

# Threat Hunting

Using Machine Learning & Threat Intelligence



Cristian Pațachia, Orange Romania  
Development & Innovation Manager

**How do you go about finding a  
needle in a haystack?**

**It's pretty simple:  
You bring a magnet**

# Agenda

- 1 New Threats: Advanced and persistent
- 2 Machine Learning: Platform, Challenges, Sources, Data, Algorithms
- 3 Threat Intelligence: Crowd-Sourced Knowledge
- 4 Detection of cyber-threats: Beyond Firewalls and IPSs
- 5 Critical Infrastructure Operators: TELCOs
- 6 RESISTO as a holistic approach to Resilience
- 7 Orange Romania: Research, Education, Innovation
- 8 Key Takeaways, Q&A

# New threats

## Advanced and Persistent

### APTs

**Advanced:** uses sophisticated vulnerability exploits, 0-day exploits. Stealthy.

**Persistent:** external command and control system that continuously monitors target(s) and exfiltrates data from target(s)

Highly dependent on social engineering, surveillance, supply-chain compromise

Usually infects multiple end-points in large networks.

Requires organization-wide efforts to detect, stop & clean-up.

# New threats

## Advanced and Persistent

### STAGES

**Initial Compromise:** use of social engineering, spear phishing, media infestation with zero-day malware

**Foothold in the stronghold:** create C2C controllable 'victim' end-point in organization's network

**Reconnaissance:** Look around, find other vulnerable 'victims', gather info about network architecture;

**Lateral movement:** expand control to other endpoints, servers, active equipment, perform data harvesting on them;

**Keep a low profile:** maintain presence, use stealthy / encrypted channels to communicate with C2C;

**Complete mission:** exfiltrate data, delete all tracks

# Machine Learning

*“Getting computers to act without being explicitly programmed”*  
Stanford University

*“(...) algorithms can figure out how to perform important tasks by generalizing from examples”*  
University of Washington



# Machine Learning

## Big Data and it's Challenges

- We generate a huge amount of data with various logical structures within multiple formats
- We generate “raw” unstructured data
- We have dozen of devices and technologies that analyze and correlate specific log and flow data. Most of them are proprietary. They speak their own language. And Syslog.
- We discuss threats using natural language
- We disseminate using language constructs that computers usually don't understand (e.g. – posting about threats on social media platforms like Twitter, using sarcasm etc.)

# Machine Learning

What do we add to the mix?



**Cellular Data**  
**Wi-Fi Data**  
**Security Data**  
**Device metrics**



**Threat Intelligence**  
**Feeds**  
**Malware Analysis**  
**Malware Samples**



**Vulnerability Scans**  
**CVEs & Zero Day**  
**Feeds**  
**Pentest Results**



# Machine Learning

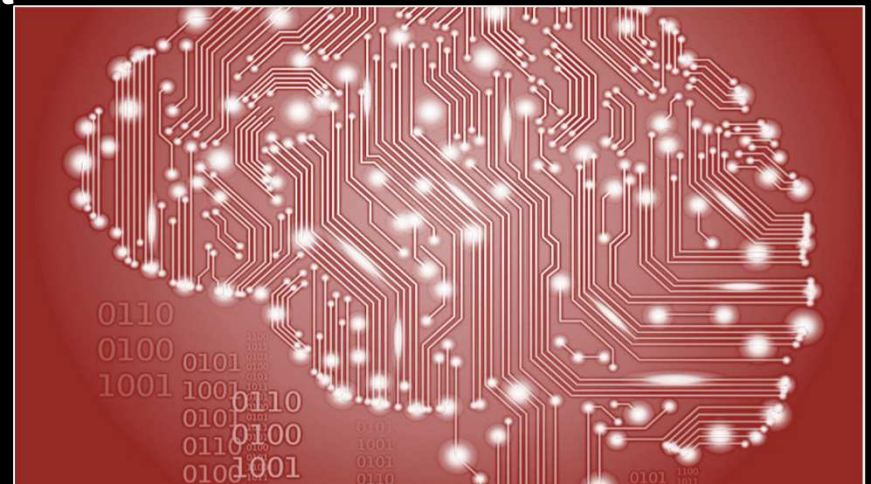
## The challenges

1. Deep Learning / Machine Learning are **JUST TOOLS**. They are NOT Silver Bullets.
2. For supervised M.L. you need reliable **labeled** sample data. And reliable is hard to find.
3. You **NEED** expert input and lots of context. The machines, by themselves won't turn dust to gold.



