



ENJOY SAFER TECHNOLOGY™

AI in cybersecurity: A tool for defenders or a weapon for attackers?

Juraj Jánošík, Head of AI/ML team

Ondrej Kubovič, Security Awareness Specialist



Artificial intelligence?



Artificial intelligence

Generally intelligent machine that can learn and make decisions independently, based on inputs from its environment – without human supervision

Machine Learning

Machine Learning

Field of computer science that gives machines the ability to find patterns in vast amounts of data, by sorting them and acting on the findings.



1950s

Machine Learning

Collaborative filtering optimization objective

→ Given $x^{(1)}, \dots, x^{(n_m)}$, estimate $\theta^{(1)}, \dots, \theta^{(n_u)}$: ~~$x \in \mathbb{R}^n$~~ ~~$\theta \in \mathbb{R}^n$~~ ~~$x \in \mathbb{R}^{n \times 1}$~~ ~~$\theta \in \mathbb{R}^n$~~

$$\min_{\theta^{(1)}, \dots, \theta^{(n_u)}} \left[\frac{1}{2} \sum_{j=1}^{n_u} \sum_{i:r(i,j)=1} ((\theta^{(j)})^T x^{(i)} - y^{(i,j)})^2 + \frac{\lambda}{2} \sum_{j=1}^{n_u} \sum_{k=1}^n (\theta_k^{(j)})^2 \right]$$

→ Given $\theta^{(1)}, \dots, \theta^{(n_u)}$, estimate $x^{(1)}, \dots, x^{(n_m)}$:

$$\min_{x^{(1)}, \dots, x^{(n_m)}} \left[\frac{1}{2} \sum_{i=1}^{n_m} \sum_{j:r(i,j)=1} ((\theta^{(j)})^T x^{(i)} - y^{(i,j)})^2 + \frac{\lambda}{2} \sum_{i=1}^{n_m} \sum_{k=1}^n (x_k^{(i)})^2 \right]$$

Minimizing $x^{(1)}, \dots, x^{(n_m)}$ and $\theta^{(1)}, \dots, \theta^{(n_u)}$ simultaneously:

$$J(x^{(1)}, \dots, x^{(n_m)}, \theta^{(1)}, \dots, \theta^{(n_u)}) = \frac{1}{2} \sum_{(i,j):r(i,j)=1} ((\theta^{(j)})^T x^{(i)} - y^{(i,j)})^2 + \frac{\lambda}{2} \sum_{i=1}^{n_m} \sum_{k=1}^n (x_k^{(i)})^2 + \frac{\lambda}{2} \sum_{j=1}^{n_u} \sum_{k=1}^n (\theta_k^{(j)})^2$$

→ $\min_{\substack{x^{(1)}, \dots, x^{(n_m)} \\ \theta^{(1)}, \dots, \theta^{(n_u)}}} J(x^{(1)}, \dots, x^{(n_m)}, \theta^{(1)}, \dots, \theta^{(n_u)})$ $\theta \rightarrow x \rightarrow \dots$

2000s



Fit to screen

Run data

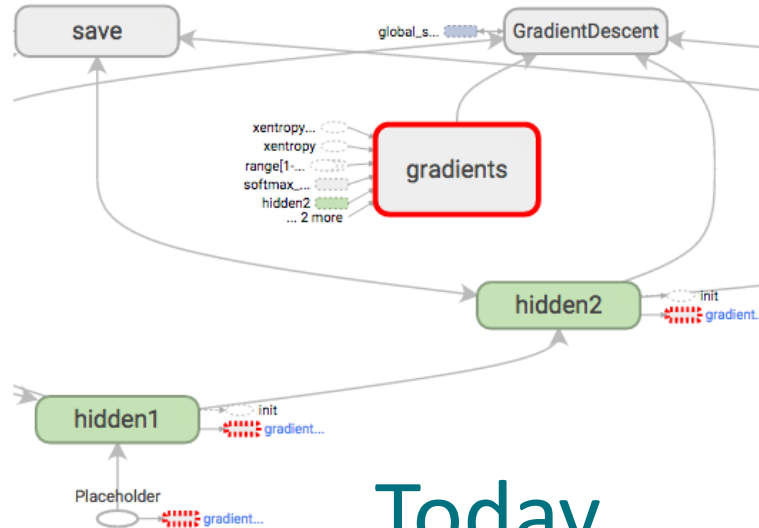
Upload

Color Structure

color: same substructure
gray: unique substructure

Graph (* = expandable)

- Namespace*
- OpNode
- Unconnected series*
- Connected series*
- Constant
- Summary
- Dataflow edge
- Control dependency edge
- Reference edge



gradients

Subgraph: 81

nodes

Attributes (0)

Inputs (7)

- xentropy_mean
- xentropy
- range[1-2]
- softmax_linear
- hidden2
- hidden1
- Placeholder

Outputs (1)

- GradientDescent

Today

How AutoML Vision^{BETA} works

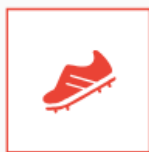
1 Upload and label images

2 Train your model

3 Evaluate



Handbag



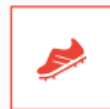
Shoe



Hat



AutoML Vision



Google



Google Search

I'm Feeling Lucky

Google offered in: [slovenčina](#)

