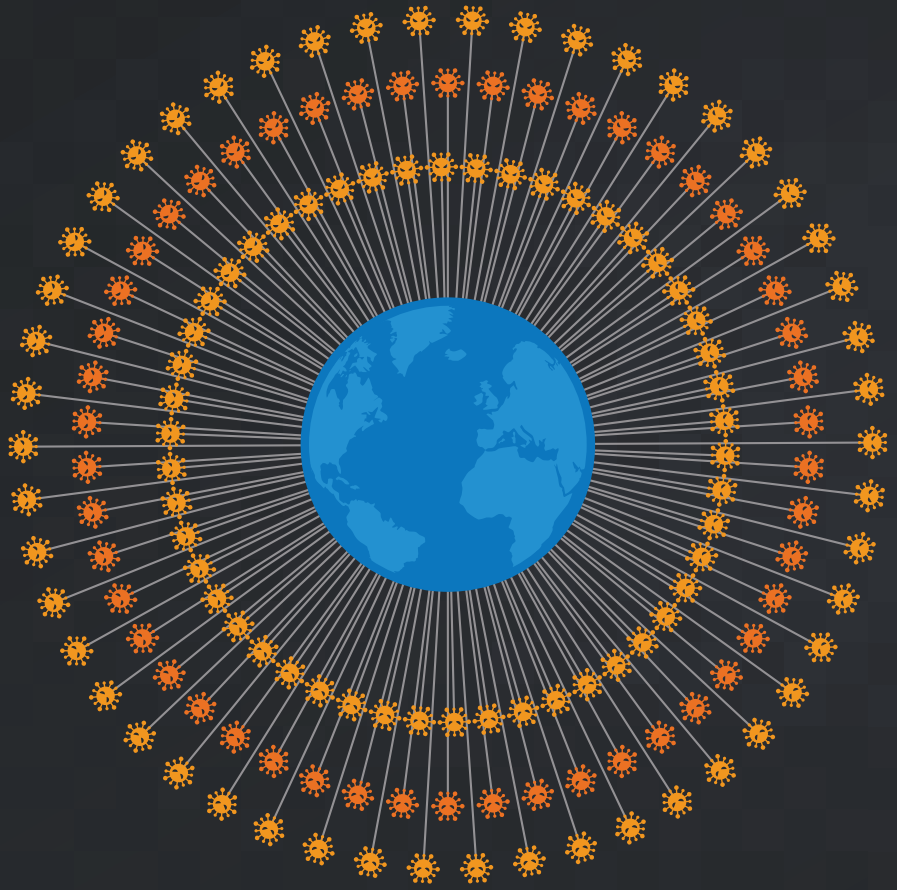
Mesterséges intelligencia alapú védelemi technológiák

HBONE Workshop 2019.

Ács György
IT biztonsági konzulens
2019. február 14.

Threat Landscape



= 10,000

1.5 Million Unique

Malware samples **DAILY**

TALOS

“Information Security is Becoming a Big Data Problem”

by Neil MacDonald, VP & Gartner Fellow

Témák

- Néhány szó az elméletről
- MI a rossz oldalon
- MI technológiák a jó oldalon
  - File vizsgálat és osztályozás
  - Ezt a domaint mire fogják használni?
  - Ismerjük fel a malware-t, visszatitkosítás nélkül

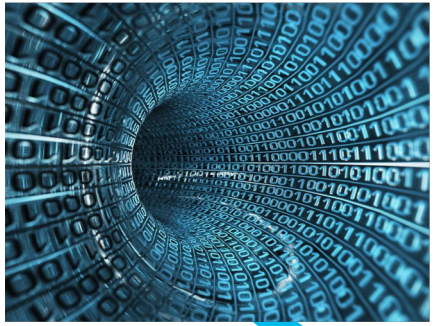
Néhány szó az elméletről

„Amíg van esélyünk megérteni, hogyan működik az MI, érdemes lenne megkísérelni megérteni!”

Ács György

Miből áll a mesterséges intelligencia?

- Adat



Algoritmus/Software

Nagy számítási kapacitás (tanítás)

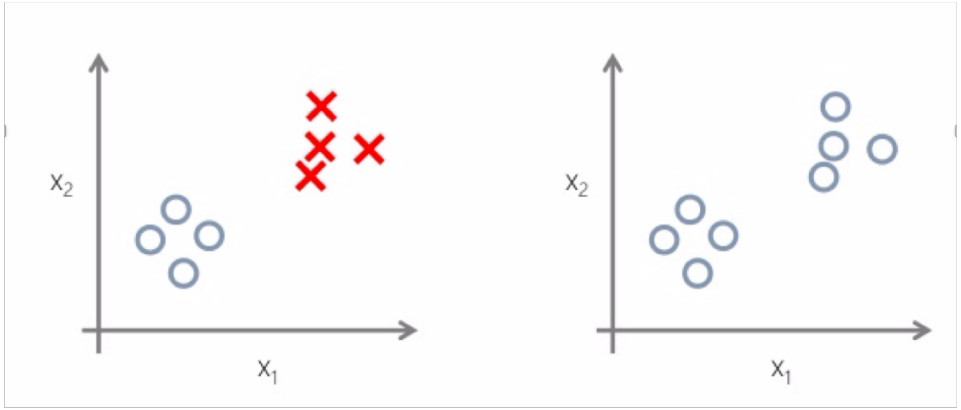
Prediction, classification, pattern discovery

scikit learn

“Feature” – ”lényegkiemelt paraméterek”, Az adat kritikus : ”DATA IS KING”
Rossz adat -> rossz végeredmény, “sok” adat kell jó végeredményhez,
adattisztítás szükséges

Gépi tanulás

- Típusok:
  - Supervised L. (felügyelt t.)
  - Unsupervised Learning
  - Reinforcement L. (megerősítéses t.)
  - Deep Learning
- Lépések:
  - „Feature” kiválasztás
  - Tanítás (címkézett adat)
  - Becslés – Prediction

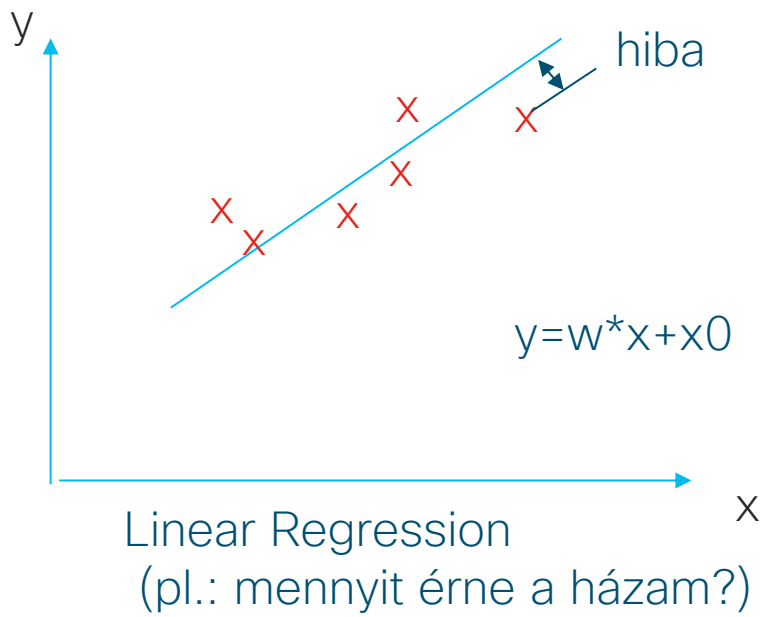


Supervised

Unsupervised

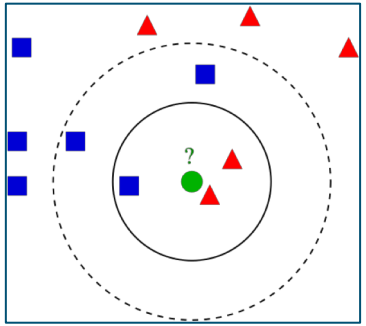
Supervised Learning - Felügyelt tanulás

- **Osztályozás (classification)**: csoportok hozzárendelése az új adathoz az előző adatok alapján
- **Regresszió**: a bementi adatból ”előrejelzi” az “y” értékeket a korábbi adatok alapján
- Minden adatnak **feature**-ei vannak (x1, x2, ..xn) és **címkéje** (y, kimenet)
- A training adat: az adat címkéje ismert

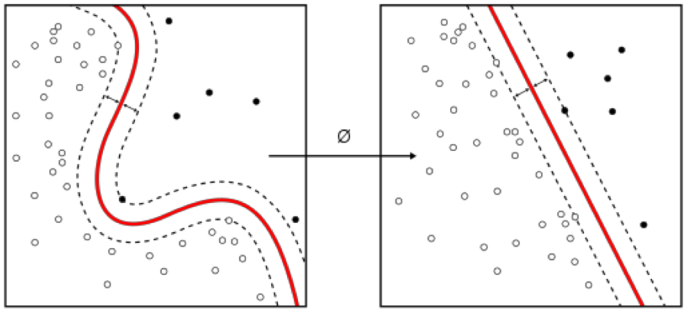


Supervised Learning – Felügyelt tanulás

- Típusai:
  - Linear Regression
  - Logistic Regression (2 kimenet lehet (malware/nem malware), pl.:joy)
  - K-Nearest Neighbor (KNN)
  - Support Vector Machine (SVM)
  - Decision trees and random forests (döntési fát épít a változókból)
  - És még sok

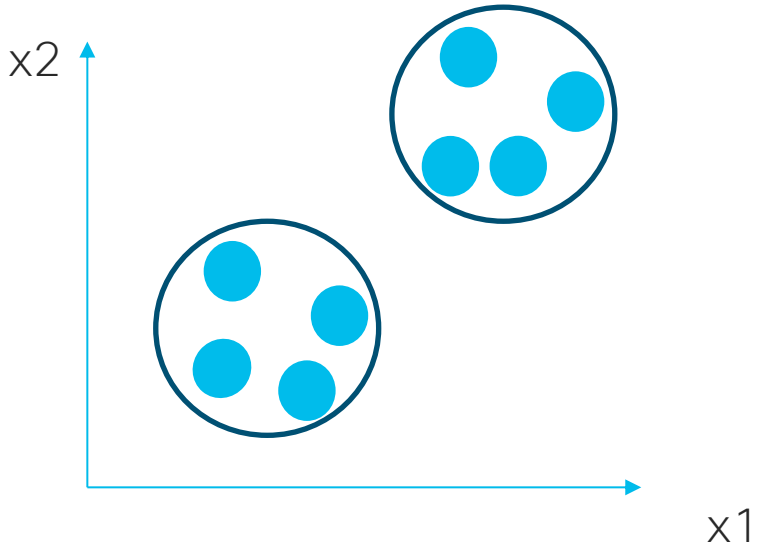


Ha k=3, akkor piros,
ha k=5, akkor kék



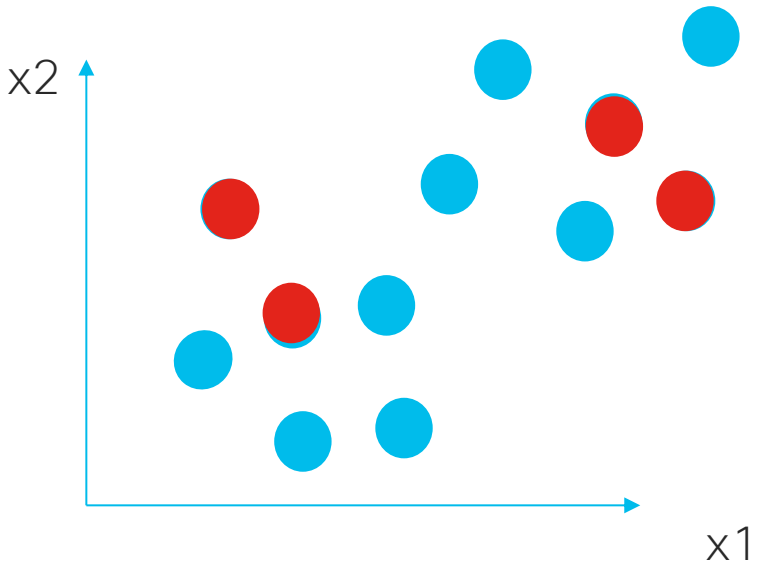
Unsupervised Learning – Felügyelet nélküli tanulás
Címkézetlen adatok vannak csak!

- **Klaszterezés**:
  - az adatok csoportosítása hasonlóság vagy különbség alapján
- **Kapcsolatok:**
  - keressünk szabályokat, amik leírják az adatokat.
  - “Aki PC-t vesz, vesz memóriát is.”



Unsupervised Learning – Felügyelet nélküli tanulás
Címkézetlen adatok vannak csak!

