DEXA 2018

“An Enhanced Cyber Attack Attribution Framework”

Dr. Nikolaos Pitropakis, Emmanouil Panaousis, Alkiviadis Giannakoulias, George Kalpakis, Rodrigo Diaz Rodriguez and Panayiotis Sarigiannidis

Presentation at TrustBus 2018
September 6, Regensburg

Contents

- ☑Advanced Persistent Threats
- ☑Other Cyber Attacks
- ☑Cyber Attack Attribution
- ☑NEON Framework

Advanced Persistent Threats

- **Advanced**
  - Because the adversary is conversant with computer intrusion tools and techniques and is capable of developing custom exploits
- **Persistent:**
  - Because the adversary intends to accomplish a mission
  - They receive directives and work towards specific goals
- **Threats**
  - Because the adversary is organized, funded and motivated.

- **APTs:**
  - Rely on multi-step attacks designed to infiltrate a system and remain there undetected for a long period of time to obtain high-value information
  - May spend a significant interval of time between different attack stages
  - Combine different attacks types, e.g., zero-day attacks (exploitation of unpatched vulnerabilities) and advanced social engineering attacks

APT Incidents

- **2009 - Stuxnet:**
  - targets SCADA systems and is believed to be responsible for causing substantial damage to Iran's nuclear programs
- **2011 - Duqu:**
  - Is a collection of computer malware discovered on 1 September 2011, thought to be related to the Stuxnet worm
  - looks for information that could be useful in attacking industrial control system
- **2012 - Flame**
  - is a modular computer malware discovered in 2012 that attacks computers running the Microsoft Windows operating system
  - was being used for targeted cyber espionage in Middle Eastern countries
- **2012 – Red October**
  - was reportedly operating worldwide for up to five years prior to discovery, transmitting information ranging from diplomatic secrets to personal information, including from mobile devices
  - the primary vectors used to install the malware were emails containing attached documents that exploited vulnerabilities in Microsoft Word and Excel

Other Cyber Attacks

- **2011 - today :**
  - Cyber attacks have been populated over the past few years
    - April 2011, Sony PlayStation suffers massive data breach, theft of names, addresses and possibly credit card data belonging to 77 million user accounts
    - 2017, the Shadow Brokers hacking group came up with a Windows platform exploit named as EternalBlue→ part of the WannaCry ransomware that affected numerous countries around the world and their critical infrastructures such as the UK's National Health System (NHS)
    - 2017 Shipping giant Maersk suffers 300 million dollars loss from Petya malware
    - July 2018, COSCO (China Ocean Shipping Company) US branch was attacked by a ransomware that resulted in the breakdown of telephone network, email servers, even the US website of the company went offline
    - 20-22 August 2018, Air Canada Suffers Data Breach - 20,000 Mobile App Users Affected

Recent cyber attacks

- **Possible solutions:**
  - Conventional incident detection and classification mechanisms but…
    - a new threat that of adversaries who aim to harm defending mechanisms that use machine learning introducing a new field of research called adversarial machine learning
    - as malicious parties become aware of the machine learning techniques used in defensive strategies they become elusive, lowering the accuracy rate of all detection capabilities
    - attackers continue to develop their new attacks based on previous → they do not reinvent the wheel → they recycle methodologies and infrastructure→ malware families and APT campaigns
  - attribution
    - the need to identify who (i.e., cyber attacker) is responsible for the orchestration of a cyber attack
  - cybersecurity situational awareness must be promoted
    - social engineering attacks take advantage of the human factor, which is referred as the weakest link

Cyber Attack Attribution

- **Attribution problem:**
  - refers to the difficulty of identifying those initially responsible for a cyber attack and their motivating factors, is a key in solidifying the threat representation
  - attribution of cyber attacks is not a straight-forward task

- DARPA splits the attribution process in three distinct phases which run in parallel:
  - **Activity Tracking and Summarization**
    - collection of information from multiple sources
  - **Data Fusion and Activity Prediction**
    - data associations are being captured across diverse data sets
  - **Validation & Enrichment**
    - adversary mistakes are being identified
    - use of analytic techniques to expose known but hidden structures

Possible solution?