NETFLIX

Verizon LTE

3:42 PM

37%



NETFLIX



TV Dramas





UBER

UBER
ADVANCED TECHNOLOGIES CENTER

UBERATC.COM/CAR

ADVANCED TECHNOLOGIES CENTER

UBER

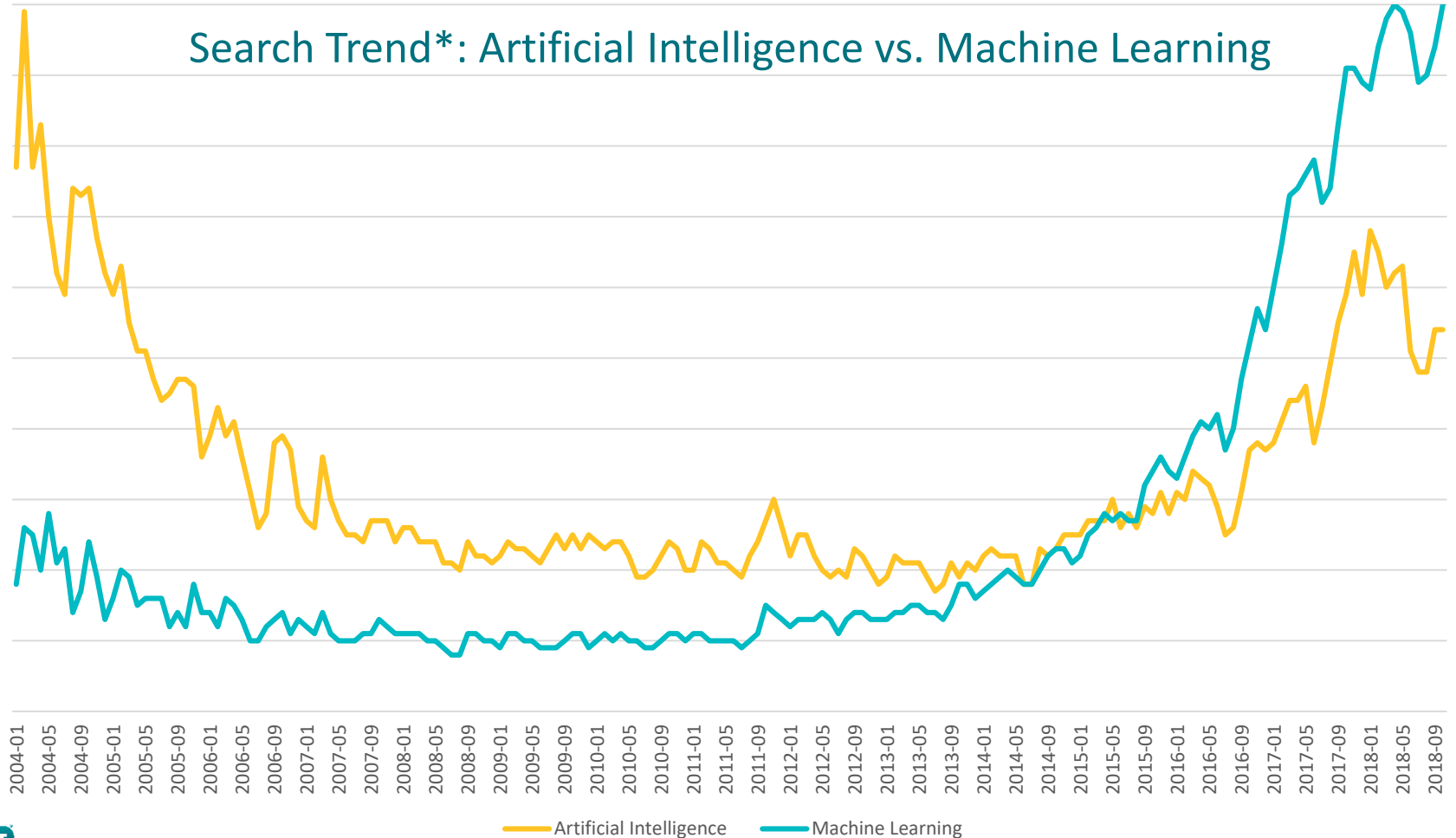
UBERATC.COM/CAR

ADVANCED TECHNOLOGIES CENTER

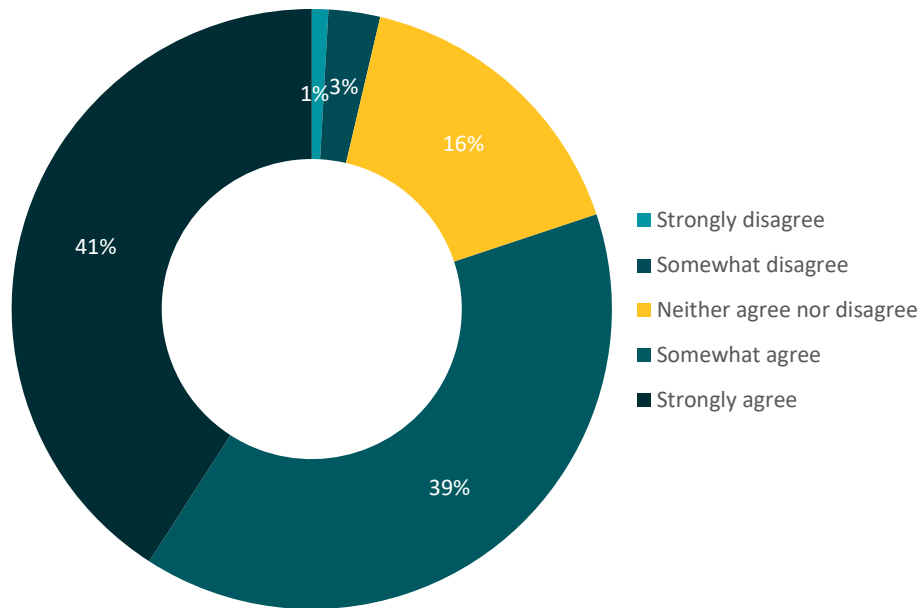
UBER

UBERATC.COM/CAR

Search Trend*: Artificial Intelligence vs. Machine Learning

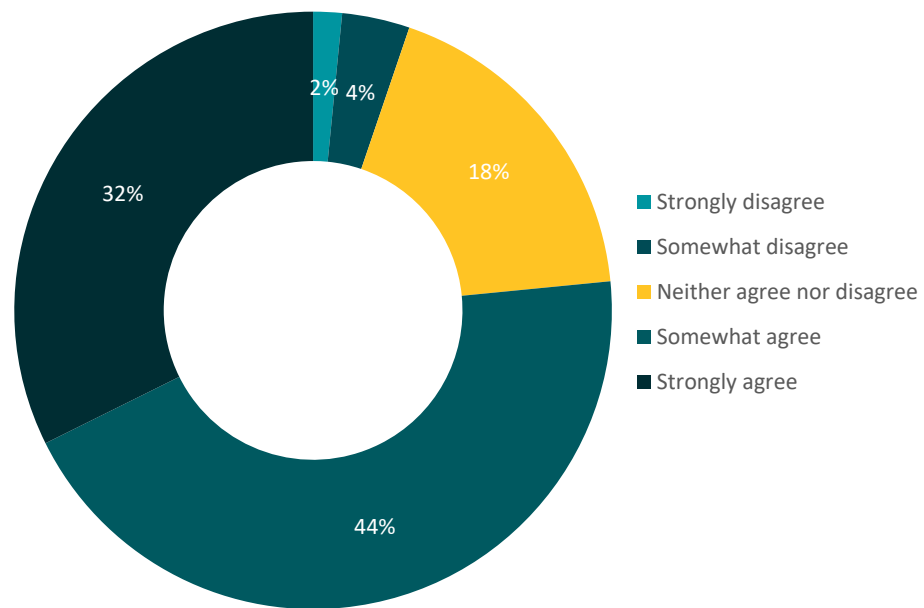


AI and ML will help/do help my organization detect and respond to threats faster



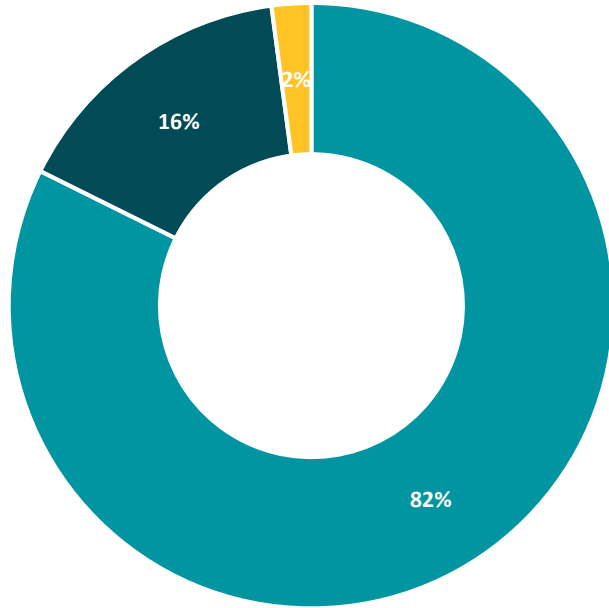
N = 900

AI and ML will help/do help solve my organization's cyber skills shortage



N = 900

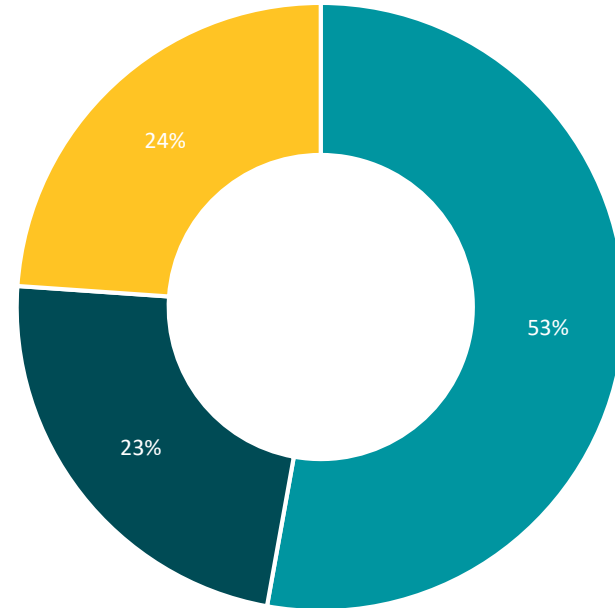
Have you/your organization implemented a cyber security product that uses ML?



- Yes
- No
- Don't know

N = 900

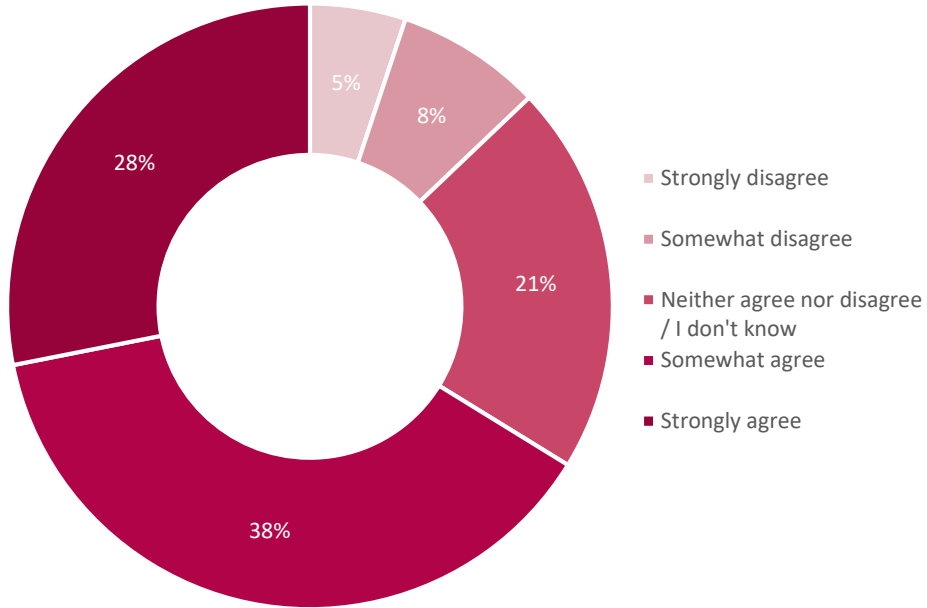
Does your organization have plans to use ML in its cyber security strategy in the next 3-5 years?



- Yes
- No
- Don't know

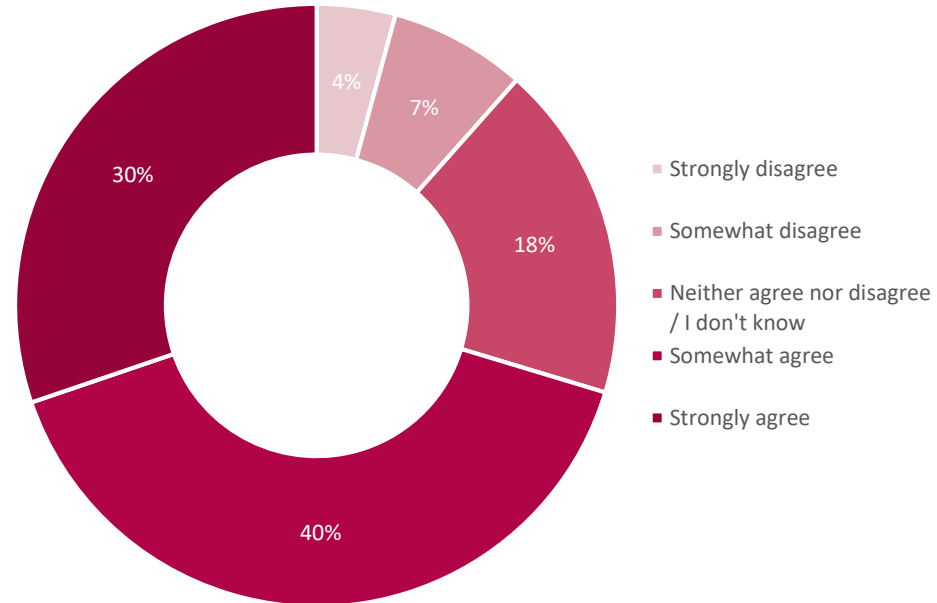
N = 159

AI will/would increase the number of attacks my organization will have to detect and respond



N = 900

AI-powered attacks will be more difficult to detect



N = 900

Adversarial Use of Machine Learning

Improved Spam

Spam in the “old days”

Domu

“Darling customer;”

“This is your functionary notify by Ceska Sporitelna to member service pay attention under...”

Predešlý oznámení mít been poslaný až k clen urcítý Žaloba Dotyk pridilil až k tato úcet.

Ackoliv clen urcítý Bezprostrední Dotyk , tebe musít obnovit se clen urcítý služba dát pozor pod ci ono vule být

Obnovit se Ted tvuj **SERVIS 24 Internetbanking.**

SERVIZ: **SERVIS 24 Internetbanking**

SKONANI: **Leden, 27 2008**

Být zavázán tebe do using SERVIS 24 Internetbanking. My ocenit tvuj obchod a clen urcítý příležitost až k slouž

Ceská Sporitelna Služba účastníkum

DULEŽITÝ Služba účastníkum HLÁŠENÍ

Low-quality Spam Today



Mrs. Ursula Alice Walton <tad@viquel.fr>

“GOOD NEWS”

To

“My dear recipient,”

“I am sure that this post would be a surprise as we never before and before...”