- Pl.: **K-means** (k-közép) :
- Próbáljuk meg ”k” csoportra osztani az adatokat
  - **Algoritmus**:
    - Válasszunk ki “k” random pontot középpontként
    - Adjuk meg az összes pontra a közelség alapján melyik középponthoz tartozik
    - Központok frissítése csoporton belül
    - Ismételje addig, amíg a centrumok nem változnak

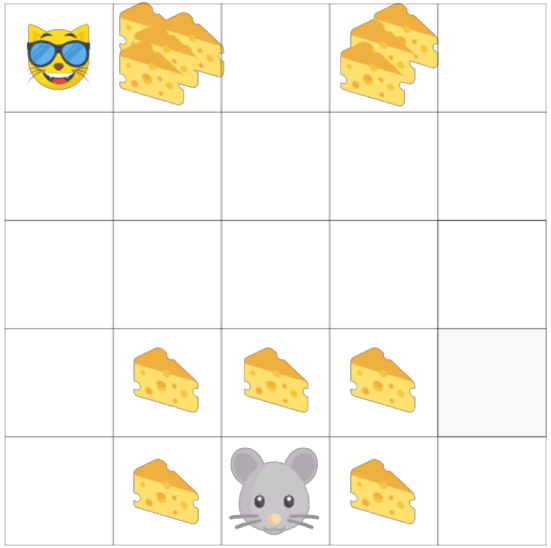



0
1:12
0
K/D/A 0/0/0
LH/DN 0/0
Axe
1
Witch Doctor
1
Dragon Knight
1
Drow Ranger
1
1
JUGGERNAUT
50
3
300
20
26
14
600 / 600
+25.7
229 / 229
+23.7
1
Drag items to add to quick buy
625


Elon Musk
@elonmusk
Follow
OpenAI first ever to defeat world's best players in competitive eSports. Vastly more complex than traditional board games like chess & Go.
5:15 PM - 11 Aug 2017
4,444 Retweets 15,026 Likes
487
4.4K
15K

Reinforcement Learning – Megerősítéses tanulás

- A világ “felfedezése”, a cselekvések “jutalmazása” adott
- Véletlenszerű lépésekkel kezdődik, minden lépésnél a jutalom eltárolása (lehet +/–, a “karakter” meghalt)
- Hozzunk létre egy stratégiát, amely magasabb jutalmat eredményez
- Folytassuk a stratégia javítását ”sok” tapasztalattal
- Példák: AlphaGo, OpenAI; magukkal több ezer játékot játszottak, amíg meg nem tanultak egy olyan stratégiát, amely közel “tökéletes” a komplex játékokban



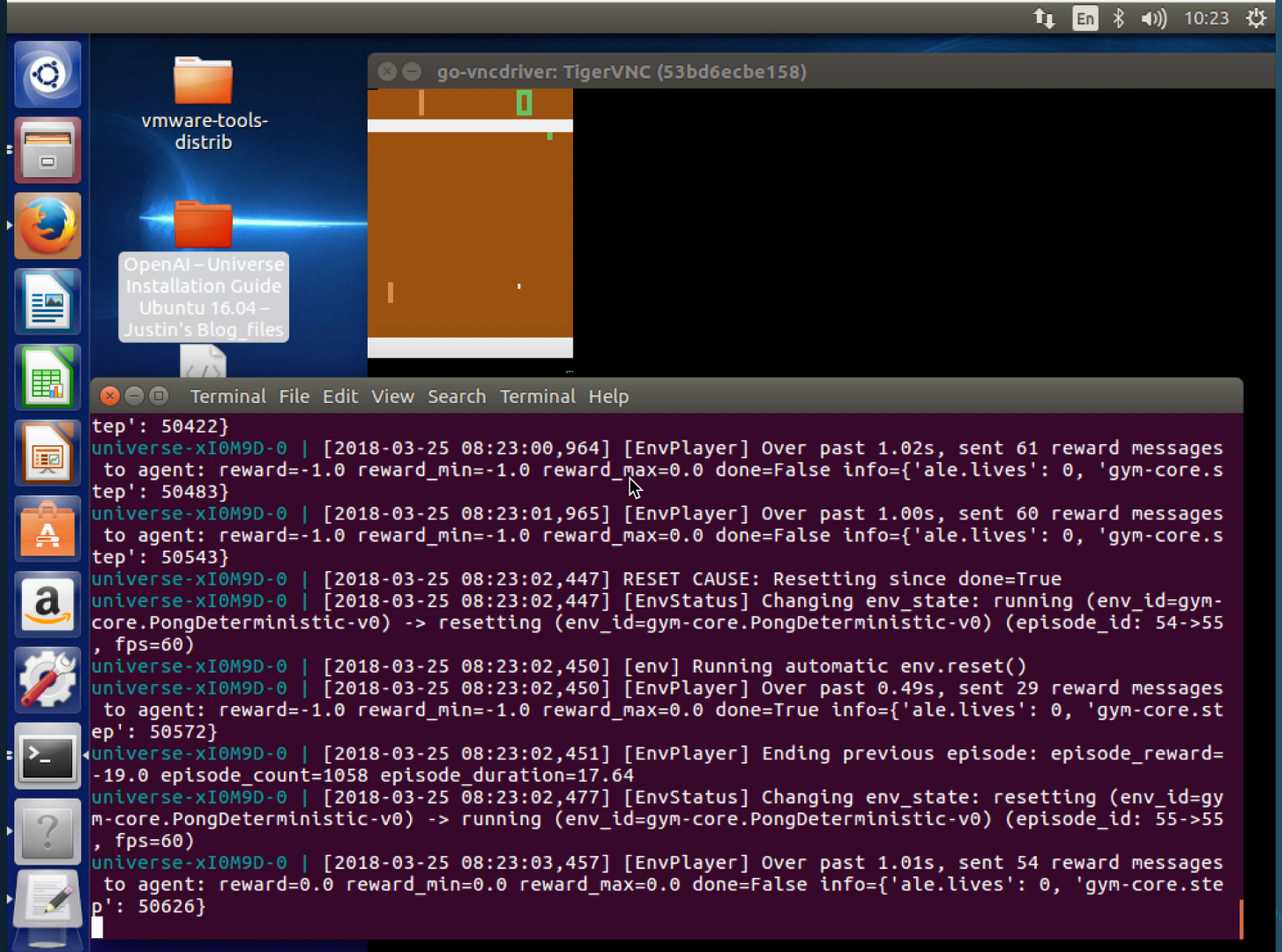
https://repl.it/@GyorgyAcs/CartPole

Deep Learning

- A legmodernebb a gépi tanulás
- Leginkább felügyelt tanuláshoz hasonlít
- Neurális hálózatokat használ (az első alkalmazás: agyi neuronok modellezésére)
- Bonyolultabb
- Több adatra van szükség
- Több időt igényel a tanítás

- Kép-, beszédfelismerés
- Natural Language Processing
- Machine translation (DeepL)
- Önvezető autók
- Stb.

MI a rossz oldalon


En
10:23
vmware-tools-distrib
OpenAI – Universe Installation Guide Ubuntu 16.04 – Justin's Blog_files
go-vncdriver: TigerVNC (53bd6ecbe158)
Terminal File Edit View Search Terminal Help
tep': 50422}
universe-xI0M9D-0 | [2018-03-25 08:23:00,964] [EnvPlayer] Over past 1.02s, sent 61 reward messages to agent: reward=-1.0 reward_min=-1.0 reward_max=0.0 done=False info={'ale.lives': 0, 'gym-core.step': 50483}
universe-xI0M9D-0 | [2018-03-25 08:23:01,965] [EnvPlayer] Over past 1.00s, sent 60 reward messages to agent: reward=-1.0 reward_min=-1.0 reward_max=0.0 done=False info={'ale.lives': 0, 'gym-core.step': 50543}
universe-xI0M9D-0 | [2018-03-25 08:23:02,447] RESET CAUSE: Resetting since done=True
universe-xI0M9D-0 | [2018-03-25 08:23:02,447] [EnvStatus] Changing env_state: running (env_id=gym-core.PongDeterministic-v0) -> resetting (env_id=gym-core.PongDeterministic-v0) (episode_id: 54->55, fps=60)
universe-xI0M9D-0 | [2018-03-25 08:23:02,450] [env] Running automatic env.reset()
universe-xI0M9D-0 | [2018-03-25 08:23:02,450] [EnvPlayer] Over past 0.49s, sent 29 reward messages to agent: reward=-1.0 reward_min=-1.0 reward_max=0.0 done=True info={'ale.lives': 0, 'gym-core.step': 50572}
universe-xI0M9D-0 | [2018-03-25 08:23:02,451] [EnvPlayer] Ending previous episode: episode_reward=-19.0 episode_count=1058 episode_duration=17.64
universe-xI0M9D-0 | [2018-03-25 08:23:02,477] [EnvStatus] Changing env_state: resetting (env_id=gym-core.PongDeterministic-v0) -> running (env_id=gym-core.PongDeterministic-v0) (episode_id: 55->55, fps=60)
universe-xI0M9D-0 | [2018-03-25 08:23:03,457] [EnvPlayer] Over past 1.01s, sent 54 reward messages to agent: reward=0.0 reward_min=0.0 reward_max=0.0 done=False info={'ale.lives': 0, 'gym-core.step': 50626}

Malware gyártás MI-val

Security

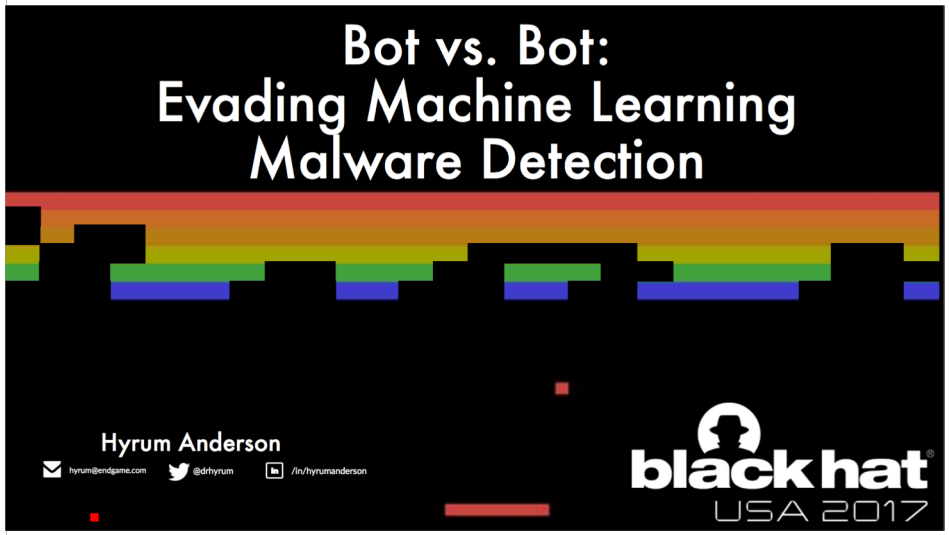
AI quickly cooks malware that AV software can't spot

Experiment used Elon Musk's OpenAI framework - no wonder he's so down on AI

By Iain Thomson in San Francisco 31 Jul 2017 at 07:02 38 SHARE ▼

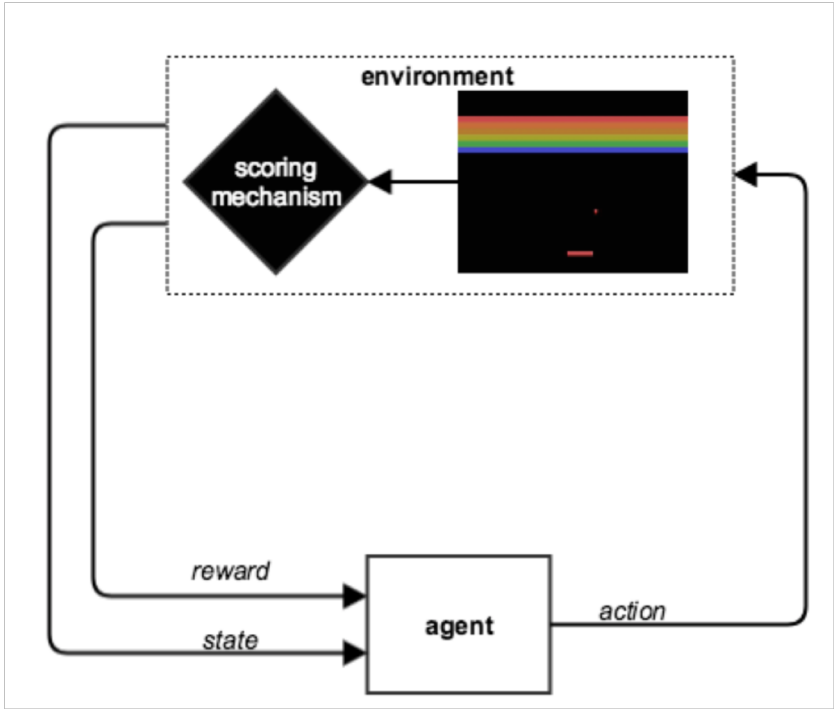
DEF CON Machine-learning tools can create custom malware that defeats antivirus software.

