MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

IN COLLABORATION WITH MIT MANAGEMENT SLOAN SCHOOL

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

MASTER MEIM 2022-2023

Artificial Intelligence and Machine Learning for Cybersecurity

SELECTED TOPICS

Luigi CATUOGNO, PhD

Associate Professor in Computer Science at Università degli Studi di Napoli Parthenope (ITALY)

1

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Summary

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Goals:

1) Why Artificial Intelligence and Machine Learning technologies turned out to be essential for cybersecurity
2) AI, ML and Automation
3) Sample applications:
   - *Intrusion Detection Systems*
   - *Ransomware Detection*
4) Conclusion

2

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Why AI and ML are essential for cybersecurity

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Cyber Attacks bring high costs for business.

It has been estimated that between 2015 and 2021, the costs of cybercrime grew by 3 trillion to over 5 trillion dollars

Including: damage and destruction of data, theft of personal/financial data, intellectual property, lost productivity, costs for restoration and investigations…

In the last four years, the global spending on cybersecurity products and services for defending against cyber-criminals has overtaken one trillion dollars

3

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Why AI and ML are essential for cybersecurity

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

For a long time, organizations have relied upon *stand-alone* security solutions, based on static and inflexible defense techniques such as *antivirus* and *firewalls*

Recently, cyber-defenses leveraging Artificial Intelligence (AI) and Machine Learning (ML) has gained center stage

Both tools have been increasingly considered as essential, due to the rapid increase in number and complexity of attacks.

In facts, these technologies have been the base upon many successful security tools have been built upon for years…

4

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Why AI and ML are essential for cybersecurity

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Traditional antivirus software is *rule-based*

- It denies any *process* to access data and computing resources whenever recognized *signature patterns* are encountered
- For example: if a *known* malware infects the computer, the antivirus installed on the system recognizes its *signature* (generally a file hash) and stops it from executing

5

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Why AI and ML are essential for cybersecurity

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Problems:

- Signature-based antiviruses *only detect known malware*
- Antiviruses do not close off the infection breach (e.g., in the web browser, a web service…) so that the vulnerability can be exploited again.

  The risk is that if a the "next time" an unknow malware comes to the system through the same point, the antivirus will not be able to detect (and defeat) it.

6

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Why AI and ML are essential for cybersecurity

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Moreover, the antivirus does not account for legitimate programs being used in malicious ways. This is the case of so called *fileless malware*

- The attacker might force legacy applications and tools (e.g., web browsers, scripting shells) to execute malicious actions.
- Due to such tools are known as "good programs", the antivirus allow them to operate, thought they have been compromised.

7

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Why AI and ML are essential for cybersecurity

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Many antivirus vendors have enriched their products capabilities with some *heuristic detection algorithms.*

Briefly, the antivirus also monitors the system operation for behaviors that could be linked to the presence and execution of malicious code

- writing/overwriting certain Windows Registry entries
- modifying certain permission on macOS devices…

In case something potentially harmful happens, the antivirus stops the process regardless it has/has not a signature related to malicious files.

However, as malware grows in complexity and sophistication, this approach become less effective…

8

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Why AI and ML are essential for cybersecurity

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Many antivirus vendors have enriched their products capabilities with some *heuristic detection algorithms*.

Briefly, the antivirus also monitors the system operation for behaviors that could be linked to the presence and execution of malicious code

- writing/overwriting certain Windows Registry entries
- modifying certain permission on macOS devices…

In case something potentially harmful happens, the antivirus stops the process regardless it has/has not a signature related to malicious files.

However, as malware grows in complexity and sophistication, this approach become less effective…

…not to metion false positives!

9

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Why AI and ML are essential for cybersecurity

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Firewalls either block or allow network traffic through certain *ports* regardless its content and nature.

e.g., nowadays almost every single networked equipment denies inbound telnet connection (TCP port 23) as it has been, in past, massively used as intrusion vector while rarely used for legitimate purposes.

10

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Why AI and ML are essential for cybersecurity

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Nevertheless, blocking certain ports is not always a viable option.

- For example, inbound traffic to ports 80(HTTP) and 443(HTTPS) must be kept opened by service providers whose users access the web services and application through their web browser.
- So that, adversaries have moved to attack their victims through such ports, being able to pass through the firewall.