“After the death of my husband, I was in a life of my life and I had a good time...”

Po smrti môjho manžela som bol v živote svojho života a ja som mal dobrý čas. , ale pred smrťou môjho ma privilegovaných rodinných príslušníkov a blízkych priateľov, a je to tak nešťastné, že môj manžel dnes s svojho neskorého manžela a ja.

V súčasnosti mám takmer polovicu nášho majetku v niekoľkých charitatívnych domoch a v niektorých krajinách jednotlivcovi, predtým sme sa v živote nestretli, pretože sme boli ešte mladí v živote, dostaneme anonym

From: ACG GmbH & Co. KG <f.brunner@acg-technologies.de>

Sent:

Subject: INVOICE-RFQ-0094-8002-008-0018LT



Filecoder

Pane,

Moja kolegyňa, ktorá má túto objednávku vybavovať, je na dovolenke.

Chcem potvrdiť údaje v tejto faktúre od vás, pred jej odovzdaním na naše oddelenie účtovníctva.

Sú podrobnosti účtu na priloženej faktúre vaše správne bankové údaje?

Ak existuje nejaká chyba, ktorá potrebuje opravu v tejto faktúre?

Potvrďte kód IBAN a swift kód.

Ak by existovala akákoľvek existujúca dohoda, dajte mi vedieť.

S Pozdravom.

(Alexander Renga)

Clumsy language,
yet grammatically
correct



ACG GmbH & Co. KG

Stolen logo of a legitimate business

ACG GmbH & Co. KG
Automation Co & GmbH,
Erlenstraße 2,
60325 Frankfurt am Main,

Actual address of a legitimate business

Customization of malware according to the target's environment

Ramnit



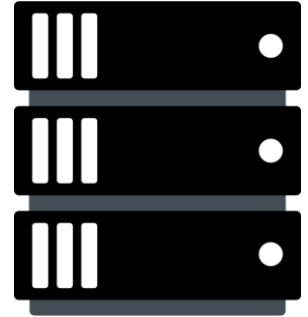
Victim 1



Victim 2



Victim 3



Attacker's
C&C

```
[ 1.614] [INFO] Number of available modules: 6
[ 1.615] [INFO] 1 Module name: AvTrust
[ 1.615] [INFO] Module description: Antivirus Trusted Module v2.0
[ 1.616] [INFO] 2 Module name: CookieGrabber
[ 1.616] [INFO] Module description: Cookie Grabber v0.2 (no mask)
[ 1.616] [INFO] 3 Module name: FFCH
[ 1.617] [INFO] Module description: FF&Chrome reinstall x64-x86 [
[ 1.617] [INFO] 4 Module name: FtpGrabber2
[ 1.618] [INFO] Module description: Ftp Grabber v2.0
[ 1.618] [INFO] 5 Module name: HOOKER
[ 1.618] [INFO] Module description: IE & Chrome & FF injector
[ 1.619] [INFO] 6 Module name: VNC IFSB
[ 1.619] [INFO] Module description: VNC IFSB x64-x86
```

Ramnit



Victim 1



Victim 2

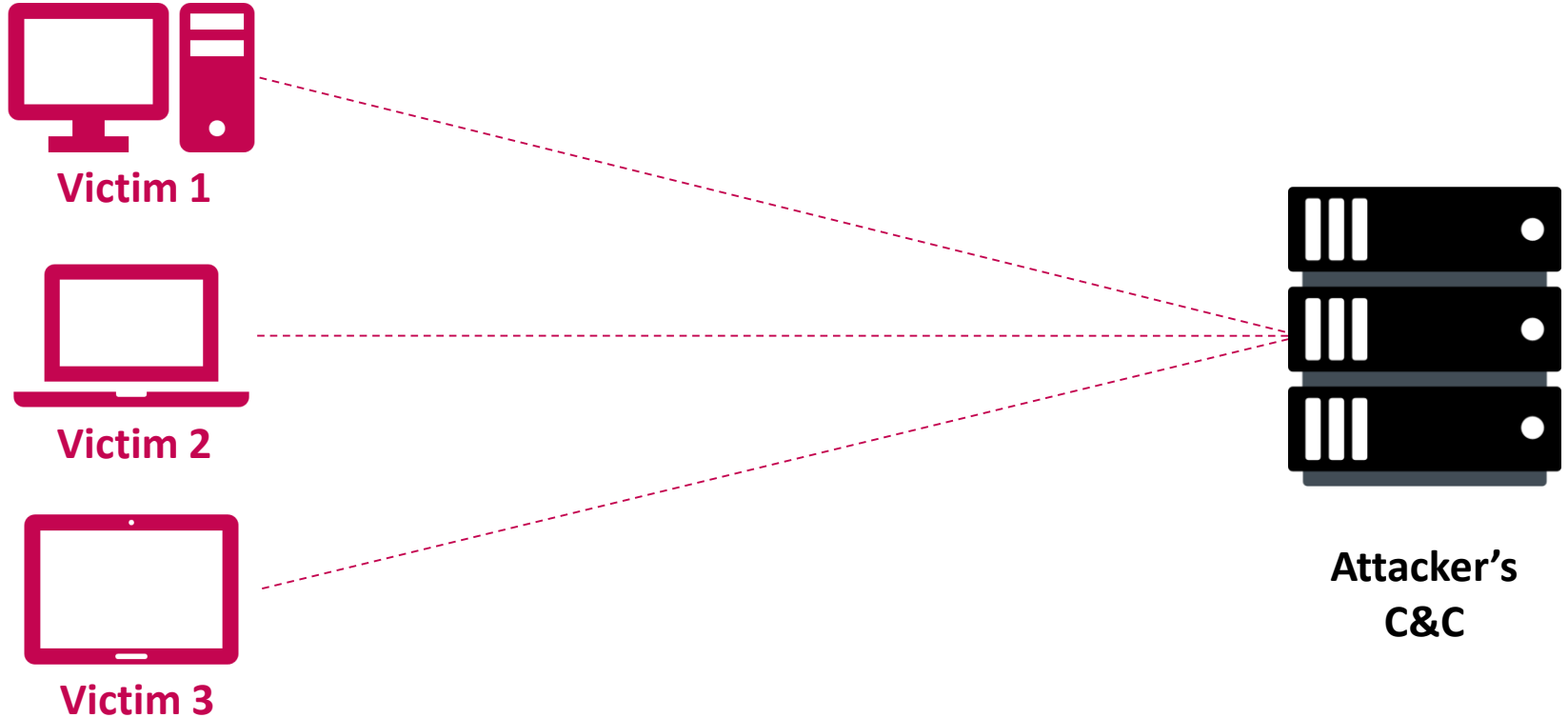


Victim 3

```
[INFO] Number of available modules: 6
[INFO] Module name: AvTrust
[INFO] Number of available modules: 4
[INFO] Module name: AvTrust
[INFO] Module description: Antivirus Trusted
[INFO] Module name: CookieGrabber
[INFO] Module description: Cookie Grabber v
[INFO] Module name: HOOKER
[INFO] Module description: IE & Chrome & FF
[INFO] Module name: VNC
[INFO] Module description: VNC (23 port) x64-
[INFO] Module description: IE & Chrome &
[INFO] Module name: VNC IFSB
[INFO] Module description: VNC IFSB x64-
```

**Attacker's
C&C**

Ramnit



Malware protecting itself

Emotet



Task Manager

File Options View

Processes Performance App history Startup Users Details Services

| Name | 5% CPU | 56% Memory | 0% Disk | 0% Network | 1% GPU | GPU Engine |
|---------------------------------------|--------|------------|---------|------------|--------|------------|
| Background processes (89) | | | | | | |
| > 64-bit Synaptics Pointing Enhanc... | 0% | 0.6 MB | 0 MB/s | 0 Mbps | 0% | |
| Active Protection System | 0% | 0.1 MB | 0 MB/s | 0 Mbps | 0% | |
| > Adobe Acrobat Update Service | 0% | 0.2 MB | 0 MB/s | 0 Mbps | 0% | |
| Application Frame Host | 0% | 7.9 MB | 0 MB/s | 0 Mbps | 0% | |
| > Auto Scroll Start Service | 0% | 0.3 MB | 0 MB/s | 0 Mbps | 0% | |
| BACK Monitor Application (32 ... | 0.1% | 6.6 MB | 0 MB/s | 0 Mbps | 0% | |
| > Camera Mute Control Service fo... | 0% | 0.7 MB | 0 MB/s | 0 Mbps | 0% | |
| COM Surrogate | 0% | 1.3 MB | 0 MB/s | 0 Mbps | 0% | |
| Communications Utility launche... | 0% | 1.4 MB | 0 MB/s | 0 Mbps | 0% | |
| > Cortana (2) | 0% | 112.6 MB | 0 MB/s | 0 Mbps | 0% | |
| CTF Loader | 0.4% | 1.7 MB | 0 MB/s | 0 Mbps | 0% | |
| Device Association Framework ... | 0% | 0 MB | 0 MB/s | 0 Mbps | 0% | |
| Device Association Framework ... | 0% | 1.1 MB | 0 MB/s | 0 Mbps | 0% | |
| > ESET Enterprise Inspector Agent | 0% | 22.0 MB | 0 MB/s | 0 Mbps | 0% | |
| ESET Main GUI | 0% | 4.8 MB | 0 MB/s | 0 Mbps | 0% | |
| > ESET Management Agent Modu... | 0% | 11.1 MB | 0 MB/s | 0 Mbps | 0% | |
| > ESET Service (2) | 0% | 34.1 MB | 0 MB/s | 0 Mbps | 0% | |
| f.lux (32 bit) | 0% | 2.5 MB | 0 MB/s | 0 Mbps | 0% | |
| > Groove Music | 0% | 0.1 MB | 0 MB/s | 0 Mbps | 0% | |