In a keynote demonstration at the DEF CON hacking convention Hyrum Anderson, technical director of data science at security shop Endgame, showed off research that his company had done in adapting Elon Musk's OpenAI framework to the task of creating malware that security engines can't spot.



https://www.theregister.co.uk/2017/07/31/ai_defeats_antivirus_software/

Atari játékok helyett malware



Környezet:
- Fallabda játék több téglasorral
- Vezérlés: jobbra, balra vagy marad
- Tégla lebontása a cél

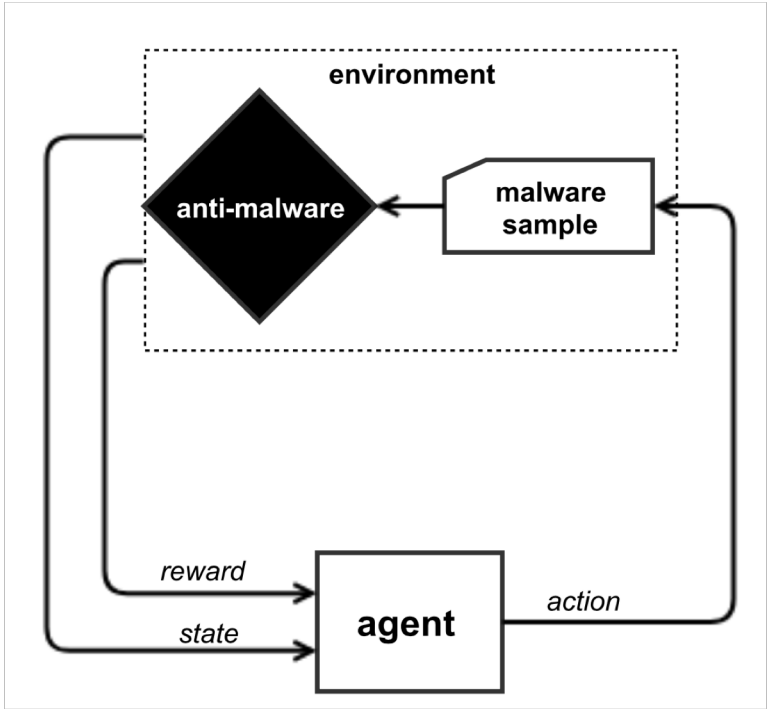
“Agent”:
- Bemenet : képpont pixelek
- Kimenet: akció (balra, jobbra)

Tanulás: több 1000 játék

Hová kellene lépni?
https://gym.openai.com/envs/Breakout-v0

https://www.blackhat.com/docs/us-17/thursday/us-17-Anderson-Bot-Vs-Bot-Evading-Machine-Learning-Malware-Detection.pdf

Atari játékok helyett malware



Környezet:

- Malware minta, Windows PE header
- Mutáció: új belépési pont, új section, véletlen importok, byte-ok, stb.

“Agent”:

- Bemenet : malware byte-ok
- Kimenet: akció (sztochasztikus)

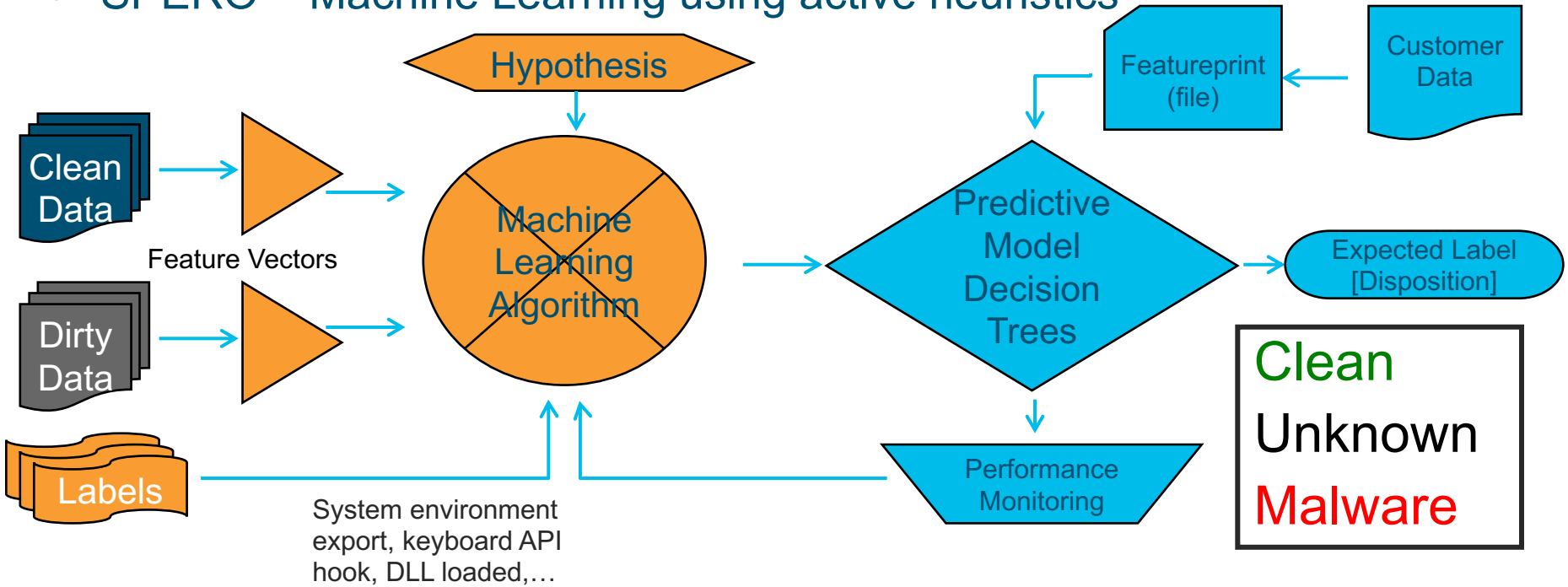
Jutalom: az AV riport szerint tiszta a file

https://www.blackhat.com/docs/us-17/thursday/us-17-Anderson-Bot-Vs-Bot-Evading-Machine-Learning-Malware-Detection.pdf

File vizsgálat és osztályozás

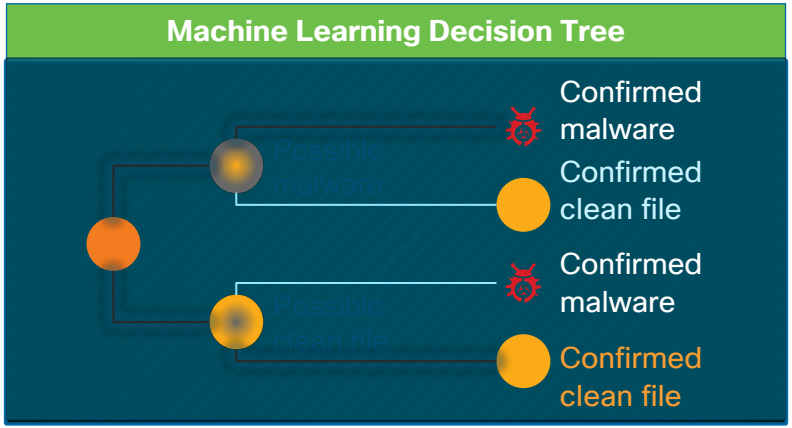
AMP, Advanced Malware Protection
Védelmi rendszer : Spero motor

- SPERO = Machine Learning using active heuristics



Védelmi rendszer : Spero vizsgálati motor

- AMP címkék = a végrehajtás során kapott jellemzők
  - Hálózati kapcsolatok?
  - Nem szabványos protokoll –alkalmazás
  - Hooking? Milyen API-kat használ?
  - Filerendszer változtatás?
    - Másolja magát
    - File-ok mozgatása
  - Más processzek indítása?



- Több, mint 400 jellemzőt elemez - azonosítja a malware-t

Jósoljuk meg, hogy Domi fog-e focizni!

- Nehéz így átlátni
- „oszd meg és uralkodj”
  - Vágjuk részhalmazokra
  - „Tiszta” halmaz? (Csak igen vagy nem)
  - Ha igen : állj
  - Ha nem : ismételd
- Nézzük, hogy az új adat melyik részhalmazba esik

Tanító minták:

9 Igen / 5 Nem

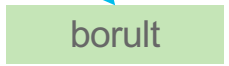
attributes

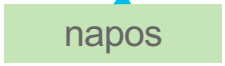
nap	időj.	pára.	szél	játék
D1	napos	magas	gyen.	Nem
D2	napos	magas	erős	Nem
D3	borult	magas	gyen.	Igen
D4	eső	magas	gyen.	Igen
D5	eső	normál	gyen.	Igen
D6	eső	normál	erős	Nem
D7	borult	normál	erős	Igen
D8	napos	magas	gyen.	Nem
D9	napos	normál	gyen.	Igen
D10	eső	normál	gyen.	Igen
D11	napos	normál	erős	Igen
D12	borult	magas	erős	Igen
D13	borult	normál	gyen.	Igen
D14	eső	magas	erős	Nem
D15	eső	magas	gyen.	???

Új adat:

https://www.youtube.com/watch?v=eKD5gxPPeY0

9 Igen / 5 Nem





nap	időj.	pára.	szél	játék
D1	napos	magas	gyen.	Nem
D2	napos	magas	erős	Nem
D8	napos	magas	gyen.	Nem
D9	napos	normál	gyen.	Igen
D11	napos	normál	erős	Igen

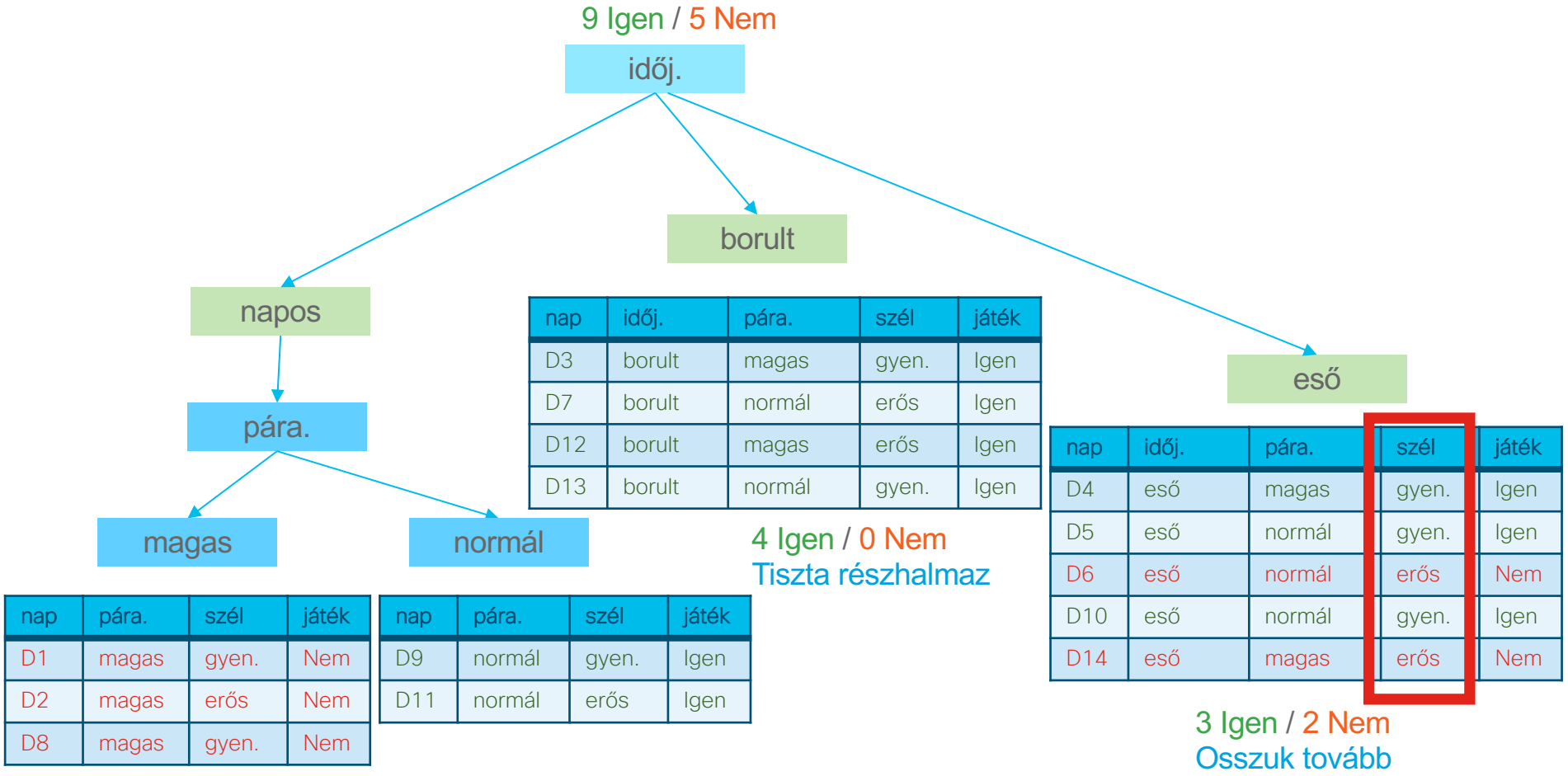
2 Igen / 3 Nem
Osszuk tovább

nap	időj.	pára.	szél	játék
D3	borult	magas	gyen.	Igen
D7	borult	normál	erős	Igen
D12	borult	magas	erős	Igen
D13	borult	normál	gyen.	Igen

4 Igen / 0 Nem
Tiszta részhalmaz

nap	időj.	pára.	szél	játék
D4	eső	magas	gyen.	Igen
D5	eső	normál	gyen.	Igen
D6	eső	normál	erős	Nem
D10	eső	normál	gyen.	Igen
D14	eső	magas	erős	Nem

3 Igen / 2 Nem
Osszuk tovább



nap	időj.	pára.	szél	játék
D3	borult	magas	gyen.	Igen
D7	borult	normál	erős	Igen
D12	borult	magas	erős	Igen
D13	borult	normál	gyen.	Igen

nap	pára.	szél	játék
D1	magas	gyen.	Nem
D2	magas	erős	Nem
D8	magas	gyen.	Nem

nap	pára.	szél	játék
D9	normál	gyen.	Igen
D11	normál	erős	Igen

nap	időj.	pára.	szél	játék
D4	eső	magas	gyen.	Igen
D5	eső	normál	gyen.	Igen
D6	eső	normál	erős	Nem
D10	eső	normál	gyen.	Igen
D14	eső	magas	erős	Nem