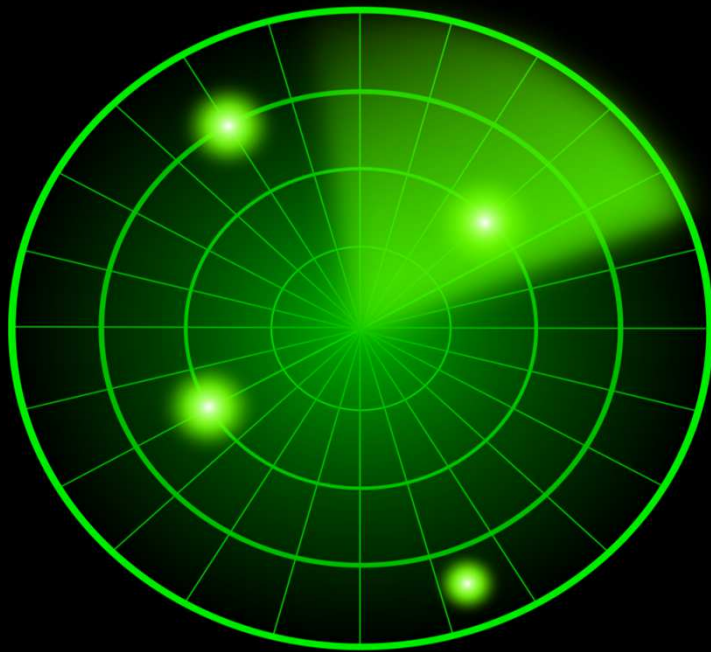
# Threat Intelligence & Crowd-Sourced Knowledge

1. We're enriching data and adding context
2. We're integrating feeds from sources like **OTX** and **VirusTotal**
3. We're working on 'reading' tweets relevant to new threats. This is tricky as machines can't understand concepts like humor or irony. It's WIP.
4. We're developing a Threat Sharing platform with local key-players in the cybersecurity space, both businesses and public-sector.



# Threat Detection

## Beyond NGFWs and IPSs



1. Signature-based detection works for most known threats
2. It won't work for new threats and APTs
3. If you don't have a signature for your new type of threat you must rely on behavior analysis.
4. We use sandboxes with near-instant spin-up and complete deletion after use. We monitor and ingest their output into our M.L. engine

# Critical Infrastructure Operators

## What is Critical Infrastructure?

Critical Infrastructure is a term used by governing bodies to describe assets that are essential for the functioning of a society and economy

## TELCOs as C.I.s

Telecommunication Operators can be assets regarded as C.I. operators because they provide services necessary for coordination and basic inter-human communications

## Specific VULNs & Threats

Large, complex technical infrastructure;



Cyber-Physical Threats

Provides services for current-gen and at least N-2 Gen devices (such as 2.5G / EDGE);



Specific vulnerabilities and threats (SS7)

Heavily reliant on human decisions



APTs relying on Social Engineering (SIM Hijacking)

# Resilience as a function of Threat Detection

## RESILIENCE

The ability (capacity) to provide and maintain an acceptable level of service in the face of threats, faults and challenges to normal operations.