Fewer details End task

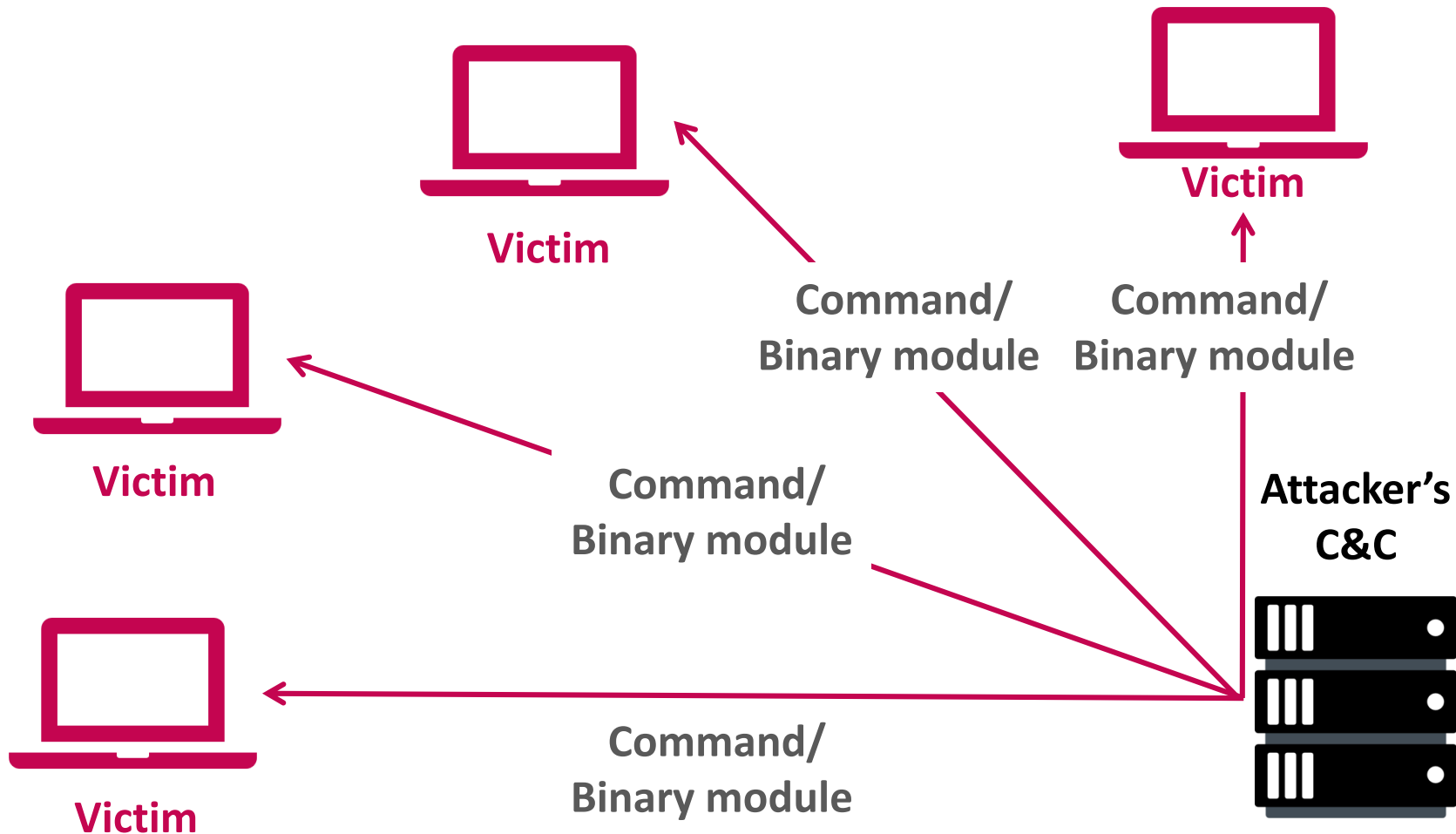


Victim

Return proList +

Attacker's
C&C





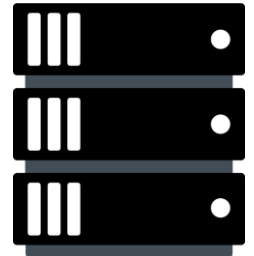
The screenshot shows the Windows Task Manager Performance tab. The 'Background processes (89)' section is expanded, showing a list of processes with their resource usage. The columns are: Name, CPU, Memory, Disk, Network, GPU, and GPU Engine. The processes listed are:

| Name | CPU | Memory | Disk | Network | GPU | GPU Engine |
|-------------------------------------|-----|--------|--------|---------|-----|------------|
| 64-bit Synaptics Pointing Enhanc... | 0% | 0.6 MB | 0 MB/s | 0 Mbps | 0% | |
| Active Protection System | 0% | 0.1 MB | 0 MB/s | 0 Mbps | 0% | |
| Adobe Acrobat Update Service | 0% | 0.2 MB | 0 MB/s | 0 Mbps | 0% | |
| Application Frame Host | 0% | 7.9 MB | 0 MB/s | 0 Mbps | 0% | |
| Auto Scroll Start Service | 0% | 0.3 MB | 0 MB/s | 0 Mbps | 0% | |

Researcher/
Honeypot



Attacker's
C&C



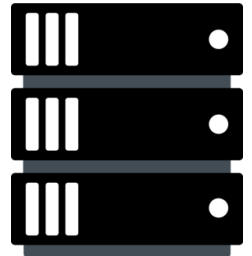
Retl Quit/Sleep

| Name | 5% CPU | 56% Memory | 0% Disk | 0% Network | 1% GPU | GPU Engine |
|--|-------------|---------------|---------------|---------------|-----------|------------|
| Background processes (89) | | | | | | |
| > 64-bit Synaptics Pointing Enhanc... | 0% | 0.6 MB | 0 MB/s | 0 Mbps | 0% | |
| Active Protection System | 0% | 0.1 MB | 0 MB/s | 0 Mbps | 0% | |
| > Adobe Acrobat Update Service | 0% | 0.2 MB | 0 MB/s | 0 Mbps | 0% | |
| Application Frame Host | 0% | 7.9 MB | 0 MB/s | 0 Mbps | 0% | |
| > Auto Scroll Start Service | 0% | 0.3 MB | 0 MB/s | 0 Mbps | 0% | |
| BACK Monitor Application (32 ...) | 0.1% | 6.6 MB | 0 MB/s | 0 Mbps | 0% | |
| > Camera Mute Control Service fo... | 0% | 0.7 MB | 0 MB/s | 0 Mbps | 0% | |
| COM Surrogate | 0% | 1.3 MB | 0 MB/s | 0 Mbps | 0% | |

Researcher/
Honeypot



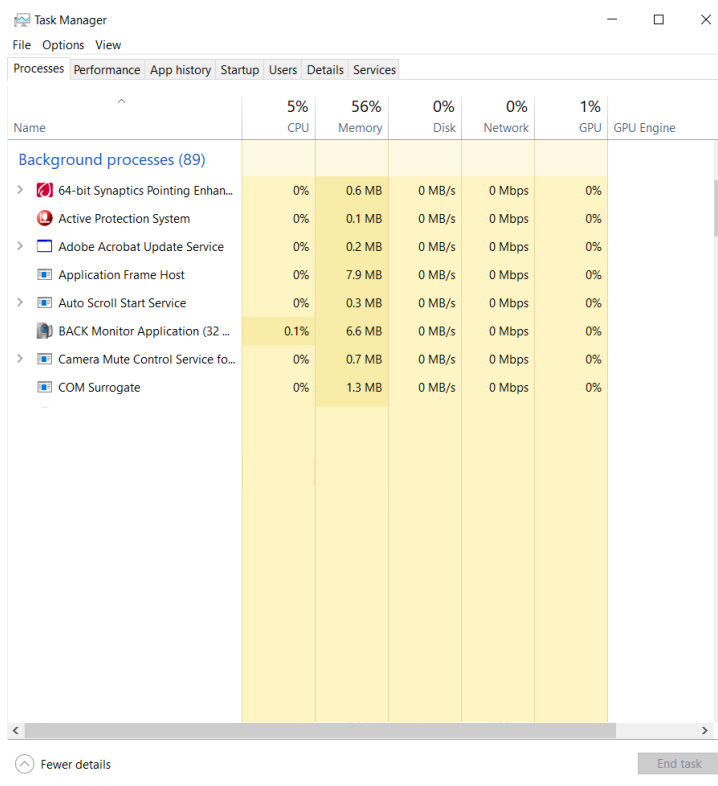
Attacker's
C&C



Re Command/
Binary module

A few days later...

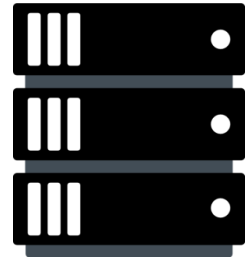
Researcher/
Honeypot



| Name | 5% CPU | 56% Memory | 0% Disk | 0% Network | 1% GPU | GPU Engine |
|---------------------------------------|--------|------------|---------|------------|--------|------------|
| Background processes (89) | | | | | | |
| > 64-bit Synaptics Pointing Enhanc... | 0% | 0.6 MB | 0 MB/s | 0 Mbps | 0% | |
| Active Protection System | 0% | 0.1 MB | 0 MB/s | 0 Mbps | 0% | |
| > Adobe Acrobat Update Service | 0% | 0.2 MB | 0 MB/s | 0 Mbps | 0% | |
| Application Frame Host | 0% | 7.9 MB | 0 MB/s | 0 Mbps | 0% | |
| > Auto Scroll Start Service | 0% | 0.3 MB | 0 MB/s | 0 Mbps | 0% | |
| BACK Monitor Application (32 ... | 0.1% | 6.6 MB | 0 MB/s | 0 Mbps | 0% | |
| > Camera Mute Control Service fo... | 0% | 0.7 MB | 0 MB/s | 0 Mbps | 0% | |
| COM Surrogate | 0% | 1.3 MB | 0 MB/s | 0 Mbps | 0% | |