11

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Why AI and ML are essential for cybersecurity

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Social engineering has turned out to be a major threat today.

Whenever attackers can't get through the firewall from the outside-in they manage to trick the insiders into doing the work for them….

12

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Why AI and ML are essential for cybersecurity

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Through a phishing email, the employee is directed to a malicious site.

The *return traffic* from such a site is allowed through the firewall. It is just the way firewalls work

Most often, the *return traffic* is an exploit for a known vulnerability and carries some additional malicious code that, once executed by the victim, will open up a backdoor on the system.

In this case, firewalls provide little, if any protection.

13

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Why AI and ML are essential for cybersecurity

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Every networked system is, in one way or another, potentially vulnerable to attacks. Nowadays, traditional protection measures such as antivirus and firewalls do no work:

- their behavior results largely *predictable* to adversaries
- they work only when they have a signature for a certain exploit. So that, whenever a new exploit appears…

14

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Why AI and ML are essential for cybersecurity

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

In contrast, solutions *leveraging AI and ML technologies*, allow to identify attacks looking beyond simple signatures:

- identifying similarities to what has happened before
- flagging things that appear to be anomalies.

15

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

AI, ML and Automation

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

16

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

AI, ML and Automation

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Although terms AI and ML are often used interchangeably, they are not the same thing…

17

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

AI, ML and Automation

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Artificial Intelligence (AI) is defined as

"the theory and development of computer systems that are able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making…"

18

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

AI, ML and Automation

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Artificial Intelligence (AI) is defined as

"the theory and development of computer systems that are able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making…"

With AI, machines simulate intelligent behaviors (including learning and problem solving), in contrast to the natural intelligence displayed by humans.

19

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

AI, ML and Automation

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Machine Learning (ML) is

"an application of AI that provide systems with the ability to automatically learn and improve form experience without being explicitly programmed."

20

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

AI, ML and Automation

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Machine Learning (ML) is

"*an application of AI that provide systems with the ability to automatically learn and improve form experience without being explicitly programmed.*"

ML focuses on the development of computer programs that *aggregate* data into a *model* and use it to learn for themselves

21

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

AI, ML and Automation

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

With ML, humans—generally analysts in the case of security– are responsible for training the machine, and the machine is capable of learning with the help of humans as feedback systems.

The more machines are trained, the *smarter* they become, as long as the training material is valuable for the tasks that the machines are supposed to focus on.

In the current defense landscape, ML is more established and, therefore, more likely to be used defensively as compared to AI.

22

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

An example: monitoring for credit card frauds

ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND AUTOMATION

Credit Card Companies monitor billions of transactions per day looking for potential fraudulent transactions;

The algorithms take in account millions of *factors* related to each transaction (*e.g., timing, circumstances, ….*)

23

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

An example: monitoring for credit card frauds

ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND AUTOMATION

In particular, some factors are obvious:

- **Location**: *e.g., if a CC is sweeped in Naples, it can't be used five minutes later in Hong Kong;*
- **Time**: *e.g., a CC can't be used contemporarily in two different transactions*

24

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

An example: monitoring for credit card frauds

ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND AUTOMATION

Other factors are less obvious:

- Considering the case that a Credit Card, that is regularly used to buy clothes at a retailer such as H&M or Kik, is suddenly used to buy clothes at Gucci.

This might raise an alert. But it is not immediately clear weather that is fraudulent activity or just someone buying clothes for a special occasion

25

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

An example: monitoring for credit card frauds

ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND AUTOMATION

Humans can not look at all the different ways the fraudulent transactions manifest themselves, so the algorithms must consider any anomalous transaction, telling apart potential frauds

The AI is part of the decision-making process.

26

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

An example: monitoring for credit card frauds

ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND AUTOMATION

Humans can not look at all the different ways the fraudulent transactions manifest themselves, so the algorithms must consider any anomalous transaction, telling apart potential frauds.

The AI is part of the decision-making process.

The ML task is "surfing" those billions of transaction each day, discovering new patterns that indicate fraud and adjusting the AI algorithms to account for the new information.

The job of ML is to bring out anomalies.

27

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Differences between AI and ML

ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND AUTOMATION

ML and AI do not always work together. Some systems take advantage of one technology or the other but not both.