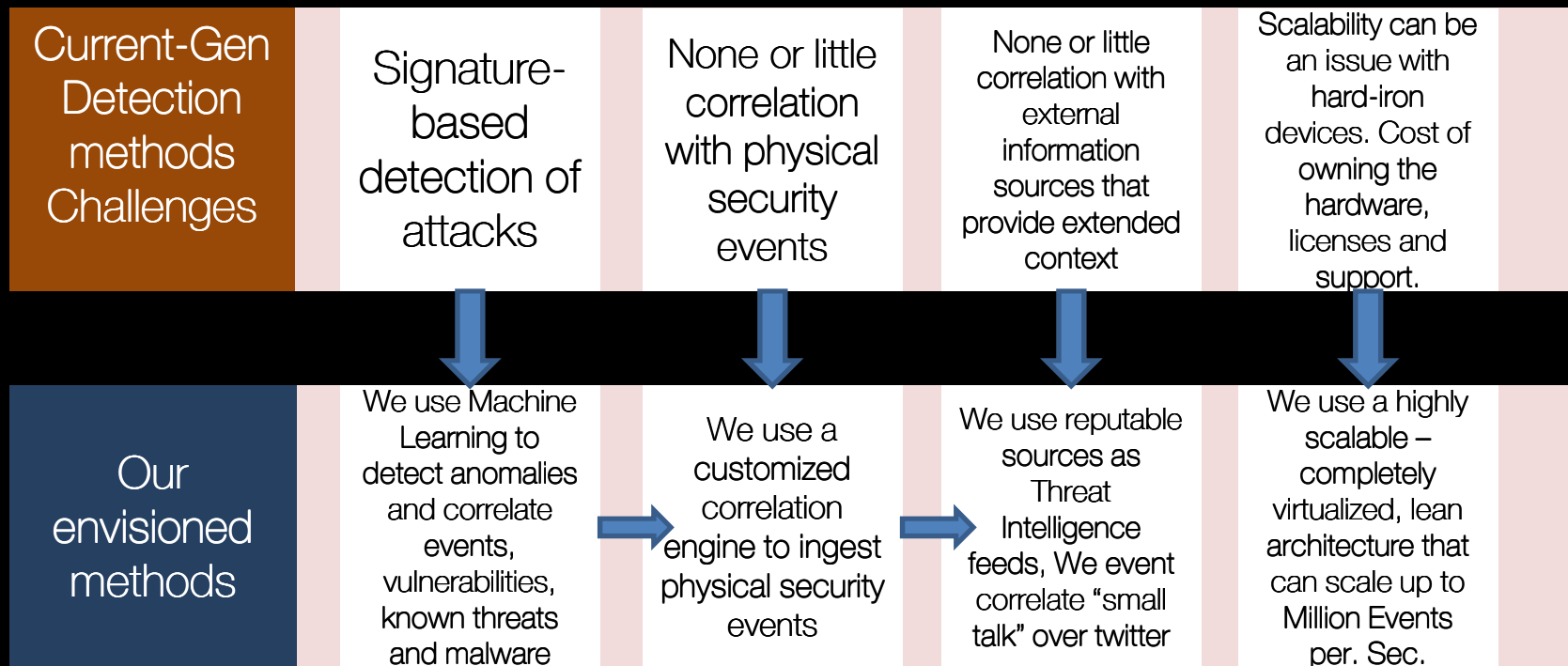
In order to increase resilience of a given system, the probable threats, challenges and risks have to be identified.

## THREAT DETECTION

- Key Factor for resilience
- Must adapt to specific threats
- Reliable in detecting complex and persistent threats

A word cloud of cybersecurity and resilience-related terms. The words are arranged in a circular pattern, with some words being larger and more prominent than others. The words include: Restriction, Adaptive, Response, Monitoring, Analytic, Positioning, Integrity, Substantiated, Dynamic, Privilege, Diversity, Segmentation, Unpredictability, and Non-Persistence. The words are in various shades of purple, blue, and yellow.

# Resilience as a function of Threat Detection





## A holistic approach to resilience enhancement

**RES**ilience enhancement and risk control platform for communication infra**ST**ructure **O**perators

Horizon 2020  
Project



Consortium  
of 19  
partners



Large  
Enterprises  
C.I. Operators  
Universities  
Research &  
Technology  
Organizations

- Holistic Approach to Situation Awareness
- Innovative Risk & Resilience & Improvement Process Management
- Decision Support System
- Protection against cyber-physical threats
- Modeled on state-of-the art technologies (Machine Learning, IoT, Block chain, Airborne Threat Detection, Holistic A-V analytics)