Attacker's
C&C

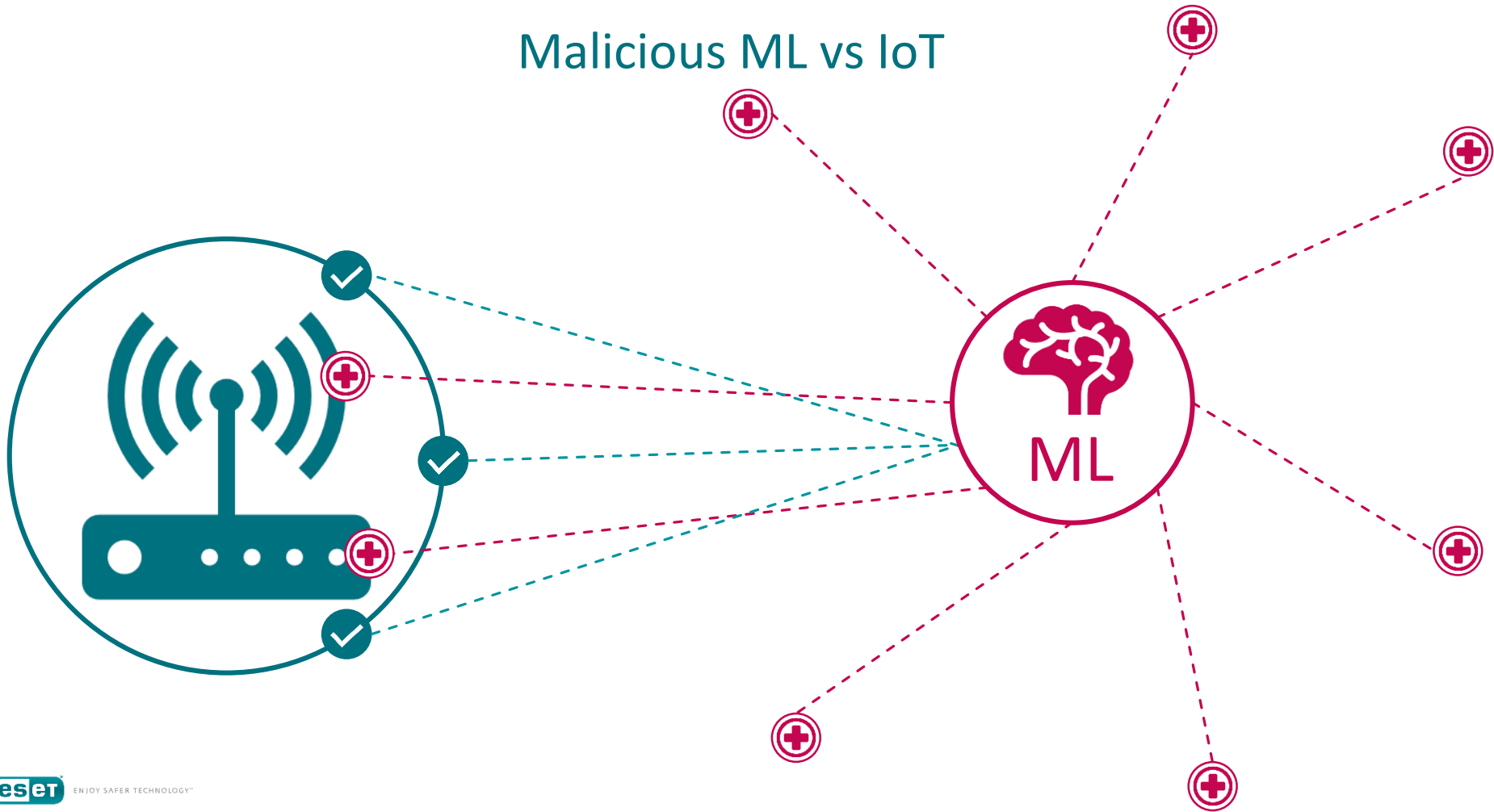


Retu Quit/Sleep



AI-powered malware vs. IoT

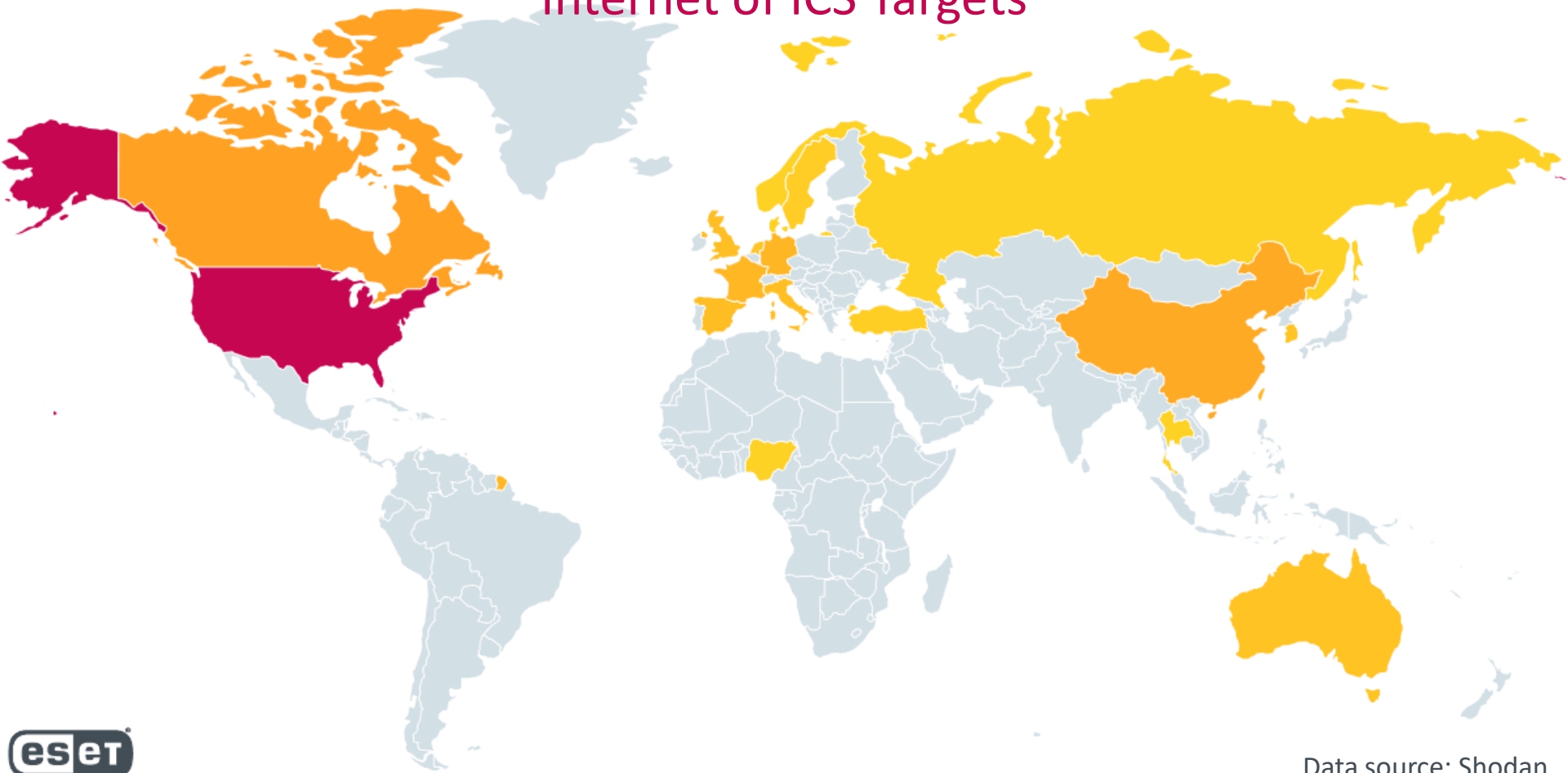
Malicious ML vs IoT



Malicious ML vs IoT



Internet of ICS Targets

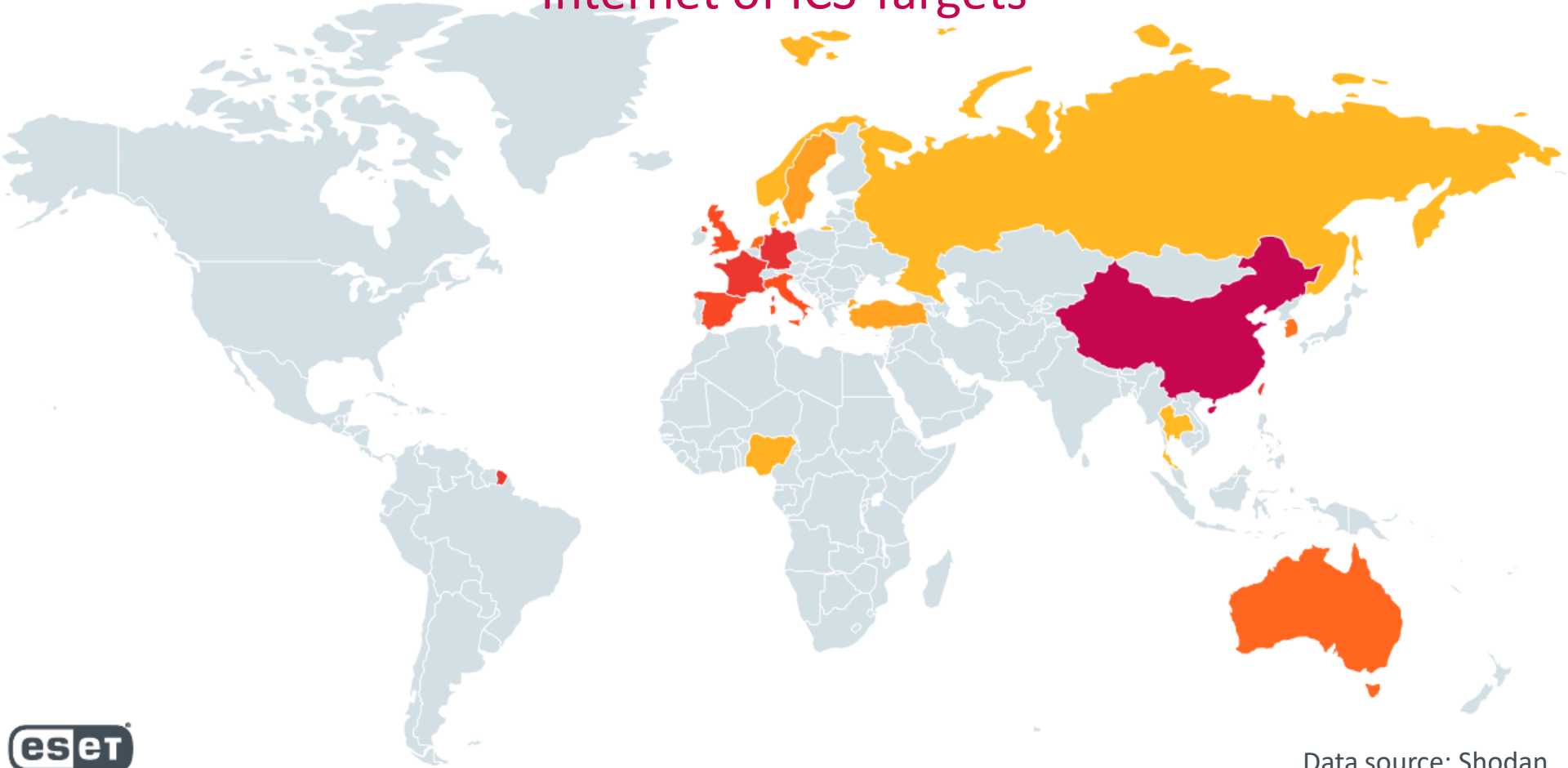


Data source: Shodan

0.26%

51.68%

Internet of ICS Targets



Data source: Shodan

ESET Augur

Machine Learning in ESET

Neural
Networks in
product

1998

DNA
Detections
(Online ML)

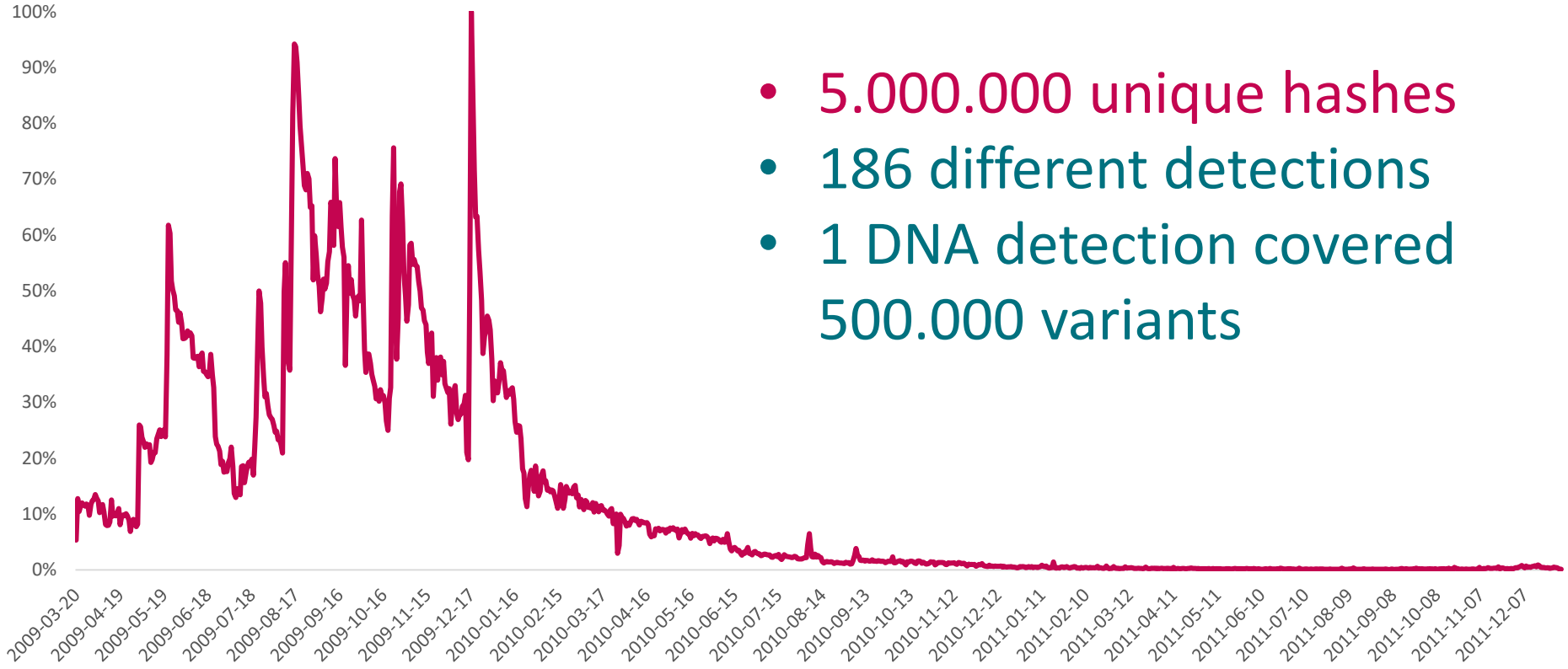
2005

DNA Detections



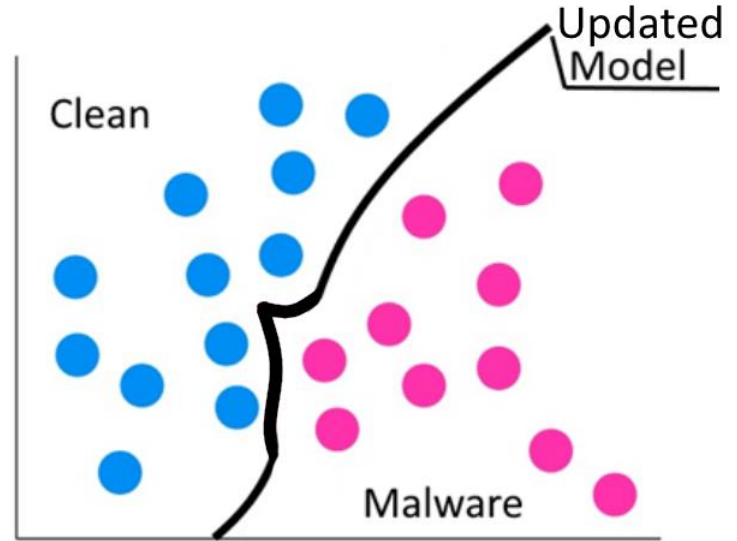
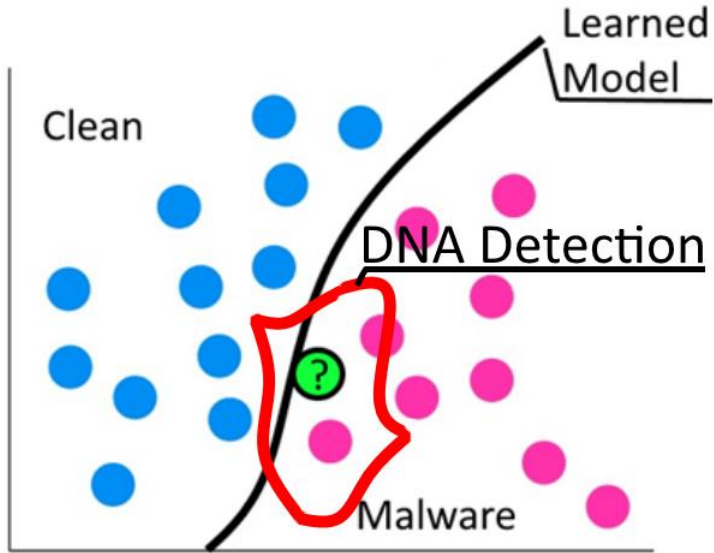
- Complex definitions of malicious behavior and malware characteristics
- Contain much more information than the indicators of compromise
- Single well-crafted DNA Signature can detect tens of thousands of related samples

Swizzor Detections



- 5.000.000 unique hashes
- 186 different detections
- 1 DNA detection covered 500.000 variants

ESET DNA Detections



Machine Learning in ESET

Neural
Networks in
product

1998

DNA
Detections
(Online Learning)

2005

Expert system
for mass
processing

2006

Algorithm placing
samples on the
map

2012



Big data and cheaper hardware

Popularity and availability of machine-learning algorithms

Machine Learning Algorithms

Deep Learning

- Deep Belief Networks (DBN)
- Convolutional Neural Network (CNN)
- Stacked Auto-Encoders

Ensemble

- Random Forest
- Gradient Boosting Machines (GBM)