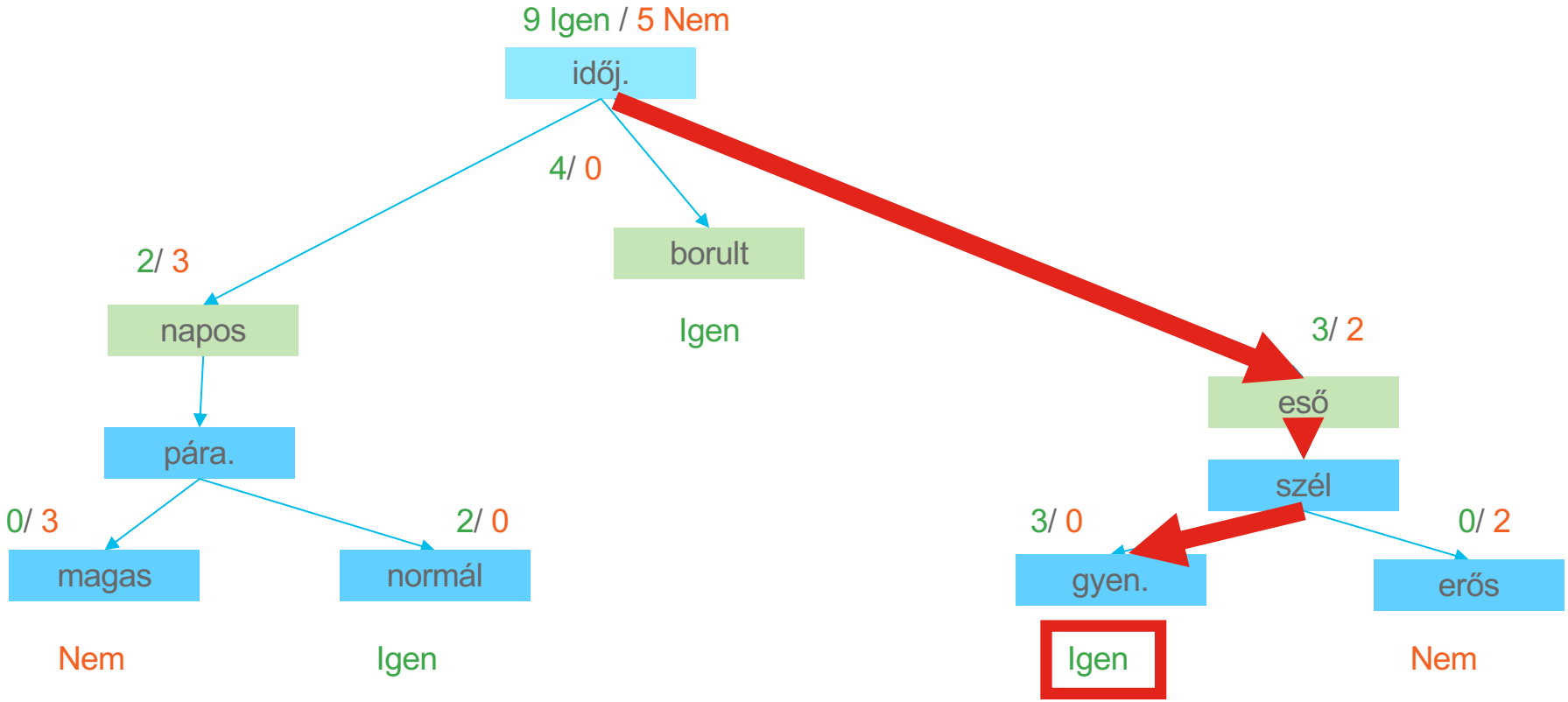
nap	időj.	pára.	szél	játék
D3	borult	magas	gyen.	Igen
D7	borult	normál	erős	Igen
D12	borult	magas	erős	Igen
D13	borult	normál	gyen.	Igen

nap	pára.	szél	játék
D1	magas	gyen.	Nem
D2	magas	erős	Nem
D8	magas	gyen.	Nem

nap	pára.	szél	játék
D9	normál	gyen.	Igen
D11	normál	erős	Igen

nap	időj.	pára.	játék
D4	eső	magas	Igen
D5	eső	normál	Igen
D10	eső	normál	Igen

nap	időj.	pára.	játék
D6	eső	normál	Nem
D14	eső	magas	Nem



Új adat:

D15	eső	magas	gyen.	Igen

Random Forests
menjünk a véletlen erdőbe!

https://colab.research.google.com/drive/1h811s2qow3a857uHnB30jofcWE-bhjC_

Ezt a domaint mire fogják használni?

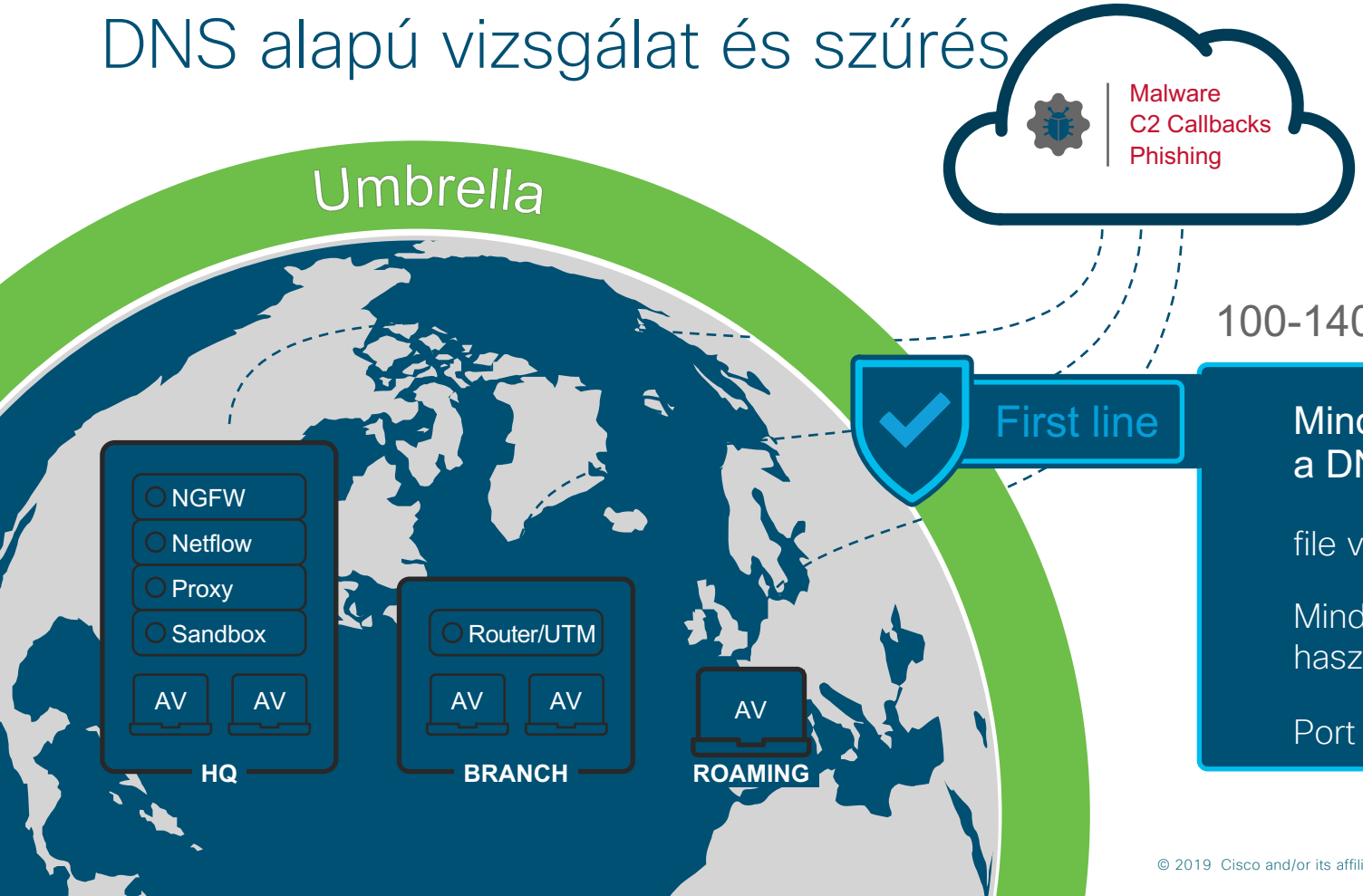
A kibertámadások elemzése

- Felderítés és infrastruktúra kiépítés
- Domain regisztráció, IP, ASN intelligencia
- Az eredmények alapján monitorozás

- Az első áldozat megjelenik

- A támadás kiterjed
- Nagy mennyiségű áldozat
- Szignatúrát fejlesztenek

DNS alapú vizsgálat és szűrés



100-140 milliárd napi kérés

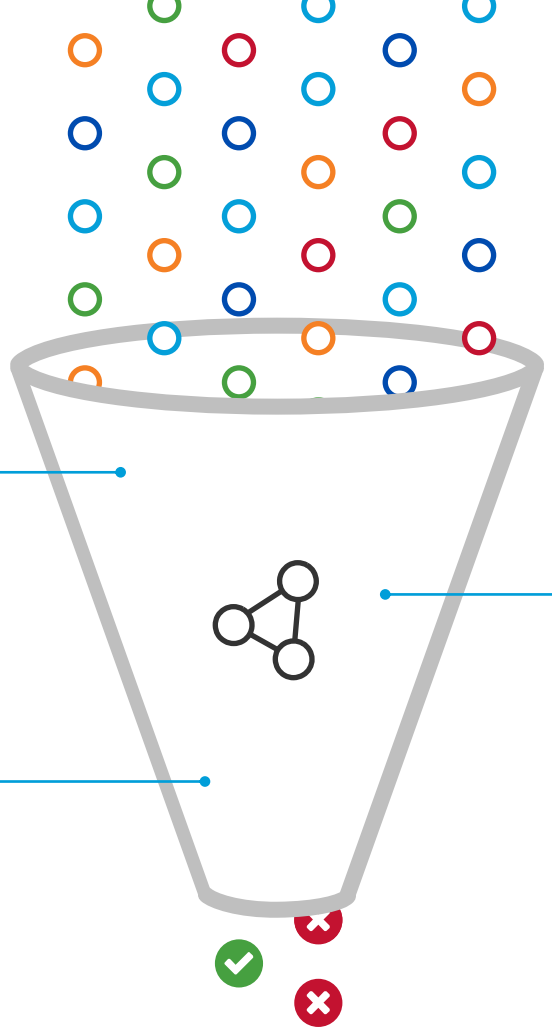
Minden kommunikáció a DNS-sel kezdődik

file vizsgálat proxy-val

Minden eszköz (IoT) használja

Port független

Statisztikai modellek



2M+ live events per second

11B+ historical events

100-140B queries per day

3 million new domains every day

Környezet miatt bűnös

- *Co-occurrence model*
- *IP Geo-Location model*
- Secure rank model
- *Sender rank model*