## Key Innovation Areas

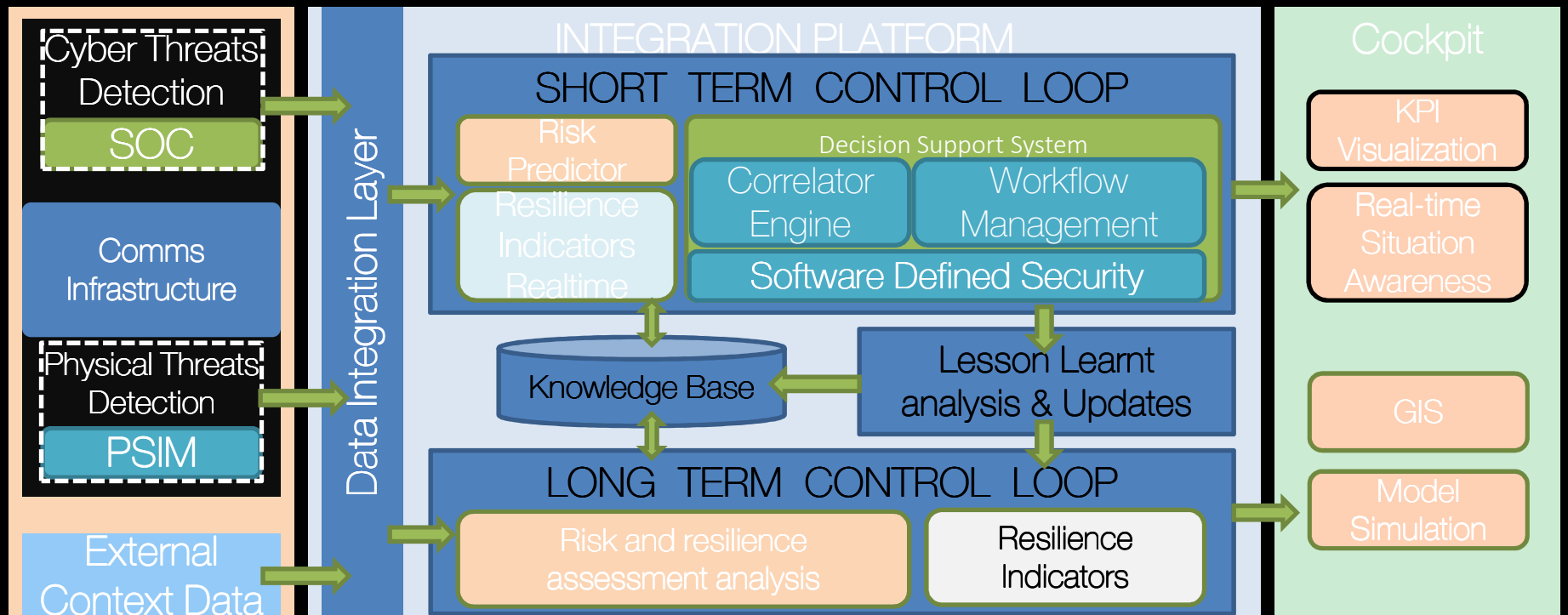
1. Enlarged Threat Landscape considered (**Cyber/ Physical/ Cyber + Physical**)
2. Holistic Approach to System Modelling
3. Integrated Risk and Resilience management
4. Convergence of **PSIM** and **Cyber Protection Technology**
5. Perspective – New challenges posed by 5G evolution (**IoT/loE, LPWAN**)
6. New Technology for detection, protection and response (**blockchain, drones, machine learning algorithms, software defined security**)
7. Cyber Intelligence

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.786409



# RESISTO

## Architecture overview



# Orange

## Education, Research, Innovation

### Orange Fab

Startups in cyber security, future of life and smart territories. From secure remote access to tissue-printing 3D printers

### Orange Educational Program

We're mentoring and training 40 scholars, students of ETTI, providing scholarships, internships, bachelors and masters degree project coordination

### Horizon 2020 Projects

We are participating in Consortiums researching new technologies, tools and methods in the IoT, 5G, Cyber-Security & Resilience, Computer Forensics

