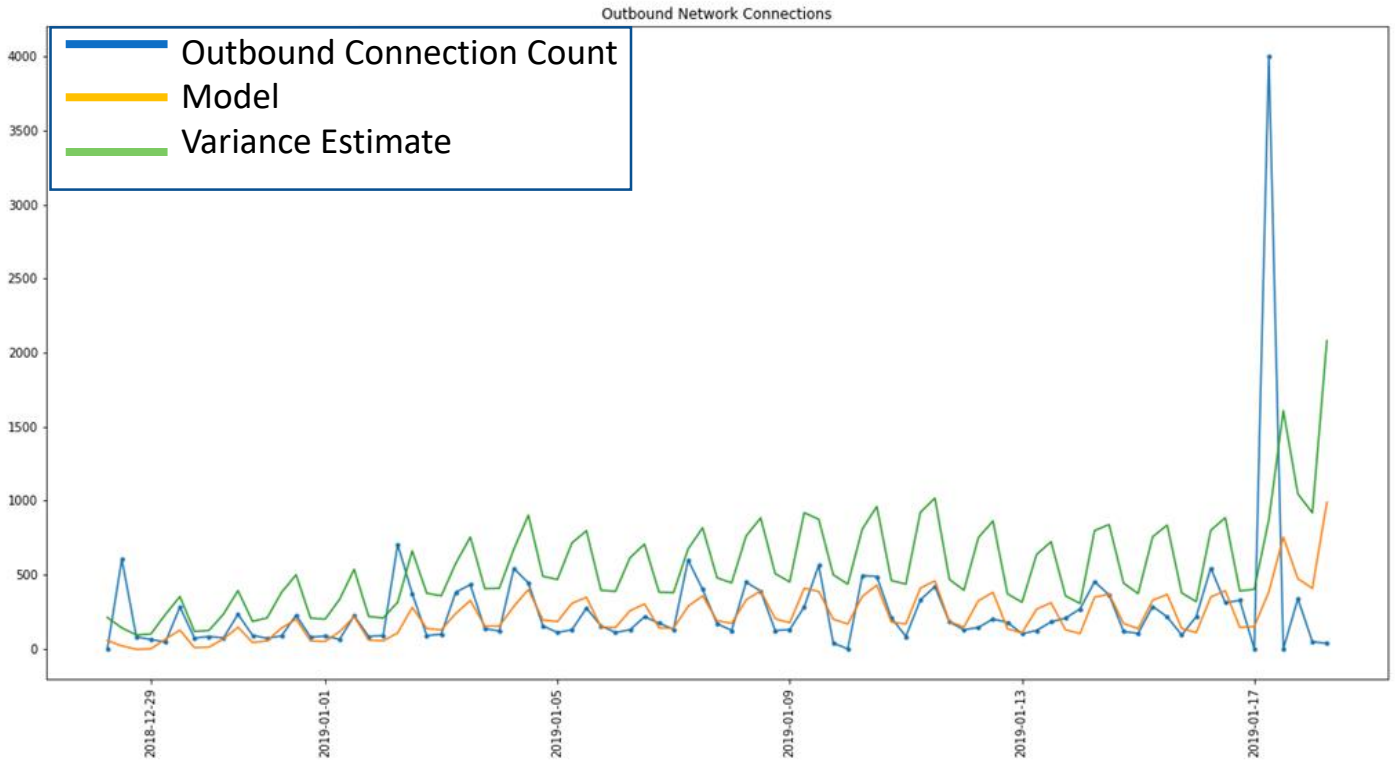
Extract features | Calculate rarity | Combine rarity | Threshold for detection

Identified Emotet and Monero ransomware malware.