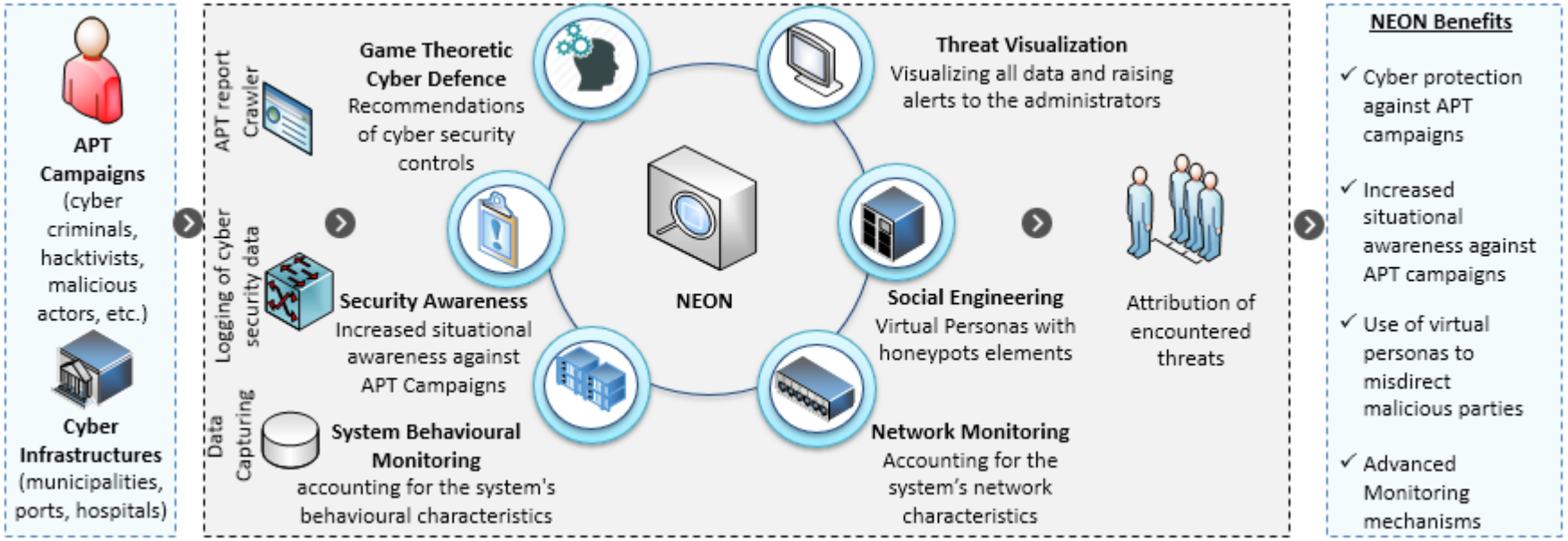
- there is no concrete methodology that attributes each attack to the malicious parties who launched it
- no methodology takes into consideration past knowledge of APT campaigns and both network and system behavioural data