- Boosting
- Bootstrapped Aggregation (Bagging)
- AdaBoost
- Stacked Generalization (Blending)

Gradient Boosted Regression Trees (GBRT)

Radial Basis Function Network (RBFN)

Support Vector Machine (SVM)

Hopfield Network

Rule System

- Cubist
- One Rule (OneR)
- Zero Rule (ZeroR)
- Repeated Incremental Pruning to Produce Error Reduction (RIPPER)

Regression

- Linear Regression
- Ordinary Least Squares Regression (OLSR)
- Stepwise Regression
- Multivariate Adaptive Regression Splines (MARS)
- Locally Estimated Scatterplot Smoothing (LOESS)

Decision Tree

- Gaussian Naive Bayes
- Multinomial Naive Bayes
- Bayesian Network (BN)
- Classification and Regression Tree (CART)
- Iterative Dichotomiser 3 (ID3)
- C4.5
- C5.0
- Chi-squared Automatic Interaction Detection (CHAID)
- Decision Stump
- Conditional Decision Trees
- M5

Instance Based

- Principal Component Analysis (PCA)
- Partial Least Squares Regression (PLSR)
- Sammon Mapping
- Multidimensional Scaling (MDS)
- Principal Component Regression (PCR)
- Partial Least Squares Discriminant Analysis
- Mixture Discriminant Analysis (MDA)
- Quadratic Discriminant Analysis (QDA)
- Regularized Discriminant Analysis (RDA)
- Flexible Discriminant Analysis (FDA)
- Linear Discriminant Analysis (LDA)
- k-Nearest Neighbour (kNN)
- Learning Vector Quantization (LVQ)
- Self-Organizing Map (SOM)
- Locally Weighted Learning (LWL)
- k-Means