Case Study 4: Brute force time series anomaly detection

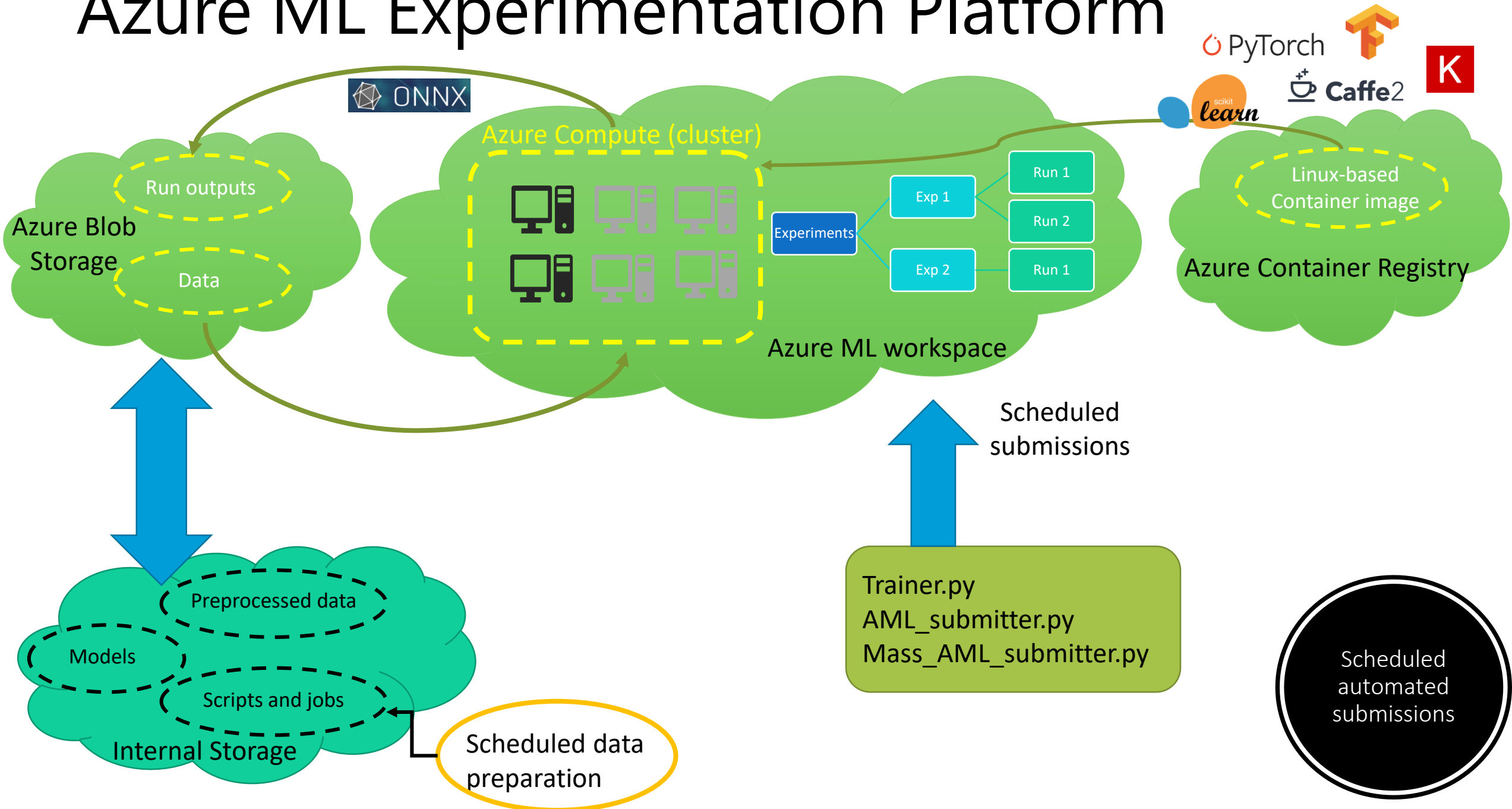
Model of login behavior

- Sensitive to time of day, day of week
- Custom self-learning model per machine
- Precision of 94%