# Orange Fab

Development, Research, Innovation

**Orange Fab**

Startups in cyber security

## **Dekeneas**

APT Hunter, Watering Hole & Cryptojacking Detection

## **Appsulate**

Advanced Secure Remote Access and Website  
Isolation platform

## **Pentest-Tools**

On-line security audit framework

# ThreatMap

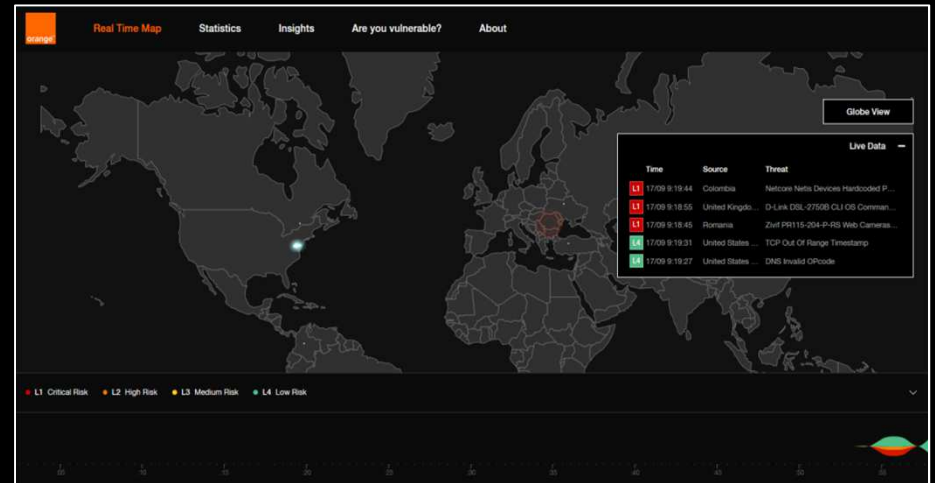
## Web Sites & IoCs

### Business Internet Security ThreatMap

We're using our own platform to monitor websites for IoCs

We're using 4 scanning engines:

- OWASP web vulnerability scanner
- CMS-specific vulnerability scanner
- APT – Watering Hole Malware scanner
- Previously Hacked? We'll interrogate the largest .ro database of hacked websites



We can scan websites linked in suspicious e-mails and actively prevent advanced phishing schemes


<https://bis-threatmap.orange.ro>

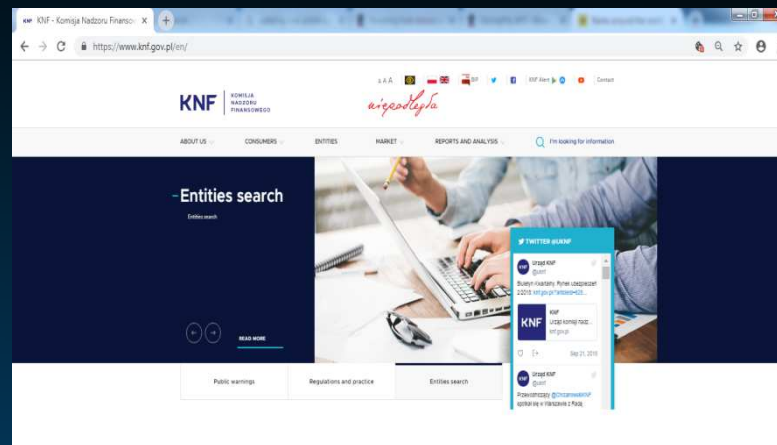
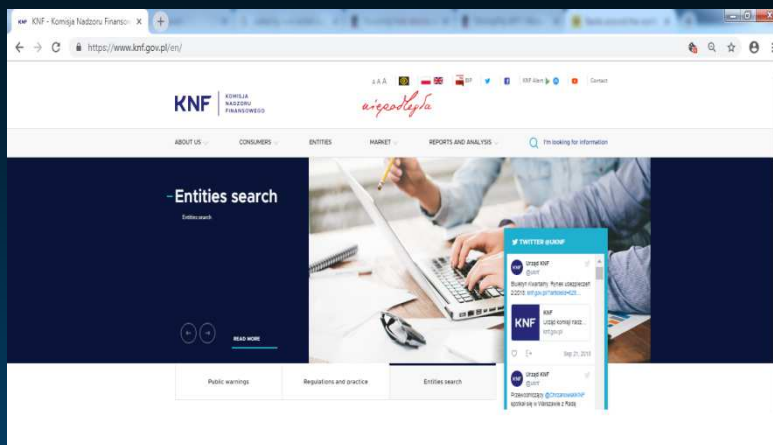