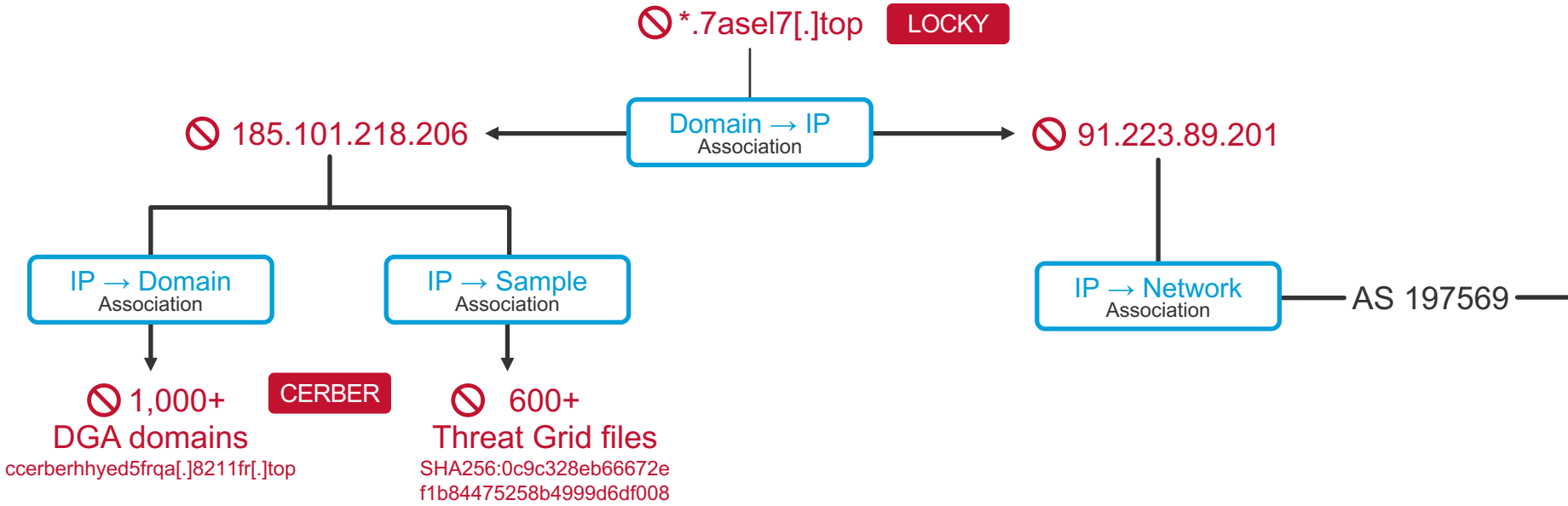
Kapcsolatok miatt bűnös

- Predictive IP Space Modeling

A bűnös mintákat mutat

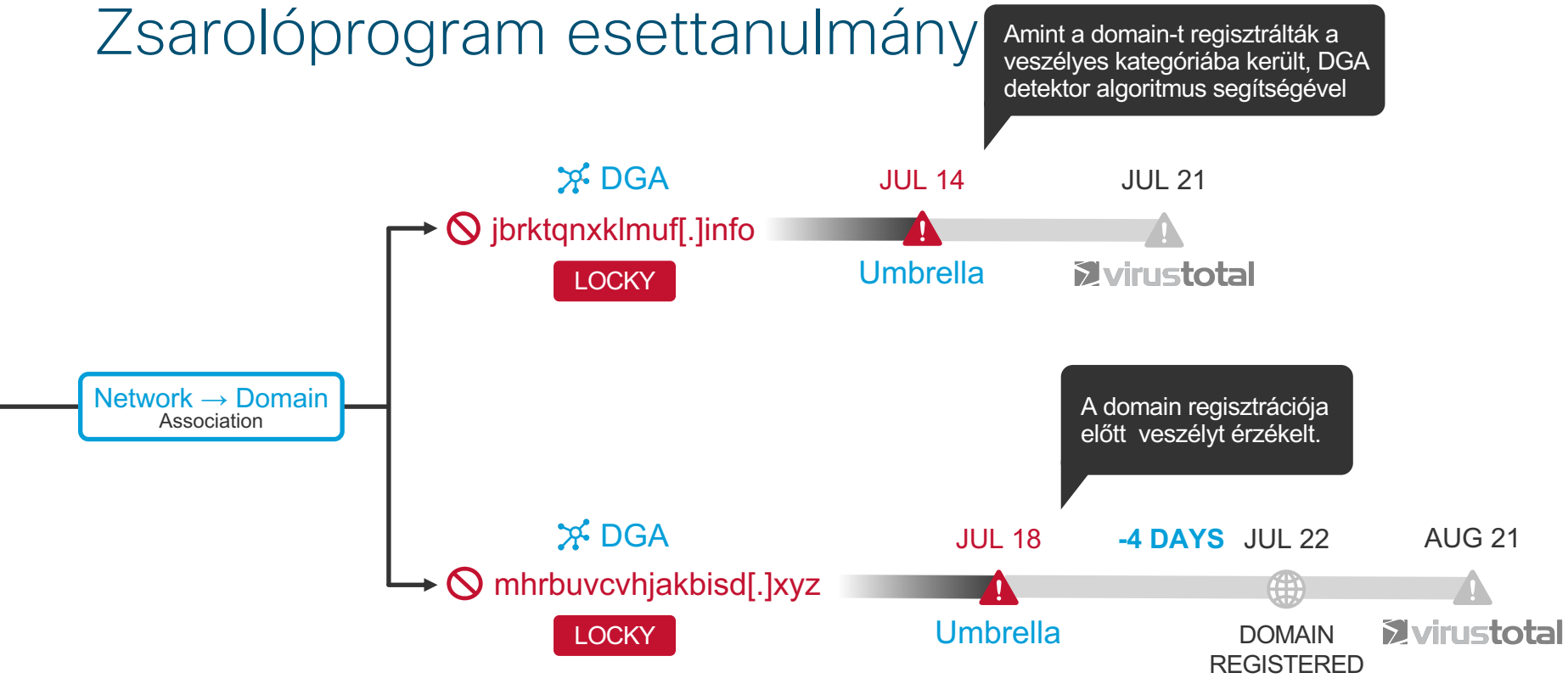
- *Spike rank model*
- *Natural Language Processing rank model*
- Live DGA Prediction

Zsarolóprogram esettanulmány



A Locky és Cerber ugyanazt az infrastruktúrát használja

Zsarolóprogram esettanulmány

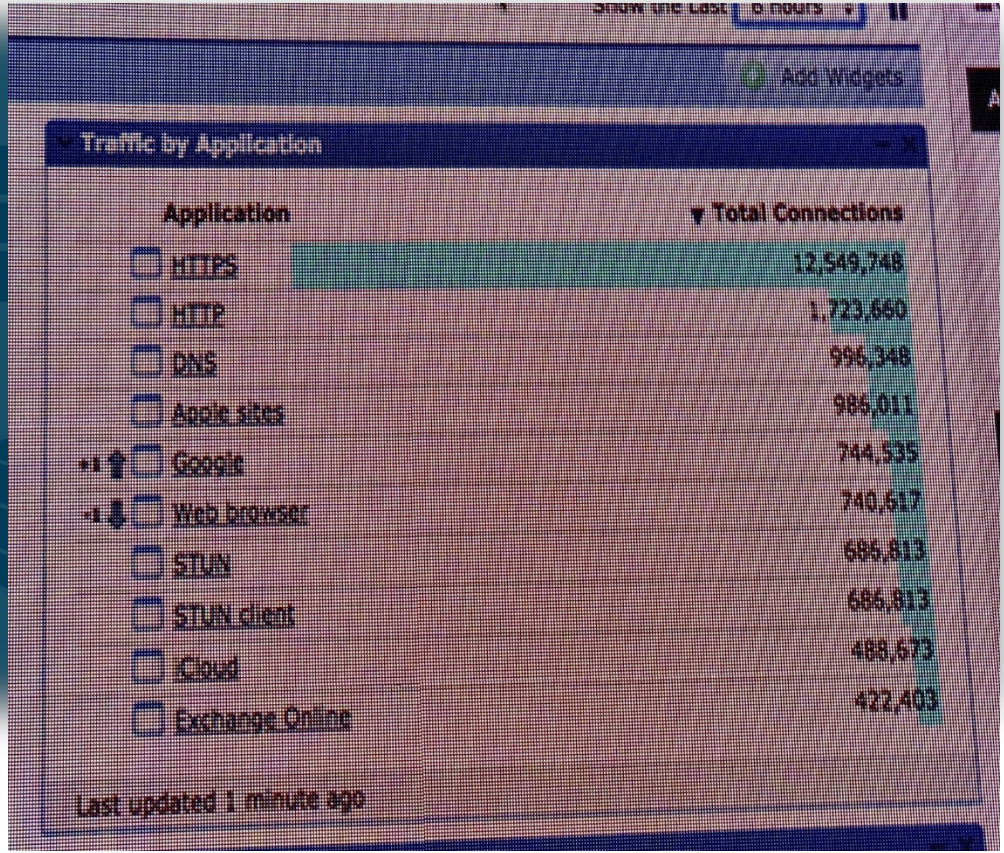


A statisztikus modellek azonosítottak és blokkoltak a DGA algoritmus által generált 2 domaint, jó néhány nappal azelőtt, hogy regisztrálták.

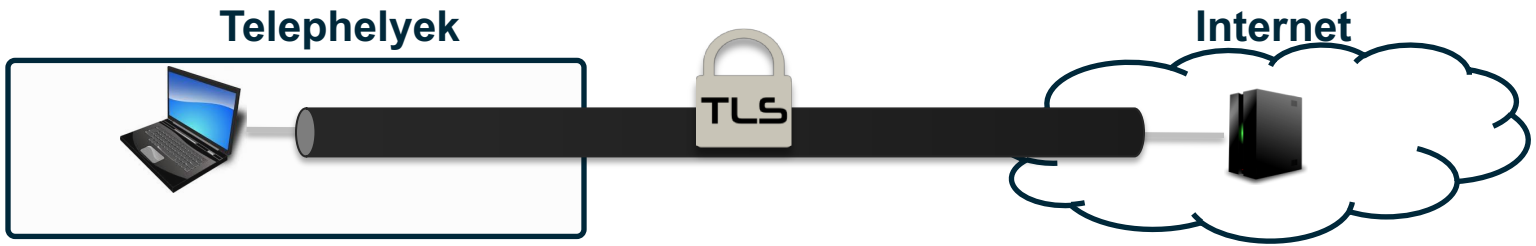
Ismerjük fel a malware-t, visszatitkosítás nélkül

Titkosítás mindenhol

Chrome will start marking all HTTP sites as not secure in July



A probléma: titkosított malware forgalom

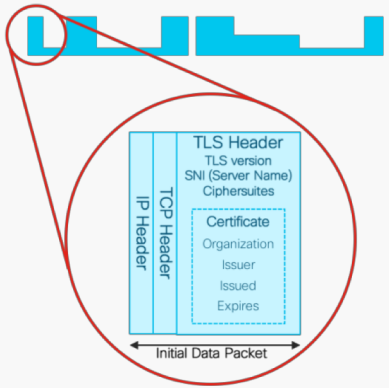


- A TLS használata többségben van már (nem baj!)
- A mintaillesztéses eljárások NEM hatékonyak
- MITM problémák
  - ”Privacy”, jogi, megvalósítási, „drága”, nem együttműködő kliensek, cert pinning, ...