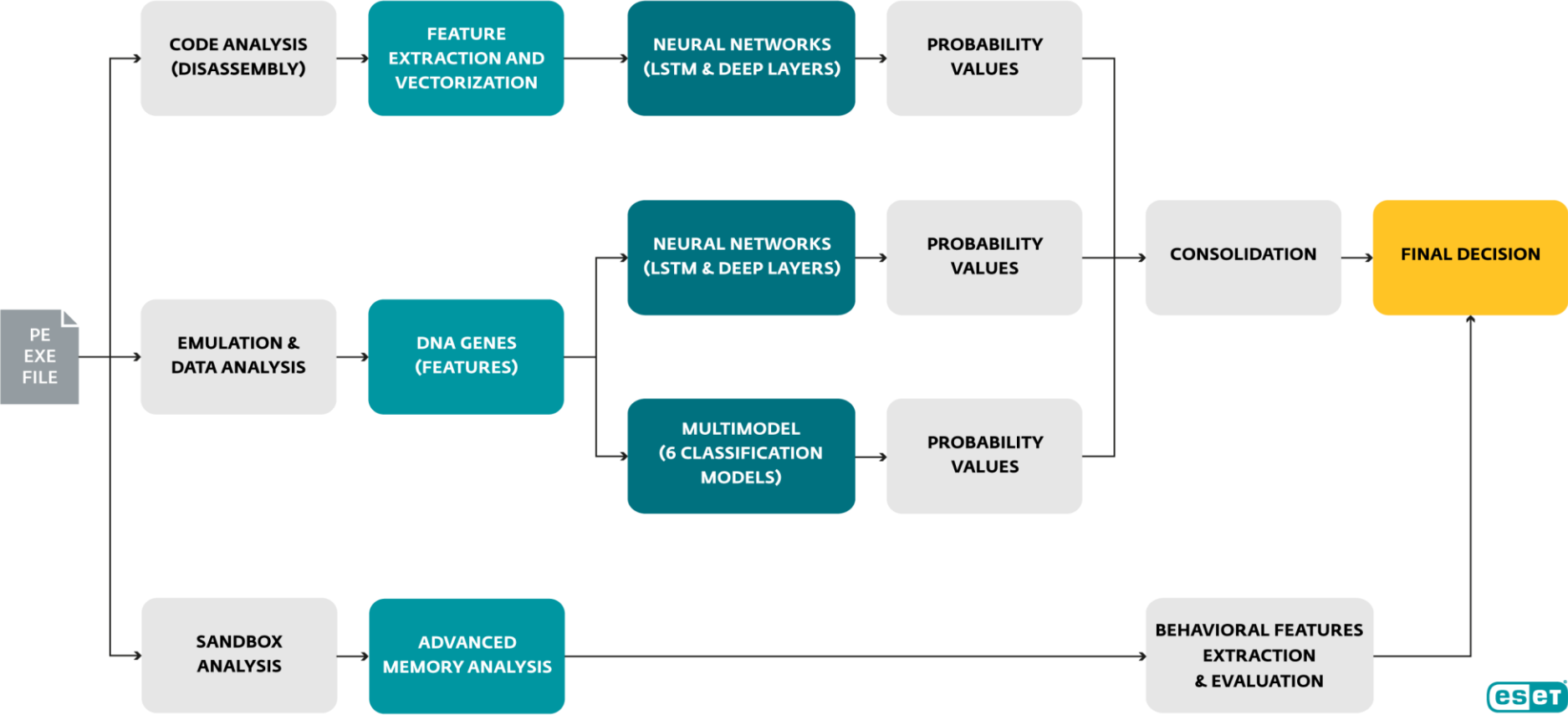
ESET's database of correctly labelled samples





Machine Learning Engine ESET Augur

How Augur processes data

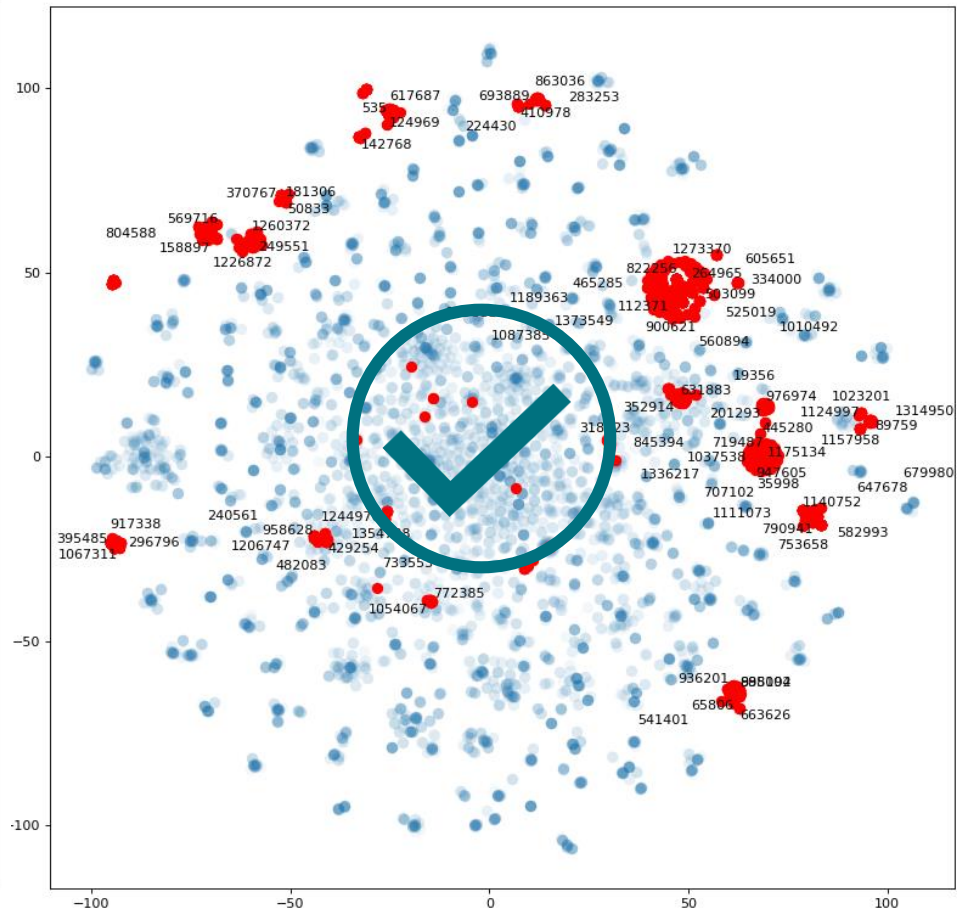
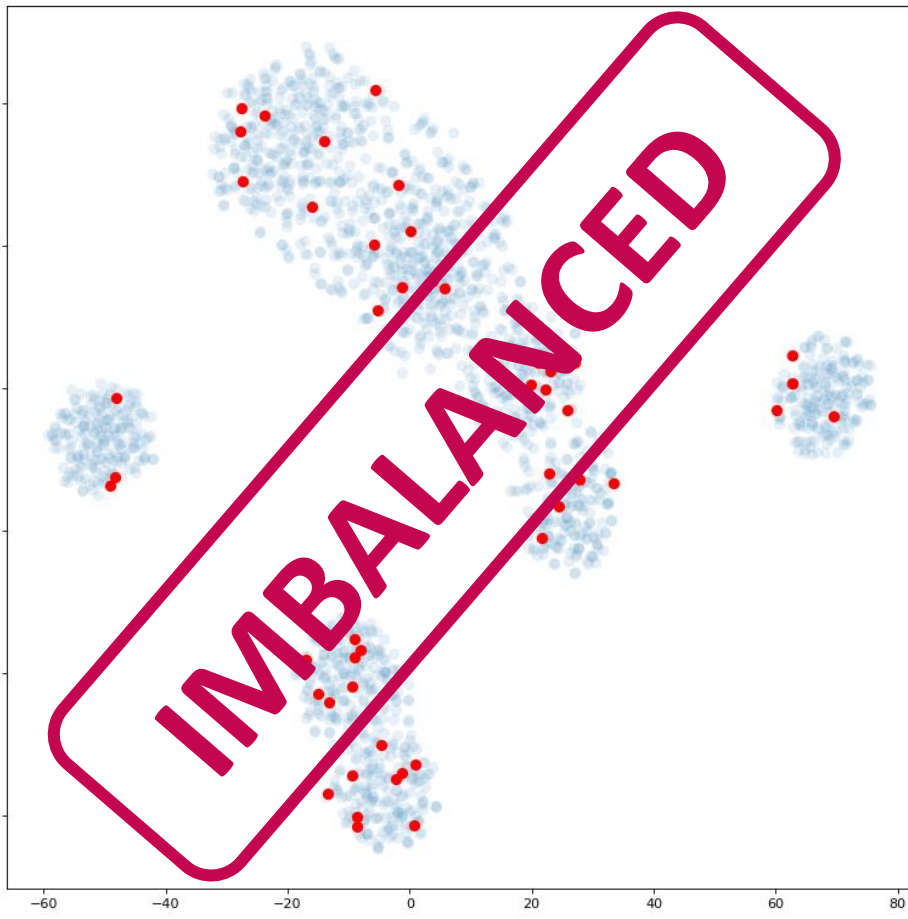