# IDEKNEAS

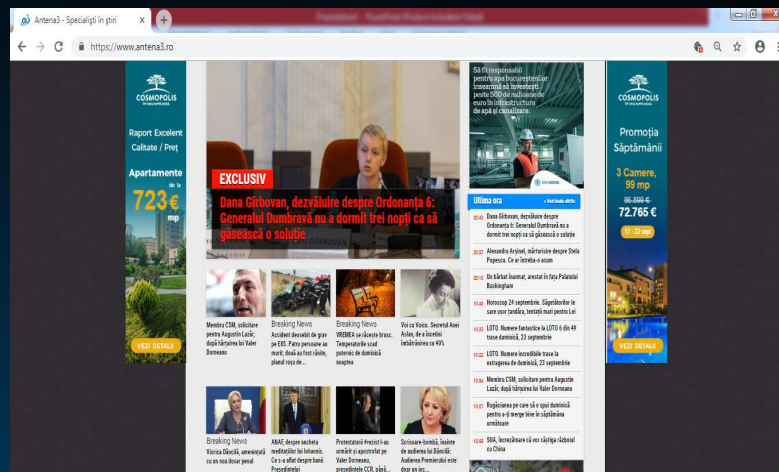
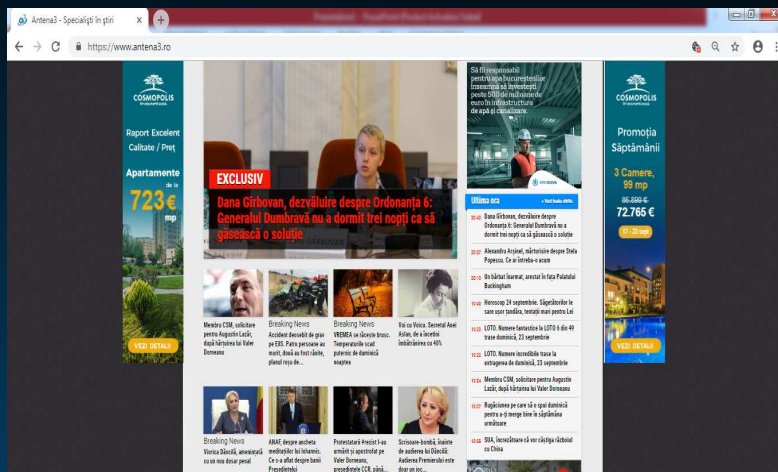
securing tomorrow

OR WHAT HAPPENS WHEN  
LEGITIMATE WEBSITES YOU  
VISIT ATTACK YOU

An abstract graphic consisting of several parallel white diagonal lines of varying lengths, slanted from the bottom-left towards the top-right. These lines are set against a dark blue background that transitions into a lighter blue gradient at the bottom.

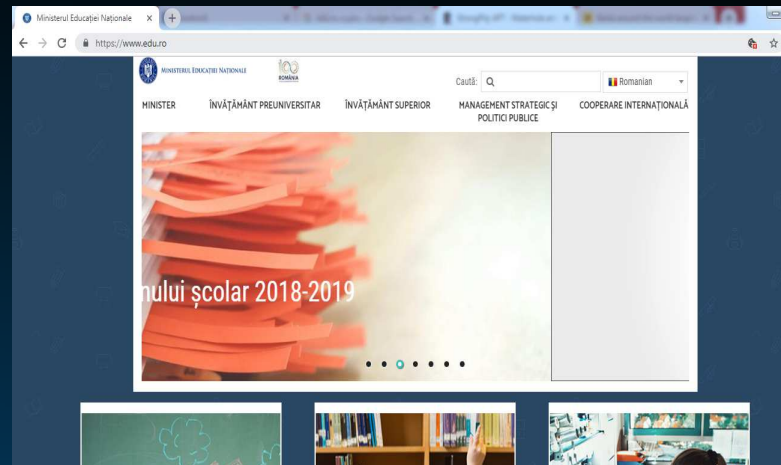
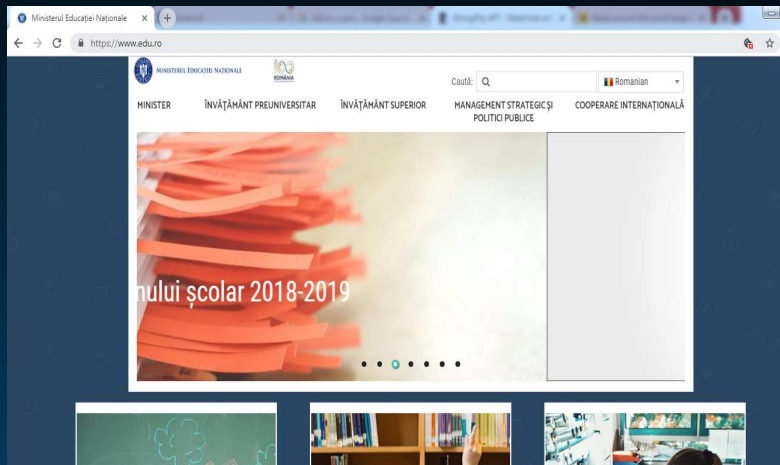