A “jellemzők” : Enhanced NetFlow és intelligencia

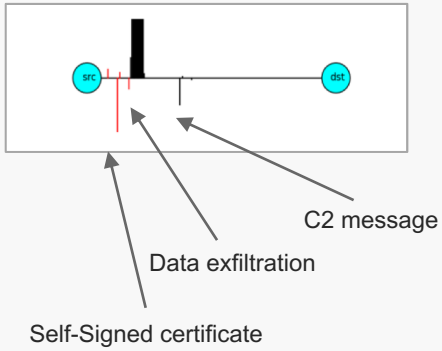
Initial Data Packet

Make the most of the unencrypted fields



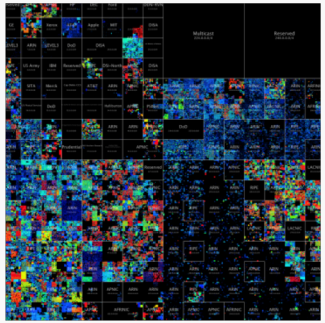
Sequence of Packet Lengths and Times

Identify the content type through the size and timing of packets



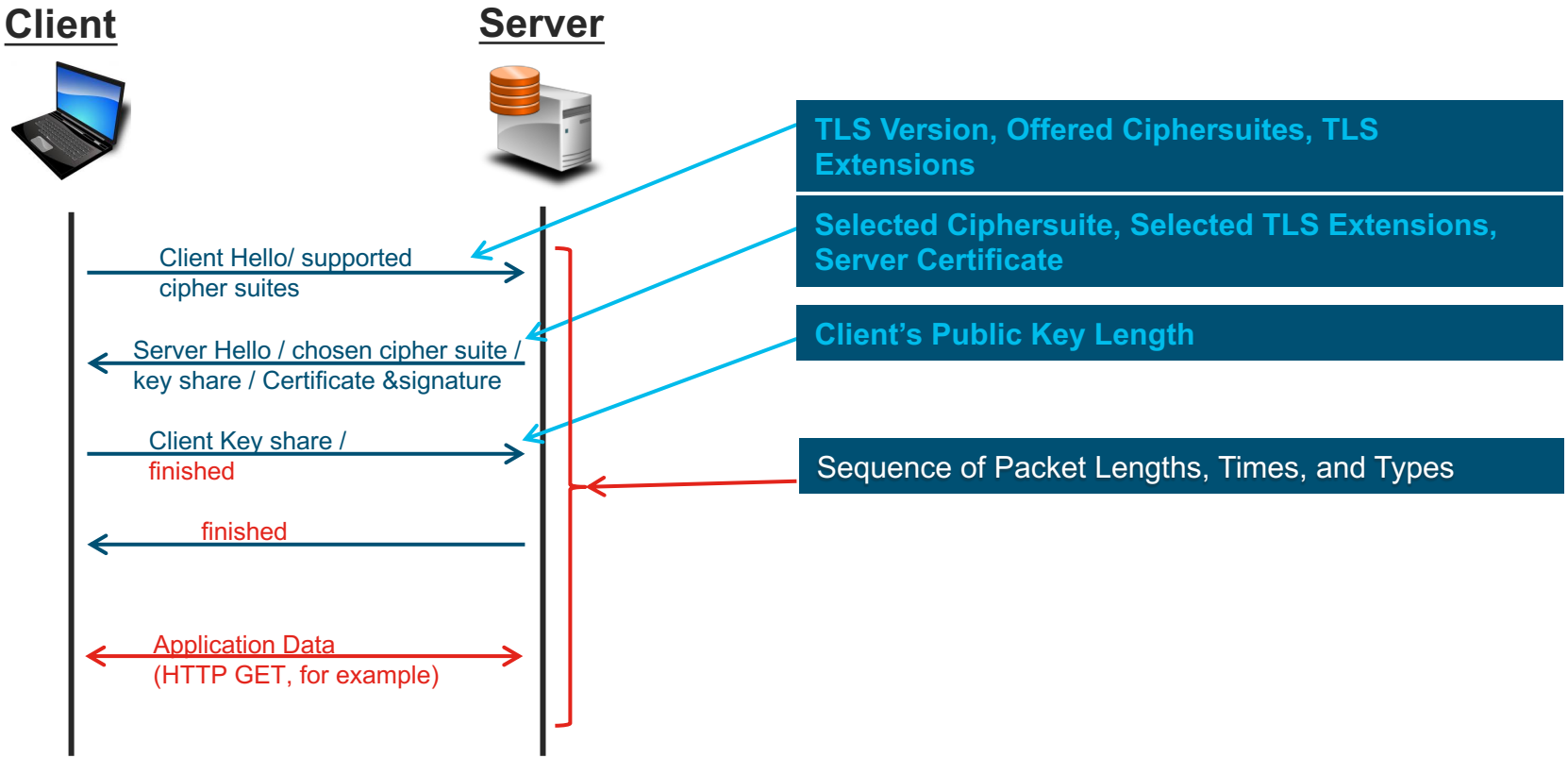
Threat Intelligence Map

Who’s who of the Internet’s dark side



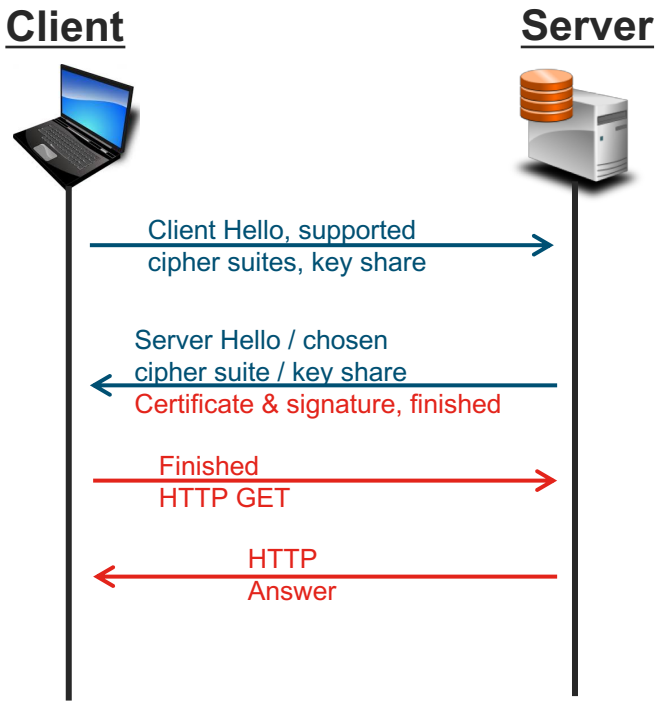
Broad behavioral information about the servers on the Internet.

Initial Data Packet, IDP, TLS 1.2



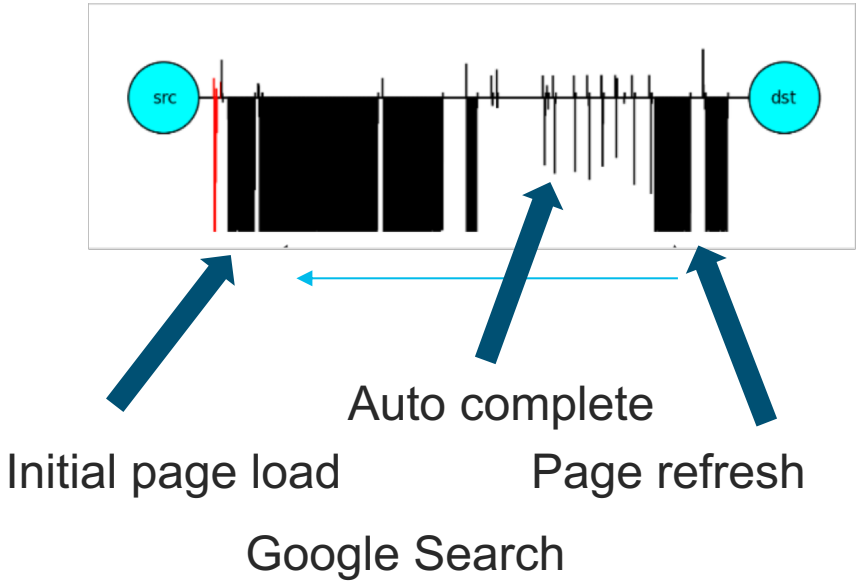
TLS 1.3

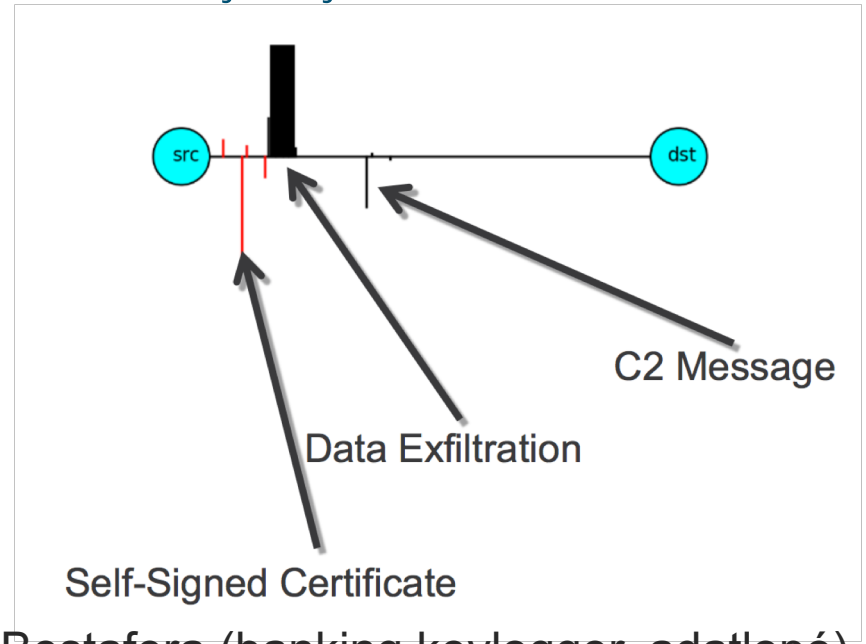
- 2018. március 21-én hagyták jóvá
- Számos biztonsági és gyorsítási fejlesztés
  - nem biztonságos régi protokollok, SHA1, DES, MD5, ... elhagyása
  - two-trip helyett one-trip, sőt, zero-trip (emlékszik a kapcsolatra), gyorsabb kapcsolatfelépítés
  - Szerver tanúsítvány is titkosított
- Böngészők elkezdték támogatni (Chrome, Firefox)
- Az ETA (eddig) gond nélkül kezeli
  - A tanúsítvány infó mégis begyűjthető aktív monitorozással és a CTA/GTA adatbázisa tartalmazza, + SPLT



Sequence of Packet Lengths and Times (SPLT)

- Az alkalmazás/felhasználó viselkedési mintáját írja le

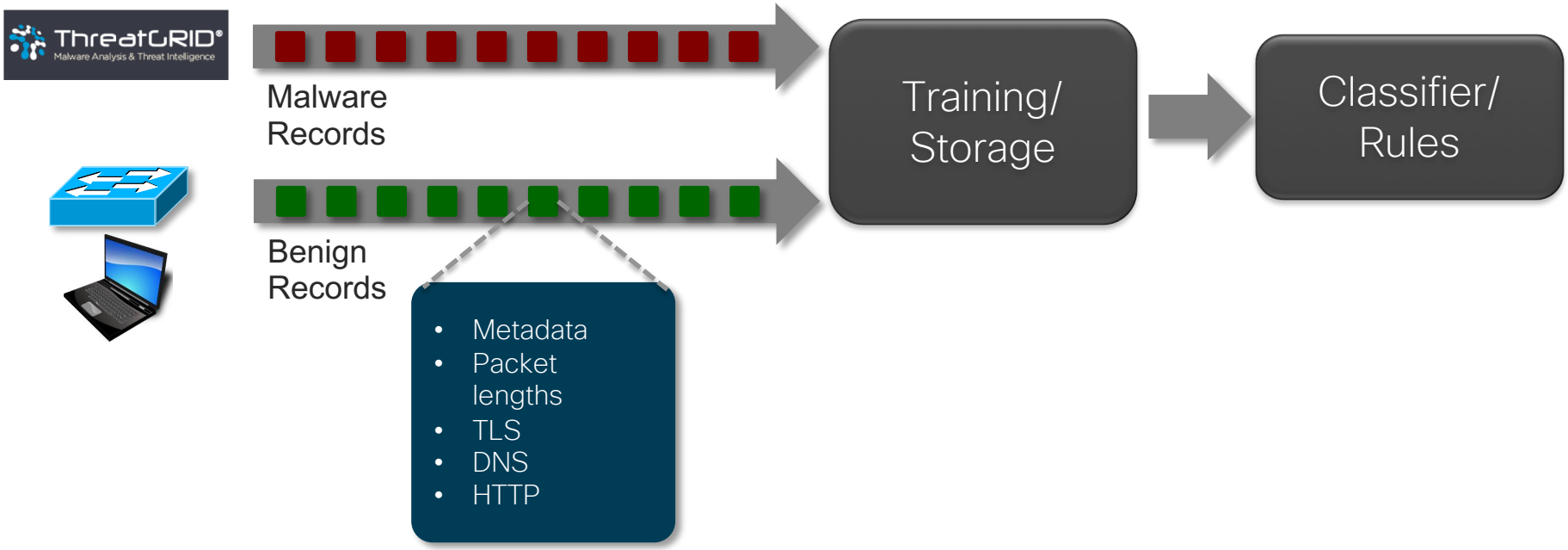




Bestafera (banking keylogger, adatlopó)

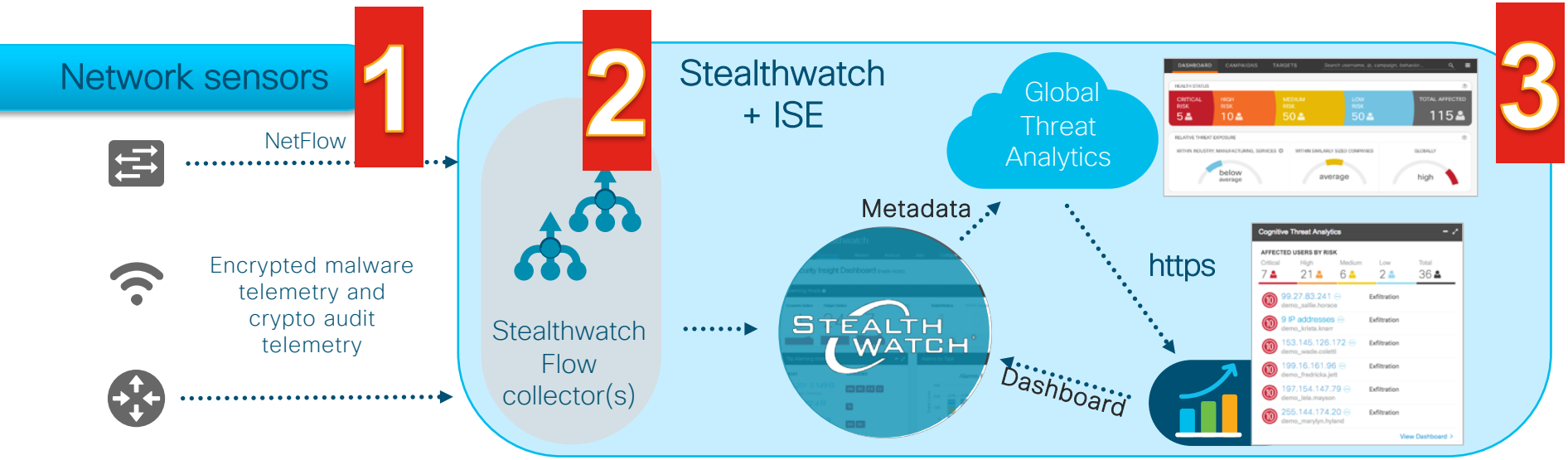
Adatgyűjtés

API, több milliárd flow



https://github.com/cisco/joy

Encrypted Traffic Analytics overview



Cisco's unique hardware and software architecture

Enhanced NetFlow with Encrypted Traffic Analytics from Cisco's newest routers and switches

Stealthwatch enhanced analytics and machine learning reduces threat investigation time

Global-to-local knowledge correlation results in higher precision of threat findings

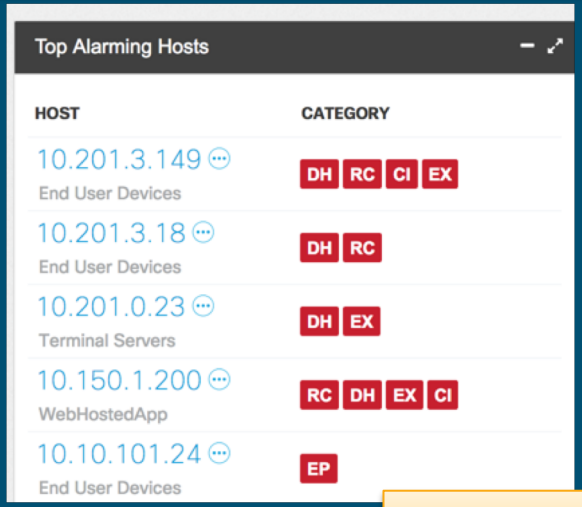
Stealthwatch Enterprise Learning Engines