NEON Framework

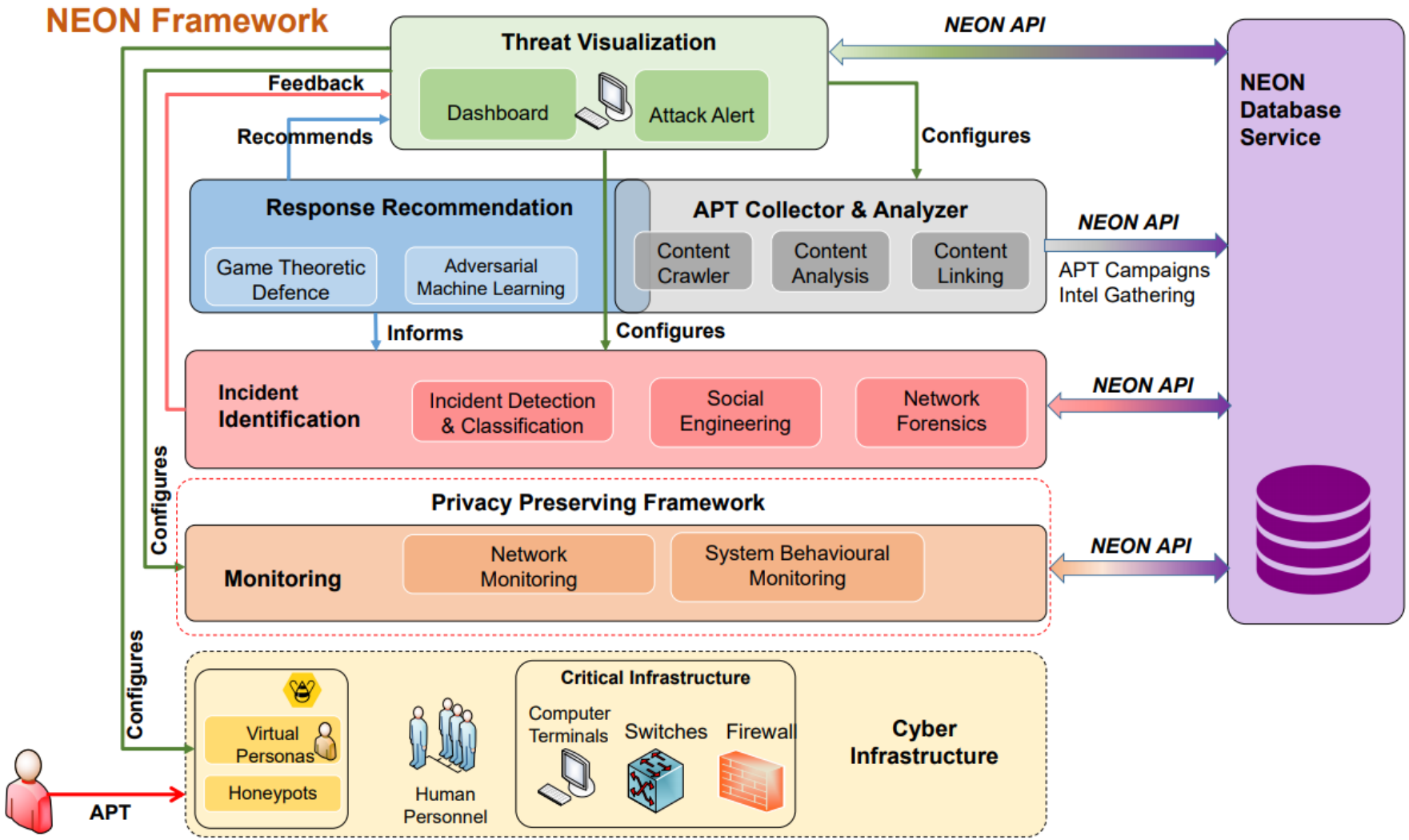
- E**n**hanced Cyb**e**r Attack Attributi**on** (**NEON**) Framework:
  - is designed to accommodate components that address the aforesaid challenges
  - leads to a user-centric automated cybersecurity platform that gathers heterogeneous data coming from APT reports and publicly available information from social media
  - using this material as ground truth, NEON correlates this with other data collected from network and system behavioural monitoring components
  - in order to increase defence against social engineering attacks, uses honeypots that attract the attention of potential attackers through the creation and management of virtual personas → accelerate the manifestation of the attacks in contained environments, drawing at the same time valuable information about the adversaries
  - uses a game theoretic approach to propose optimal cybersecurity actions against adversaries