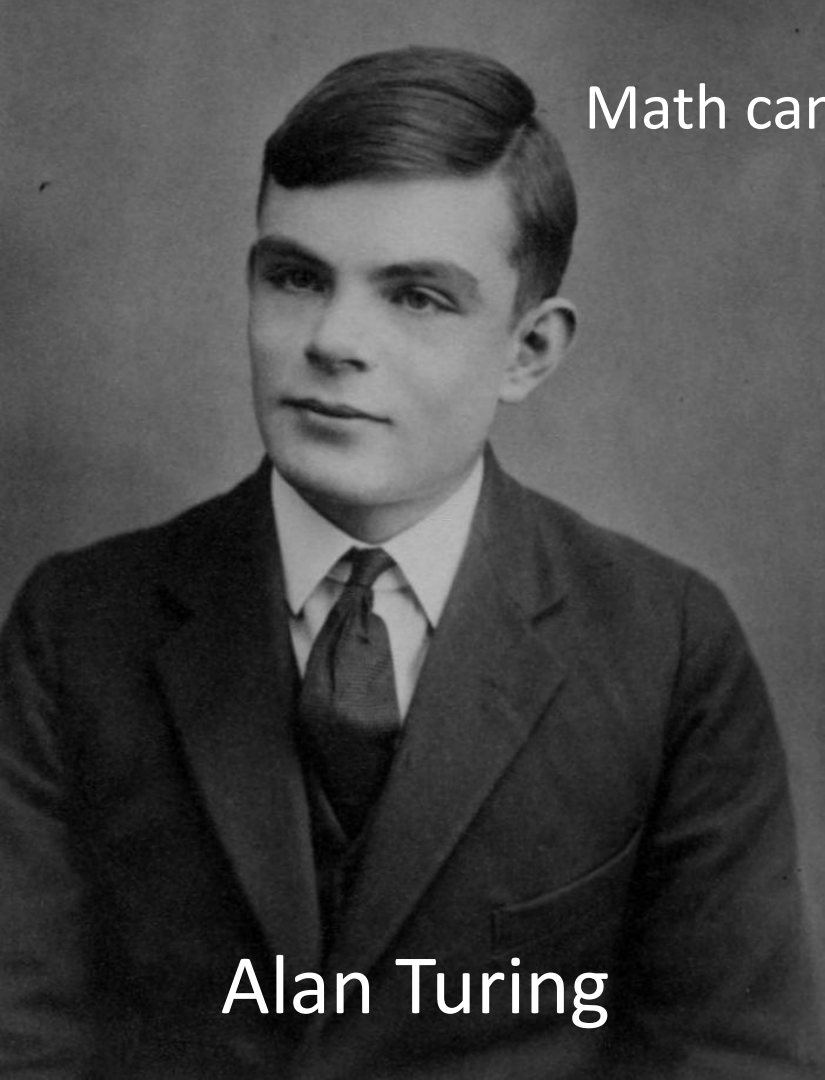
Test of Older (March 2017) Augur Model

| Malware (May 2017 June 2018) | Detection ratio | Number of samples | Number of samples detected by Augur |
|---------------------------------|-----------------|----------------------|---|
| NotPetya | 98.8 | 86 | 85 |
| BadRabbit | 100 | 20 | 20 |
| Crysis | 100 | 30 | 30 |
| WannaCryptor | 100 | 67 | 67 |

Limits of Machine Learning

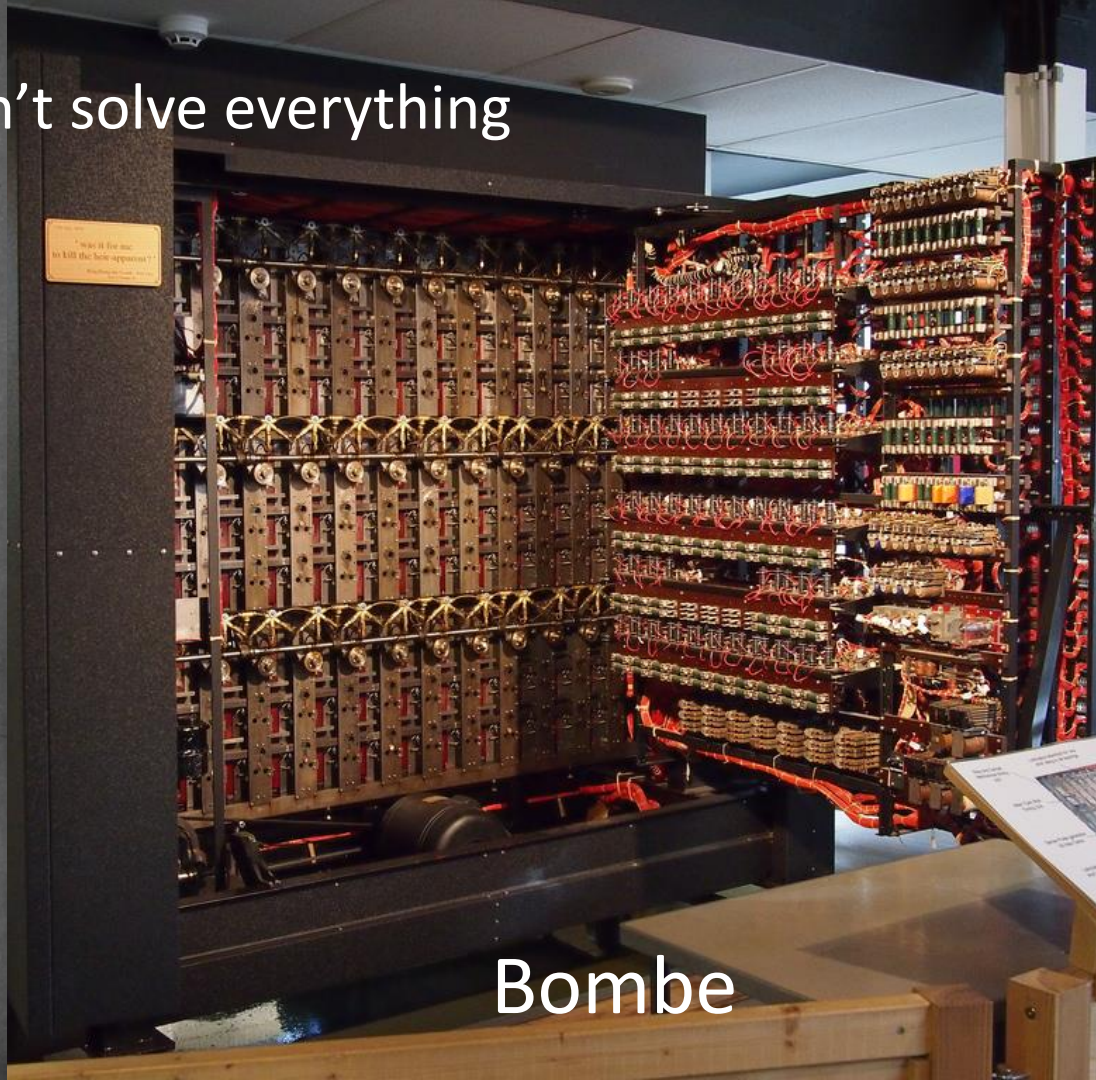
Training Set





Alan Turing

Math can't solve everything



Bombe

Intelligent Adversary



Intelligent Adversary

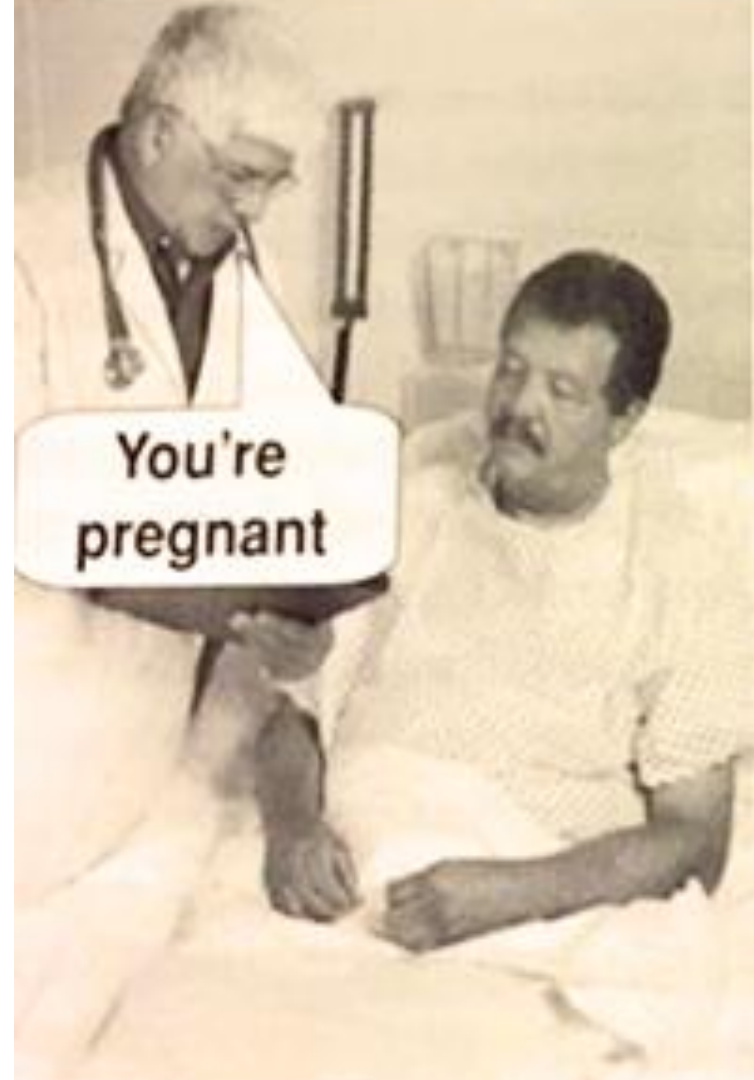


Intelligent Adversary



False Positives

- Malware **may break** business
(ransomware)
- False positive **will break** business
(OS, Files, IP/URLs)



Silver Bullet?







**Botnet
Protection**


**Ransomware
Shield**


**In-product
Sandbox**


**Cloud Malware
Protection System**


**Advanced
Memory
Scanner**


UEFI Scanner


Dna Detection


Exploit Blocker


**Network Attack
Protection**


**Reputation
& Cache**

-  PRE EXECUTION
-  EXECUTION
-  POST EXECUTION

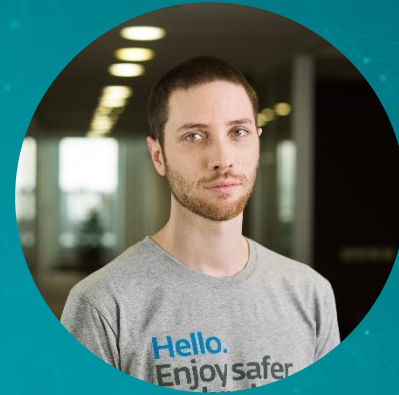
- What is AI and ML
- How is AI/ML implemented in ESET products
- How could attackers use AI/ML against their victims in the future
- Available on [WeLiveSecurity.com](https://www.welivesecurity.com) or via our PR gurus





Juraj Janosik

Head of AI/ML Team



Ondrej Kubovic

Security Awareness Specialist