DO YOU SEE THE DIFFERENCE?



DO YOU SEE THE DIFFERENCE?





DO YOU SEE THE DIFFERENCE?

February 13, 2017 By Pierluigi Paganini

Last week, several Polish banks **confirmed** their systems were infected with a malware after their staff visited the site of the Polish Financial Supervision Authority.

The interesting aspect of the attack is that crooks used the Polish financial regulator, the Polish Financial Supervision Authority (KNF), to spread the malware.

In order to avoid spreading the malware, the authorities took the decision to shut down the entire network at the KNF "in order to secure evidence."



Symantec Official Blog

6 views

## Attackers target dozens of global banks with new malware

Watering hole attacks attempt to infect more than 100 organizations in 31 different countries.

Symantec Security Response

By: Symantec Security Response

Created 12 Feb 2017

0 Comments

繁体中文 日本語

0

0

0

0

0

Like

Organizations in 31 countries have been targeted in a new wave of attacks which has been underway since at least October 2016. The attackers used compromised websites or "watering holes" to infect pre-selected targets with previously unknown malware. There has been no evidence found yet that funds have been stolen from any infected banks.

The attacks came to light when a bank in Poland discovered previously unknown malware running on a number of its computers. The bank then shared indicators of compromise (IOCs) with other institutions and a number of other institutions confirmed that they too had been compromised.

As reported, the source of the attack appears to have been the website of the Polish financial regulator. The attackers compromised the website to redirect visitors to an exploit kit which attempted to install malware on selected targets.

Symantec has blocked attempts to infect customers in Poland, Mexico and Uruguay by the same exploit kit that infected the Polish banks. Since October, 14 attacks against computers in Mexico were blocked, 11 against computers in Uruguay, and two against computers in Poland.

# USUALLY VICTIMS LEARN OF THE ATTACK FROM THE NEWS



WHEN THE HARM HAS ALREADY  
BEEN DONE

**UPDATE** Antivirusii avertizează că site-ul Ministerului Educației **vă folosește calculatorul** pentru a mina criptomonede / De ce în codul sursă al paginii edu.ro apar trimeri către site-uri cu escorte din Turcia?

de Adrian Vasilescu HotNews.ro  
Miercuri, 30 mai 2018, 23:54 Economie | Telecom

Facebook Twitter Email Mai multe... 991

Antivirusul Avira și AdGuard, o aplicație care blochează reclamele, avertizează că site-ul Ministerului Educației [www.edu.ro](http://www.edu.ro) folosește în prezent resursele computerelor de pe care este accesat pentru a mina criptomonede, potrivit testelor HotNews.ro pe mai multe PC-uri și browsere web. Dezactivarea aplicației AdGuard arată cum la accesarea site-ului [www.edu.ro](http://www.edu.ro) gradul de folosire a procesorului PC-ului urcă instant la 100%, lucru ce ar fi imposibil să se întâmple la accesarea unei simple pagini web.