- *To the best of our knowledge, NEON is the first framework that has been designed with the ultimate goal to perform enhanced attribution of APT campaigns*

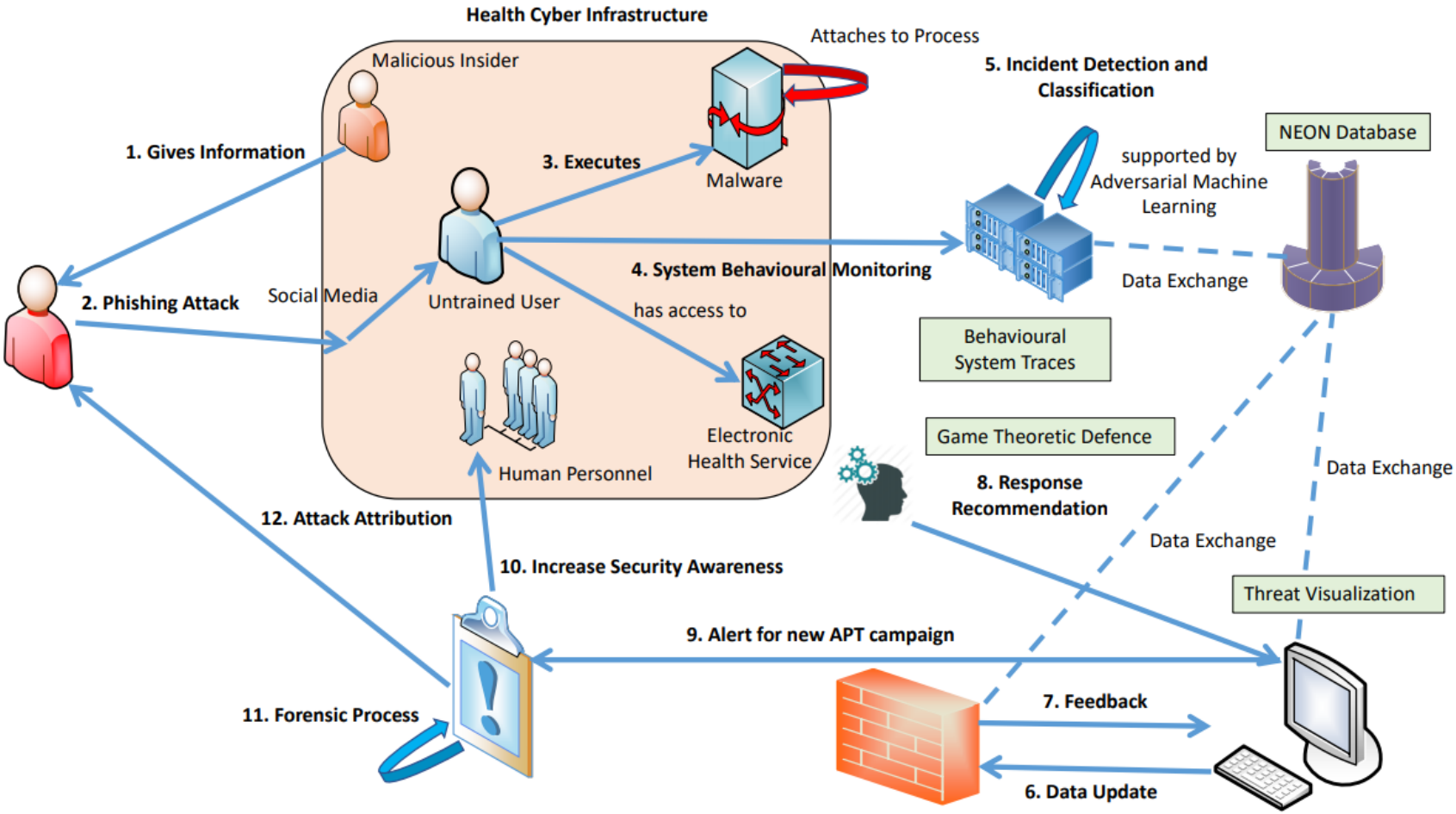
NEON Framework



NEON Architecture



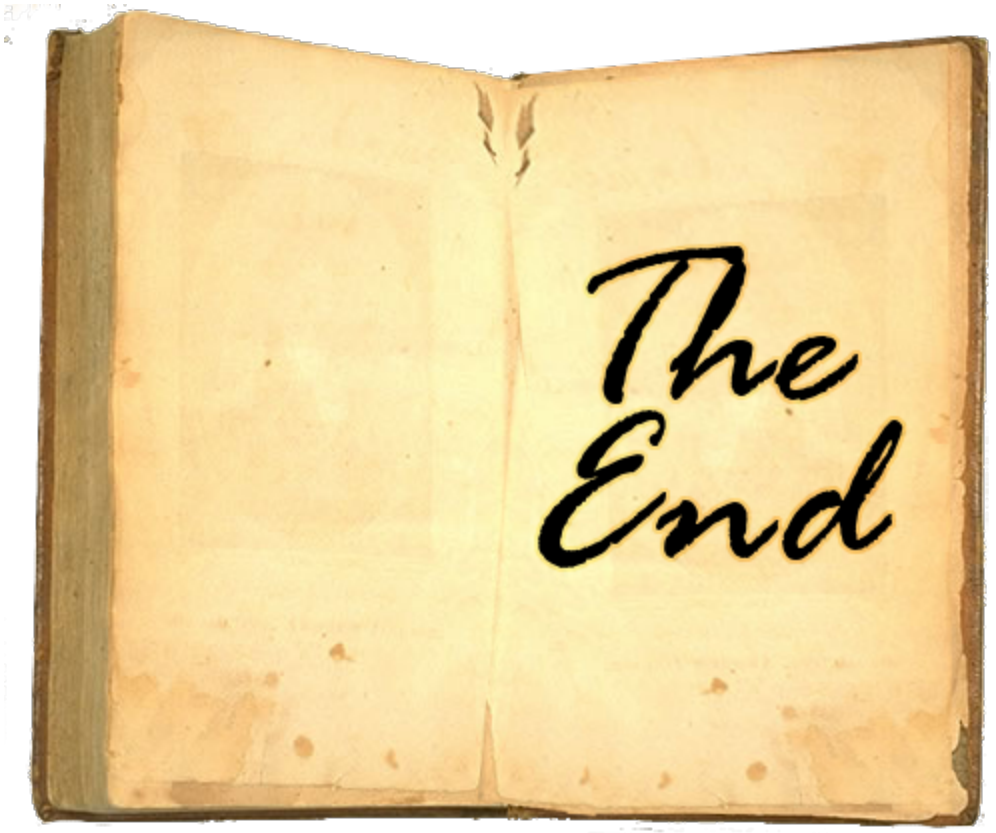
Healthcare Use case of NEON



Conclusions and future work

- Enhanced attack attribution frameworks are in their infancy
- APT becomes the most prominent threat paradigm
- To address challenges that emerge from the above, we propose the NEON framework:
  - Its primary target is the collection and representation of intelligence about APT campaigns and then the correlation with monitoring activities
  - Honeypots with the help of virtual personas improve the detection capabilities of zero-day exploits and social engineering attacks
  - Game theoretic defences are incorporated into NEON to mitigate the actions of sophisticated APT attackers
  - adversarial machine learning supports data trustworthiness thus facilitating accurate APT detection and attribution
  - threat management console visualizes and pronounces the situational awareness of people and critical infrastructures in NEON

- Our plan is to develop the individual NEON components in the following order: (i) APT Collector & Analyzer, (ii) Monitoring, (iii) Incident Identification, (iii) Response Recommendation, and (iv) Threat Visualization.


The
End