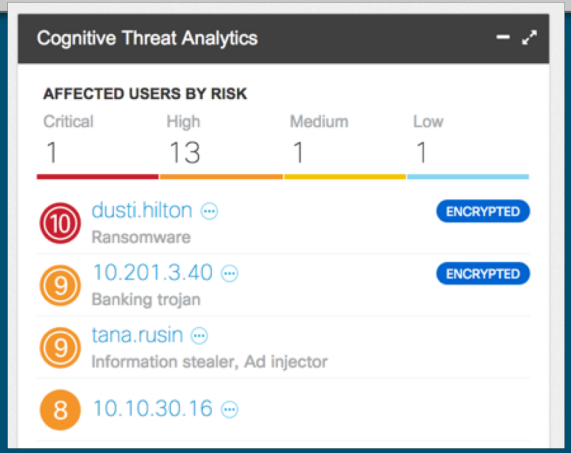
Stealthwatch "On Box"

- Behavioural Analysis
- Anomaly detection through statistical learning
- **Unsupervised Learning Engine**
- User Defined Behaviour Analysis

Global / Cognitive Analytics

- Cloud Hosted
- **Multi-layer Machine Learning**
- Anomaly detection through statistical learning
- **Supervised Learning Engine**
- Malware classification





Stealthwatch Enterprise

Stealthwatch

Dashboards | Monitor | Analyze | Jobs

Desktop Client

Security Insight Dashboard | Inside Hosts

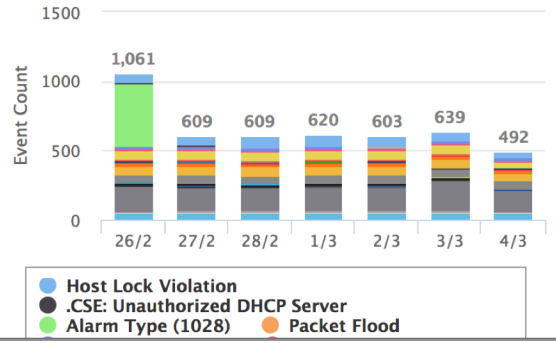
Alarming Hosts

Concern Index	Target Index	Recon	C&C	Exploitation	DDoS Source	DDoS Target	Data Hoarding	Exfiltration	Policy Violation	Anomaly
21	0	124	0	14	0	2	4	1	4	1

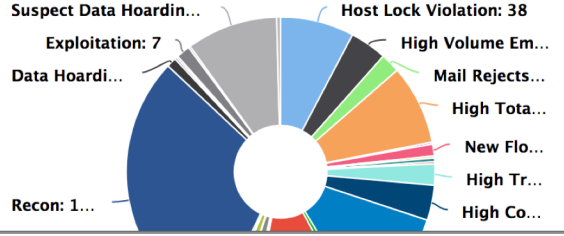
Top Alarming Hosts

HOST	CATEGORY
10.201.3.149	DH RC CI
10.10.30.11	AN
10.201.3.18	DH RC
10.201.0.23	DH
209.182.184.2	RC
10.201.0.15	RC
10.150.1.200	RC DH EX CI

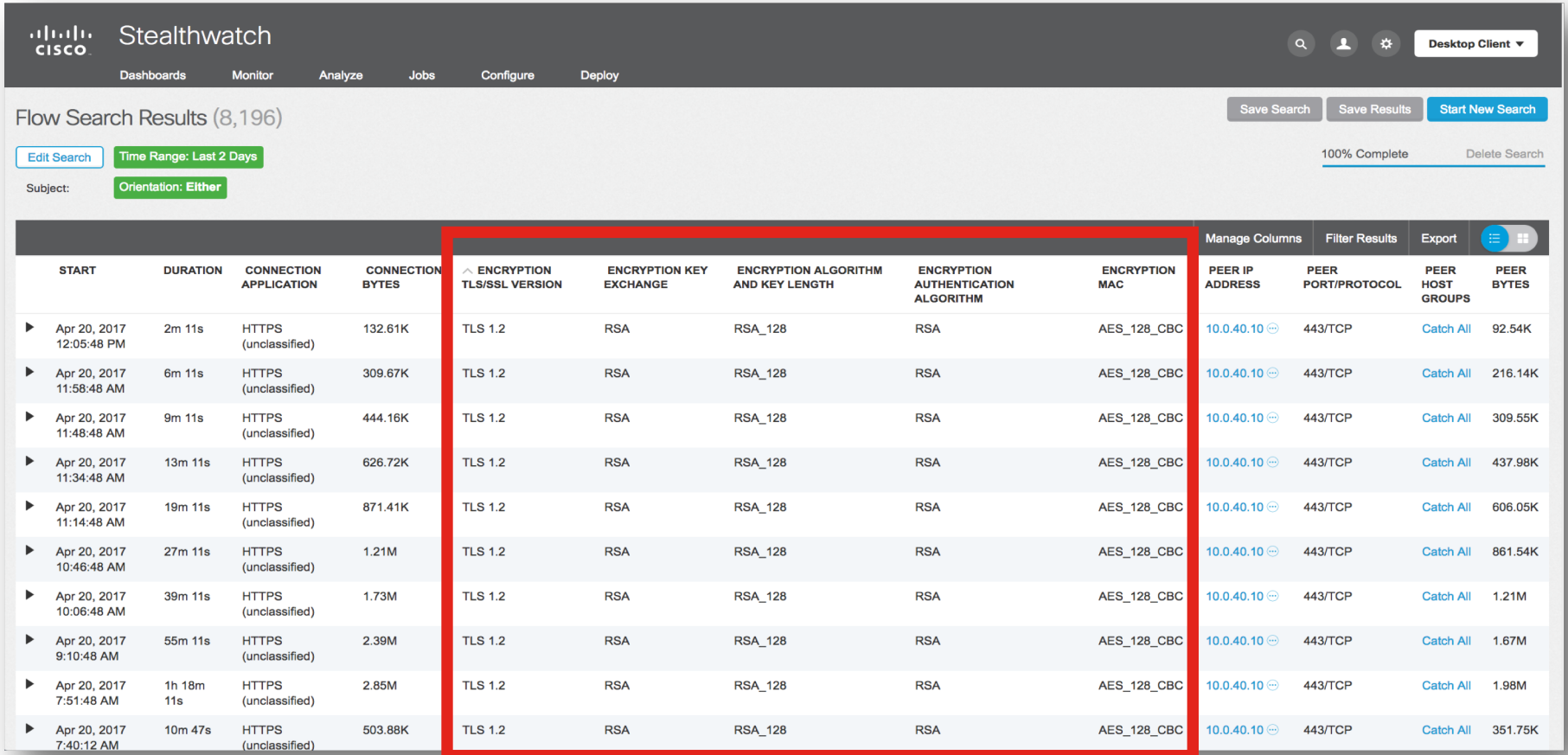
Alarms by Type



Today's Alarms

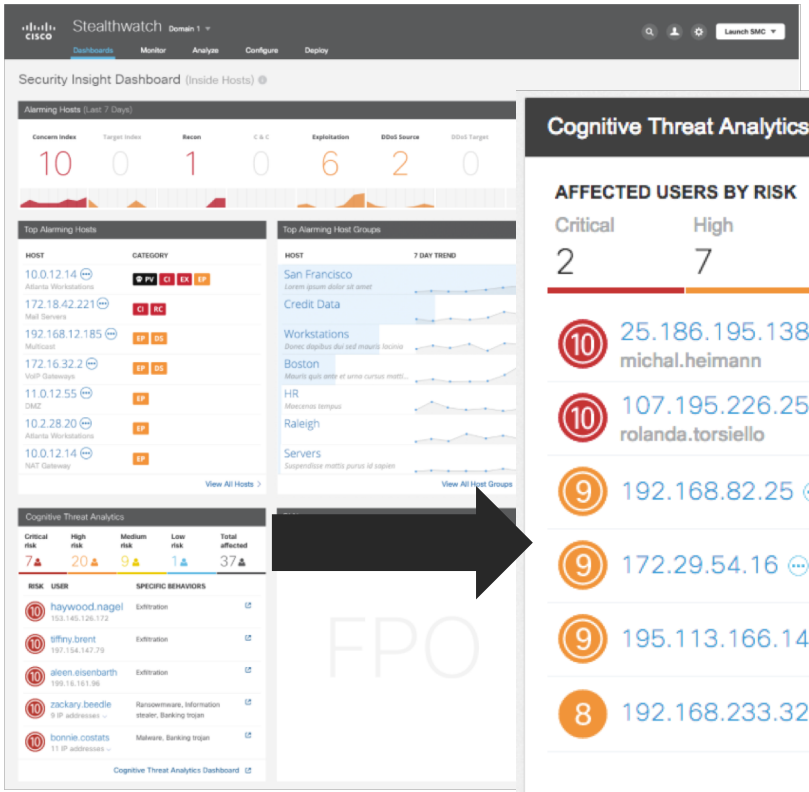


Titkosítási adatok az összes hálózati kapcsolatban

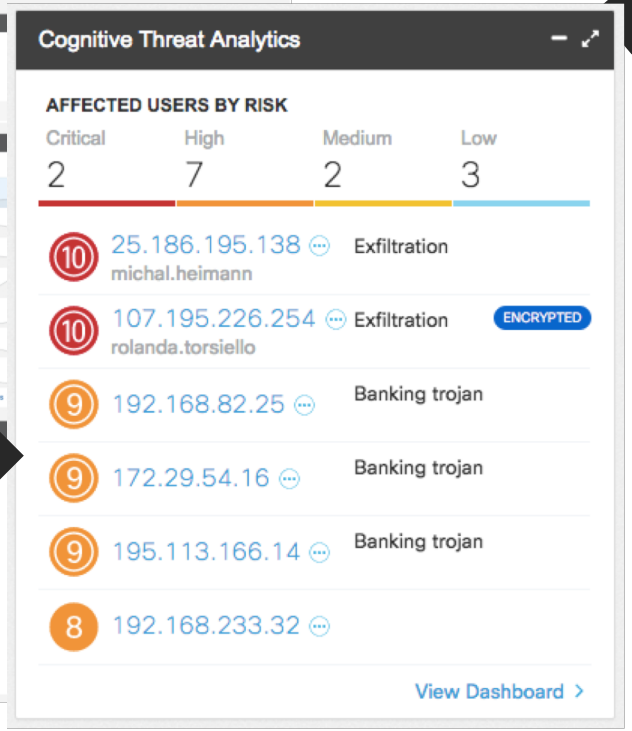


START	DURATION	CONNECTION APPLICATION	CONNECTION BYTES	ENCRYPTION TLS/SSL VERSION	ENCRYPTION KEY EXCHANGE	ENCRYPTION ALGORITHM AND KEY LENGTH	ENCRYPTION AUTHENTICATION ALGORITHM	ENCRYPTION MAC	PEER IP ADDRESS	PEER PORT/PROTOCOL	PEER HOST GROUPS	PEER BYTES
Apr 20, 2017 12:05:48 PM	2m 11s	HTTPS (unclassified)	132.61K	TLS 1.2	RSA	RSA_128	RSA	AES_128_CBC	10.0.40.10	443/TCP	Catch All	92.54K
Apr 20, 2017 11:58:48 AM	6m 11s	HTTPS (unclassified)	309.67K	TLS 1.2	RSA	RSA_128	RSA	AES_128_CBC	10.0.40.10	443/TCP	Catch All	216.14K
Apr 20, 2017 11:48:48 AM	9m 11s	HTTPS (unclassified)	444.16K	TLS 1.2	RSA	RSA_128	RSA	AES_128_CBC	10.0.40.10	443/TCP	Catch All	309.55K
Apr 20, 2017 11:34:48 AM	13m 11s	HTTPS (unclassified)	626.72K	TLS 1.2	RSA	RSA_128	RSA	AES_128_CBC	10.0.40.10	443/TCP	Catch All	437.98K
Apr 20, 2017 11:14:48 AM	19m 11s	HTTPS (unclassified)	871.41K	TLS 1.2	RSA	RSA_128	RSA	AES_128_CBC	10.0.40.10	443/TCP	Catch All	606.05K
Apr 20, 2017 10:46:48 AM	27m 11s	HTTPS (unclassified)	1.21M	TLS 1.2	RSA	RSA_128	RSA	AES_128_CBC	10.0.40.10	443/TCP	Catch All	861.54K
Apr 20, 2017 10:06:48 AM	39m 11s	HTTPS (unclassified)	1.73M	TLS 1.2	RSA	RSA_128	RSA	AES_128_CBC	10.0.40.10	443/TCP	Catch All	1.21M
Apr 20, 2017 9:10:48 AM	55m 11s	HTTPS (unclassified)	2.39M	TLS 1.2	RSA	RSA_128	RSA	AES_128_CBC	10.0.40.10	443/TCP	Catch All	1.67M
Apr 20, 2017 7:51:48 AM	1h 18m 11s	HTTPS (unclassified)	2.85M	TLS 1.2	RSA	RSA_128	RSA	AES_128_CBC	10.0.40.10	443/TCP	Catch All	1.98M
Apr 20, 2017 7:40:12 AM	10m 47s	HTTPS (unclassified)	503.88K	TLS 1.2	RSA	RSA_128	RSA	AES_128_CBC	10.0.40.10	443/TCP	Catch All	351.75K