Tools and platform

Azure ML Experimentation Platform



Tools of the trade

Experimentation & training

Models

Production scoring

Current

Cosmos

tLC

AE

tLC

C#

JS

tLC

Evolution

python

K

Keras

Microsoft CNTK

TensorFlow

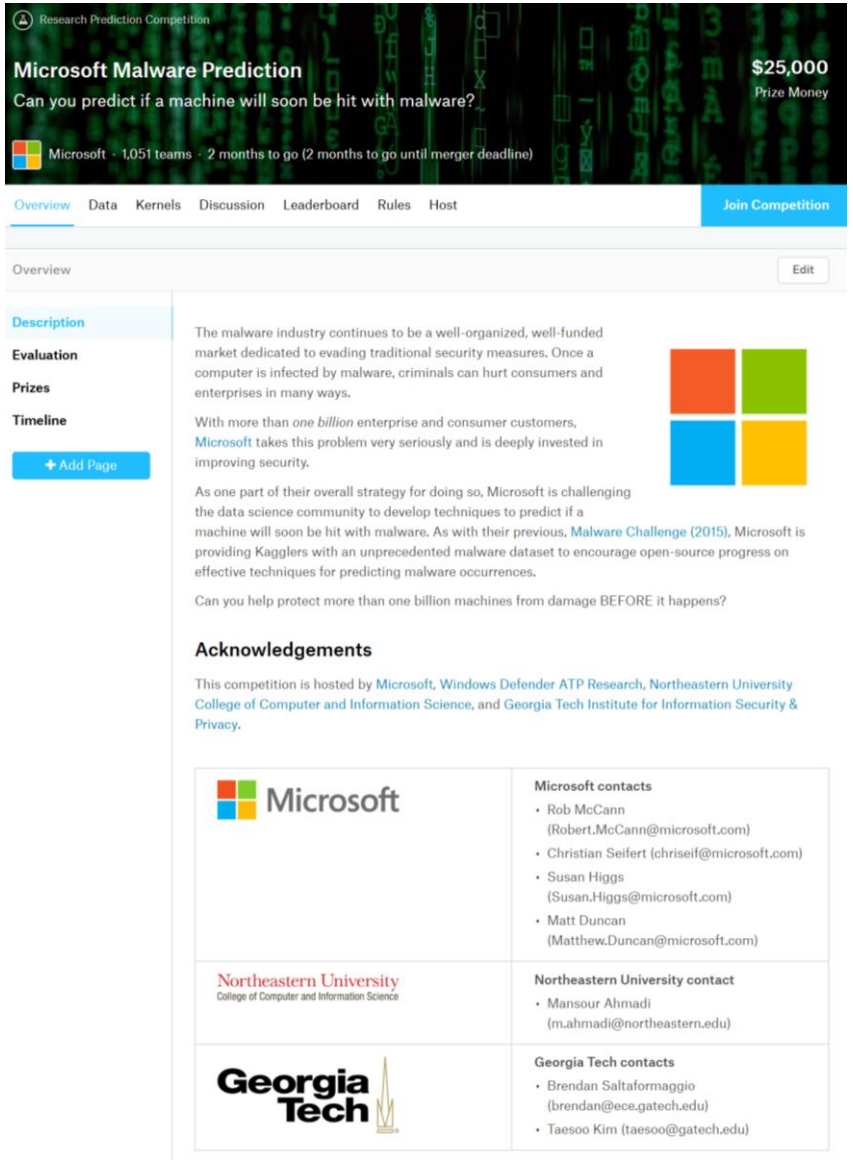
ONNX

C#

ML.NET

Community

Kaggle data science competition



New 2018-19 Competition: Anticipate malware based on machine state

- Effort started w/ internship, competition running 12/13/18 - 3/13/19
- https://www.kaggle.com/c/microsoft-malware-prediction

Collaborative!

- Academic partners (Northeastern, Georgia Tech, UW, UW Tacoma)
- Microsoft partners (ILDC, MSRA)
- >2,426 teams & >300 forum discussion threads, > 3,000 posts
- Winning submissions are being reviewed and hold promise for product impact
  - Durability over time is a focus

Newsworthy!

- Our blog
- Academic/MS partner ann.
- ZDNet, Tom's Hardware
- Bleeping Computer, Neowin, ...

Internships

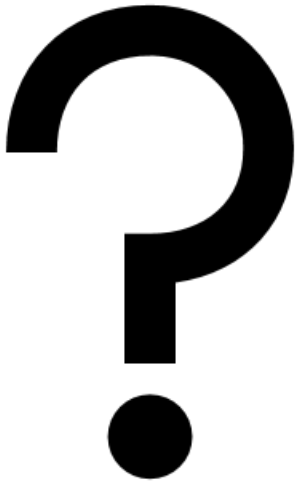
Undergraduate and graduate internships

- 12 weeks paid internship
- Access to real-world attack data
- Work on cool problems
- Often partners with Microsoft Research

Published papers

- Danny Hendler, Shay Kels, and Amir Rubin. "Detecting malicious PowerShell commands using deep neural networks." ACM, 2018.
- Jack W. Stokes, De Wang, Mady Marinescu, Marc Marino, Brian Bussone. “Attack and Defense of Dynamic Analysis-Based, Adversarial Neural Malware Detection Models.” MILCOM, 2018.
- Yehonatan Cohen, Danny Hendler, and Amir Rubin. "Detection of malicious webmail attachments based on propagation patterns." Knowledge-Based Systems, 2018.
- Rakshit Agrawal, Jack W. Stokes, Mady Marinescu, Karthik Selvaraj. “Robust Neural Malware Detection Models for Emulation Sequence Learning.” MILCOM, 2018.
- Md Amran Siddiqui, Jack W. Stokes, Christian Seifert, Evan Argyle, Robert McCann, Joshua Neil, Justin Carroll. “Detecting Cyber Attacks Using Anomaly Detection with Explanations and Expert Feedback.” ICASSP, 2019.
- Rakshit Agrawal, Jack W. Stokes, Mady Marinescu, Karthik Selvaraj. “Neural Sequential Malware Detection with Parameters.” ICASSP 2018
- Rakshit Agrawal, Jack Stokes, Karthik Selvaraj, Mady Marinescu. "Attention In Recurrent Neural Networks For Ransomware Detection." ICASSP 2019

Contact:
chriseif@microsoft.com
Linkedin
cseifert



Questions