Edu.ro minează criptomonede  
Foto: Captura edu.ro

- **UPDATE (joi, 31 mai)** Specialiștii Centrului Național de Răspuns la Incidente de Securitate Cibernetică (CERT-RO) au declarat pentru HotNews.ro că în prezent sunt în discuții cu oficialii Ministerului Educației pentru identificarea cauzelor incidentului. Site-ul [www.edu.ro](http://www.edu.ro) este găzduit pe serverele Institutului Național de Cercetare în Informatică (ICI).

**CryptoPotato** BEGINNERS GUIDES ICO LIST BTC ANALYSIS **NEWS** BUY LANGUAGE Q

BTC \$6755.543 ETH \$245.229 XRP \$0.588 BCH \$501.066 LTC \$63.105 NEO \$19.729  
HOME - CRYPTO NEWS - ROMANIAN MINISTRY OF EDUCATION WEBSITE FORCES USERS TO MINE CRYPTOCURRENCY

 **1xBit.com** EXPLORE THE FUTURE OF BETTING! PLACE A BET BET WITH CRYPTOCURRENCIES

## Romanian Ministry of Education Website Forces Users to Mine Cryptocurrency

AUTHOR: MATTHEW NORTH - LAST UPDATED ON MAY 31, 2018 @ 12:22 UTC

The Romanian Ministry of Education's website, [edu.ro](http://edu.ro), is reportedly using its visitor's computers to mine cryptocurrency, as per a local news source [Hotnews.ro](http://Hotnews.ro).

It's claimed that the antivirus software companies *Avira* and *Adguard* warned users that the site has been flagged for mining cryptocurrencies. *Hotnews.ro* also reported that agents from the Romanian National Computer Security Incident Response Team are looking into the matter.

Earlier this year, the same agency reported that over 150 domains with the *.ro* extension were being used to host the *ConHive* script to mine the *Monero* cryptocurrency. *CoinHive* has since been classified as malware by leading antivirus companies such as *Malwarebytes* as it covertly mines cryptocurrencies without the user's permission. The act of mining can also reportedly overload the victim's CPU, which prevents the computer from functioning normally.

JOIN OUR COMMUNITY



 **Bitcasino.io** Get up to m\$ 1,000 FREE! Play now

 **etoro** INSTANT ORDERS INTUITIVE PLATFORM SECURED ACCOUNT

# SO WHAT HAPPENS? HOW CAN WE DEFEND AGAINST WATERING HOLES & CRYPTOJACKING?

- ▶ Attacker compromises a well known legitimate website
- ▶ He injects HTML elements or scripts in the website so that it redirects users to an exploitation server
- ▶ The exploitation server uses previously unknown or unpatched exploits to attack the redirected users and plant malware in their devices
- ▶ In the case of cryptojacking attack the injected HTML simply uses user's devices resources to mine for crypto

## WHAT IS A WATERING HOLE ATTACK?

- ▶ Usually, they are detected either during the exfiltration phase or when a user notices his device's resources going nuts
- ▶ Usually, this happens too late
- ▶ Most of the times they go undetected

HOW ARE THESE ATTACKS DETECTED  
BY TRADITIONAL MEANS?



- ▶ Firewalls won't catch it because the attack is executed via an existing, permitted session
- ▶ There is no signature for the exploit or the dropped malware
- ▶ The HTML elements are usually highly obfuscated so no signature can be generated also
- ▶ They only target specific IP or device configuration
- ▶ Newer cryptojacking attacks are either obfuscating contacting domains, using proxies and also are very careful with the resources they use on exploited [devices](#)

## WHY ARE THESE ATTACKS SO HARD TO DETECT & MITIGATE?



- ▶ We analyzed over 40,000 malicious Javascript and HTML elements
- ▶ We extracted a set of features that are common based on the probability of appearance of certain instructions and coding styles and relations between them
- ▶ We trained our machine learning algorithm using 70% of the malicious HTML elements and Javascripts and HTML elements from ALEXA TOP100 websites
- ▶ We tested our neural network with the remaining 30% of malware samples
- ▶ For further analysis we run the suspicious samples in a sandbox network
- ▶ The knowledge is used to further train the machine learning algorithm
- ▶ 100% detection rate, 1% false positive, 0% false negative

## WHAT IS DEKENEAS?



# Thank you