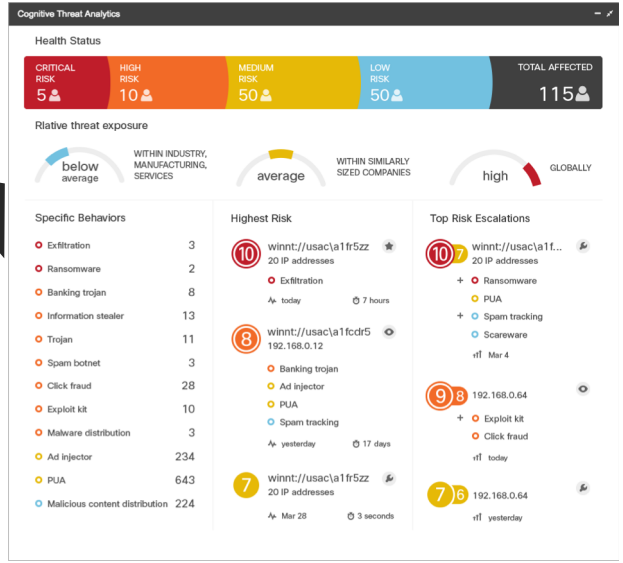
Encrypted Traffic Analytics



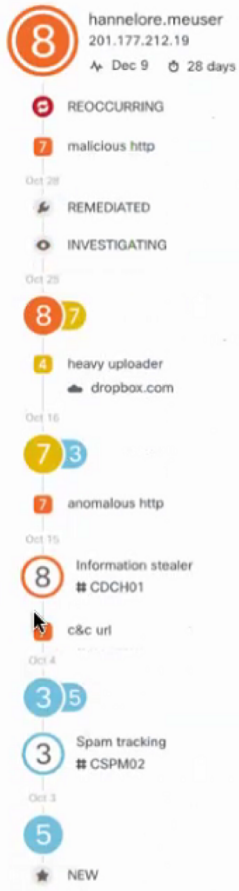
Cognitive Analytics



Expanded CTA Dashboard View



Cognitive Threat Analytics



FTP (uncla...	138B	< 0.01	138B		0B	
ICMP	992B	< 0.01	890B		102B	
Undefined...	549.07KB	< 0.01	548.67KB		408B	
Undefined...	277.39MB	< 0.01	235.45MB		41.94MB	
HTTP (uncl...	620B	< 0.01	144B		476B	
FTP (uncla...	138B	< 0.01	138B		0B	
ICMP	992B	< 0.01	890B		102B	
Undefined...	549.07KB	< 0.01	548.67KB		408B	
Undefined...	277.39MB	< 0.01	235.45MB		41.94MB	
HTTP (uncl...	620B	< 0.01	144B		476B	
FTP (uncla...	138B	< 0.01	138B		0B	
ICMP	992B	< 0.01	890B		102B	

10

MALWARE ENCRYPTED

100% confidence, in #CWNC01

NEW / TRIAGE

AFFECTING

dusti.hilton

10.201.3.51

OCCURRENCE

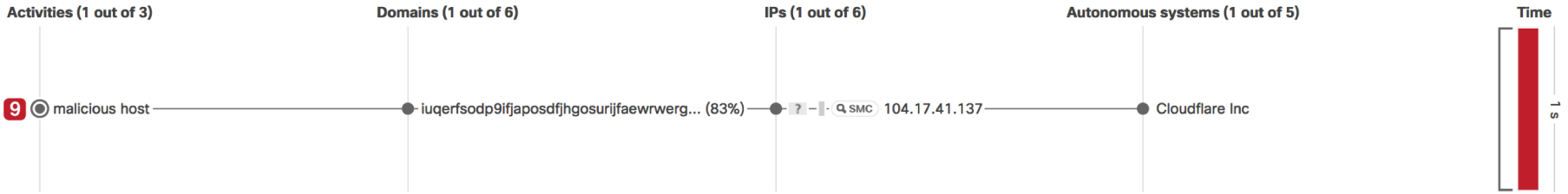
7 days

Apr 9 - Apr 16

Add notes...

ACTIVITIES AND FLOWS

SEVERITY FILTER: 9 8 7 6 5 4 3 2 1 Hide related

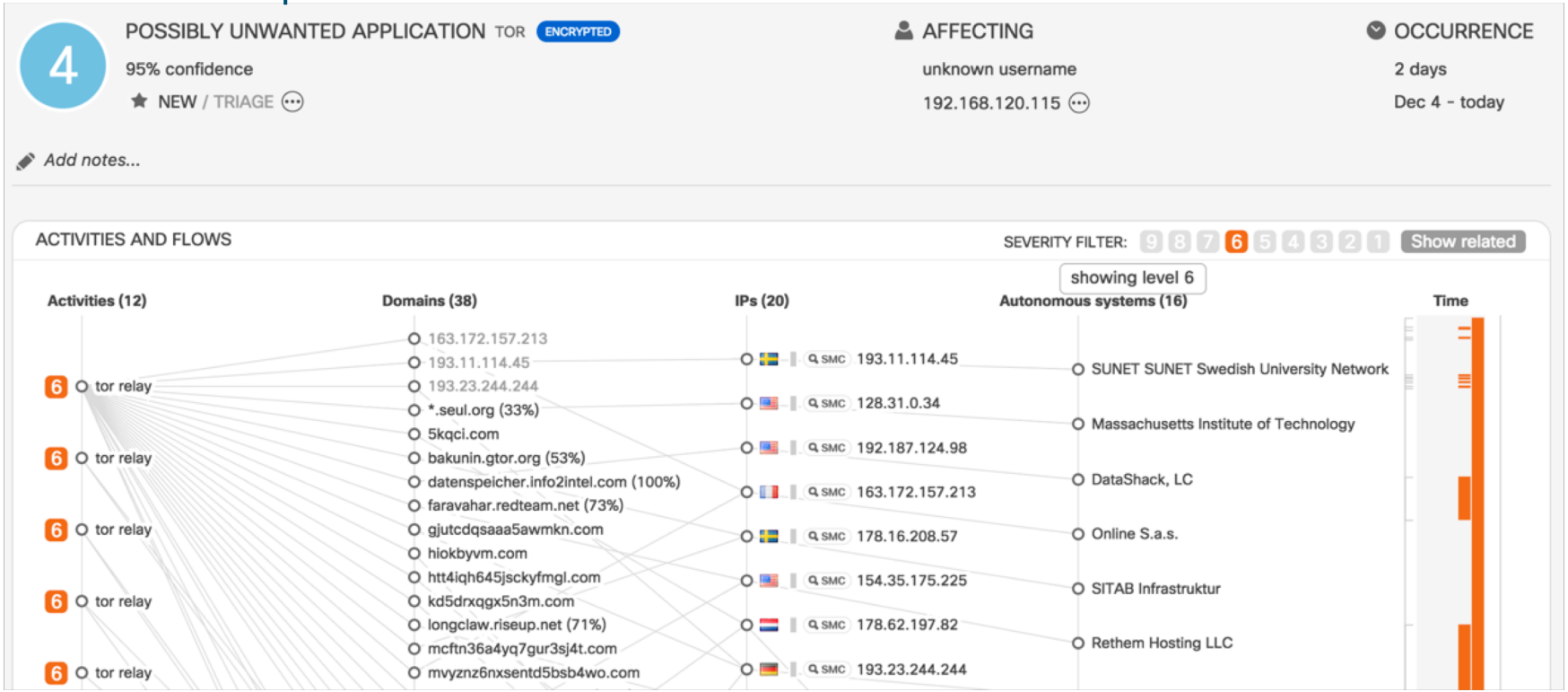


UPLOAD	DOWNLOAD	REQUESTS	DURATION	USER AGENTS	NO REFERRER	HTTP
1.6 KiB	1.2 KiB	1	1 second	0	100%	0

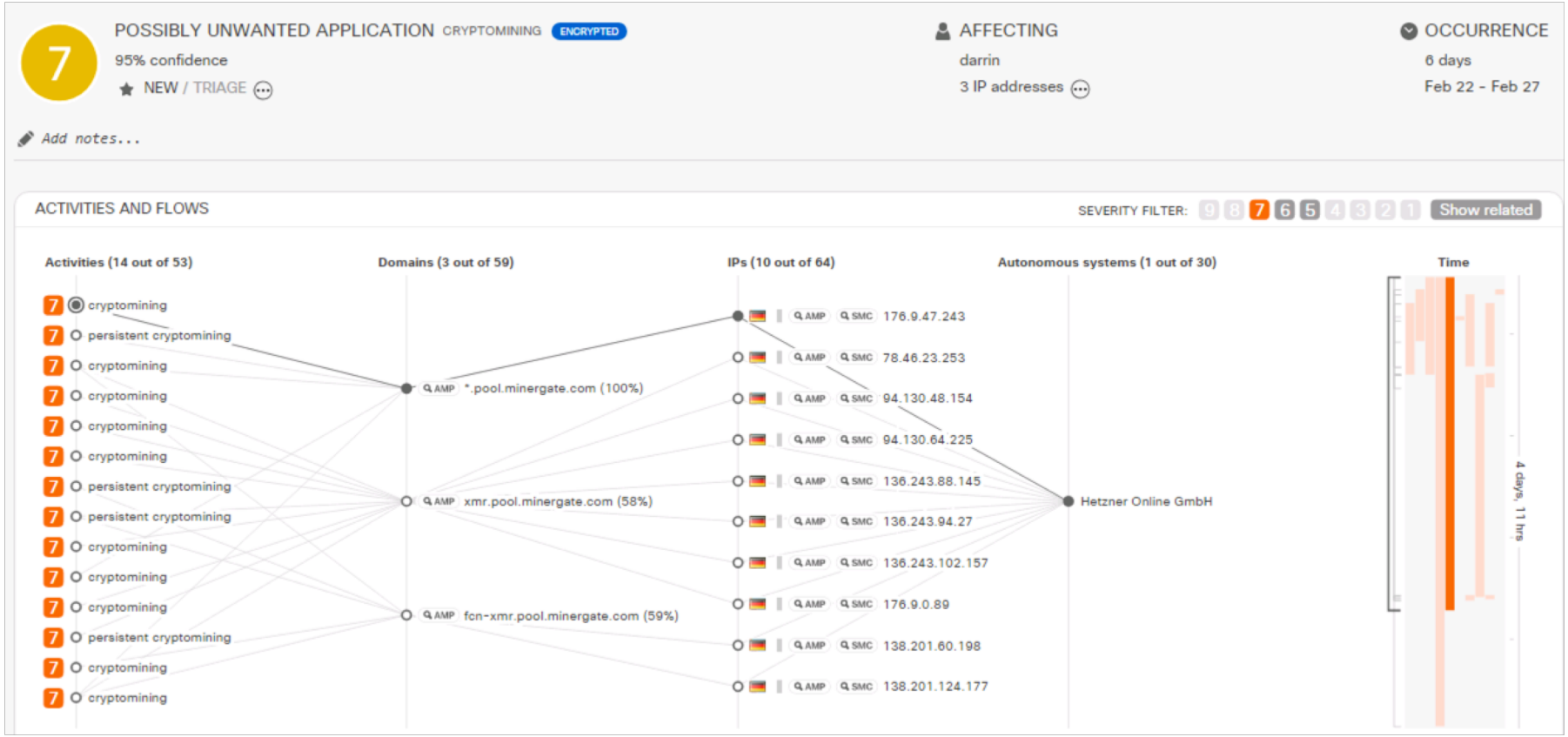
Client IP, Server IP, URL, SHA Filter

TYPE	TIMESTAMP	CLIENT	CLIENT F	CLIENT L	SERVER IP	SEI	NETWO	URL	DUR	PASSIVE DNS H	CLIEN	CLIEN	SERVI	SERVE
N	Apr 16, 2018 00:06:09 GMT+02:00	10.201.3.51	49476	49481	SMC 104.17.41.137	80	TCP		1 s	iuqerfsodp9ifjaposc	1628		712	5

TOR kapcsolatok

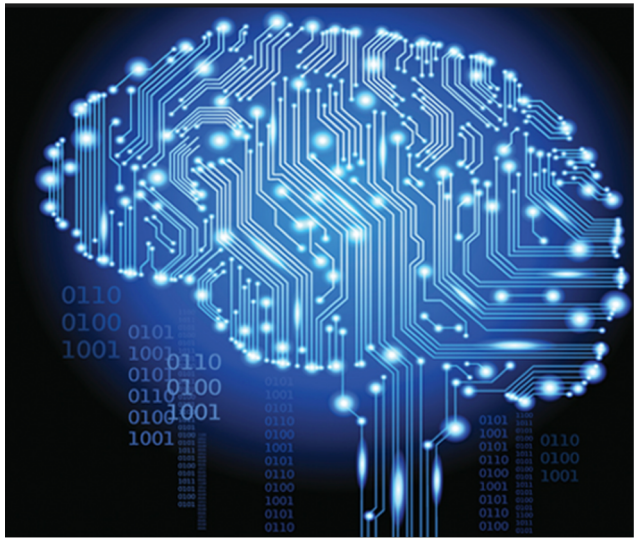


Cryptomining!



Az MI segíthet

- A kiberbűnözők nem fognak morális kérdést csinálni az MI-ből
- Csak ”emberi erőforrással” nem lehet lépést tartani a nagyszámú malware áradattal
- MI-val hatékonyabb felderítés, megelőzés lehetséges
- malware file felismerés, DNS forgalom,
- titkosított forgalomban,
- HTTP logok, NetFlow információk feldolgozása
- Új megközelítéseket is adhat (DeepMind go példa)



Köszönöm