- *Security information and event managers (SIEMs) use ML to search through millions of log events to build alerts, but the security operation center (SOC) analyst sees only the alerts.*
- *Social networks use AI for several tasks, for example to help to identify and tag people in pictures. The technology is invisible to the user, they just know that when they upload a picture, their friends are automatically tagged in it.*

28

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Automation

ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND AUTOMATION

Automation can be seen as

"the technique, method or system of operating or controlling a process, reducing human intervention to a minimum"

29

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Automation

ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND AUTOMATION

Automation can be seen as

"the technique, method or system of operating or controlling a process, reducing human intervention to a minimum"

Automation is just manual rules and processes repeated automatically but nothing is learned, as in the case with ML and AI.

30

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Automation

ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND AUTOMATION

Automation is often the result of AI and ML systems within an organization.

- For example, *an organization might use AI and ML to identify suspicious activity and then uses automation to raise alerts on the activity or even take action to stop it.*

Automation is what humans see of AI and ML systems.

Automation driven by AI and ML backend systems is one of the biggest growth areas in cybersecurity.

31

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Challenges in Adopting AI and ML

ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND AUTOMATION

Ensuring to be able to collect data (datasets for training and testing systems) to feed into AI and ML systems. Data might not be made available by vendors/stakeholders/parthners for several reasons *(e.g., for privacy)*.

Data interoperability. Being able to handle data in in whatever format it is presented and to understand its structure so that it can be parsed and correlated against other data types being ingested by the AI and ML system.

32

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Challenges in Adopting AI and ML

ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND AUTOMATION

AI and ML systems require a lot of maintenance at least initially.

It is vital to feed these systems with the right data and to continuously verify that the system produces its output according the expectations.

Managing the training process of an AI and ML system in order to make it to better understand the kind of results the analysts are looking for, is a crucial task to be accomplished *continuously and carefully*

33

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Intrusion Detection

A CASE STUDY

34

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Intrusion Detection

We can define an **intrusion** as

"a set of actions aimed at compromising the security goals: integrity, confidentiality or availability of of any data asset, computing resource, networking resource"

Intrusion Detection is the

"complex of processes aiming at identifying and responding to intrusions"

35

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Intrusion Detection System (IDS)

An **Intrusion Detection System (IDS)** is

"a set of tools, and techniques which identifies –preferably in real time – any occurring intrusion (i.e., unauthorized use, misuse and abuse of a computing resource.)"

It is a *reactive* (rather than *proactive*) measure for system security

36

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Intrusion Detection System (IDS)

IDSes can be classified according different criteria:

With respect to the way the intrusions are modeled and recognized

- **Misuse Detection** (a.k.a., signature-based ID): seeks for known ad well-defined intrusion patterns.
- **Anomaly Detection** builds up a model of the expected system operation profile, then monitors and detects any significant deviation form such profile.

37

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Intrusion Detection System (IDS)

IDSes can be classified according different criteria:

With respect to the *deployment*

- **Host-based (HIDS):** Detects attacks against a single host.
- **Distributed IDS (DIDS):** Gathers audit data from multiple hosts and possibly the network that connects them. Detects attacks involving multiple hosts.
- **Network-based (NIDS):** Uses network traffic as the audit data source, retrieving information from the networking hardware rather than the hosts. Detects attacks from network.

38

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Intrusion Detection System (IDS)

FUNDAMENTAL ASSUMPTIONS

IDSes are based upon two fundamental assumptions:

I. *Every event happening on the system is observable*

II. *Normal and intrusive activities have distinct evidence*

39

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Intrusion Detection System (IDS)

CONCEPTS

Source data:

- *audit files*: system and application log files
- data from resources allocation/usage accounting

Such data are massive and disaggregated and therefore they are used to build derived values (features) intended to be informative and non-redundant.

40

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Intrusion Detection System (IDS)

CONCEPTS

Feature:

- individual measurable property or characteristic of any event.
- capture evidences of occurring events (both normal and intrusive)
- The number and type of features depends on the treatment method and the *model*

41

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Intrusion Detection System (IDS)

CONCEPTS

Model:

- Tabular, logical, mathematical tool which correlate evidences to categorize the events.

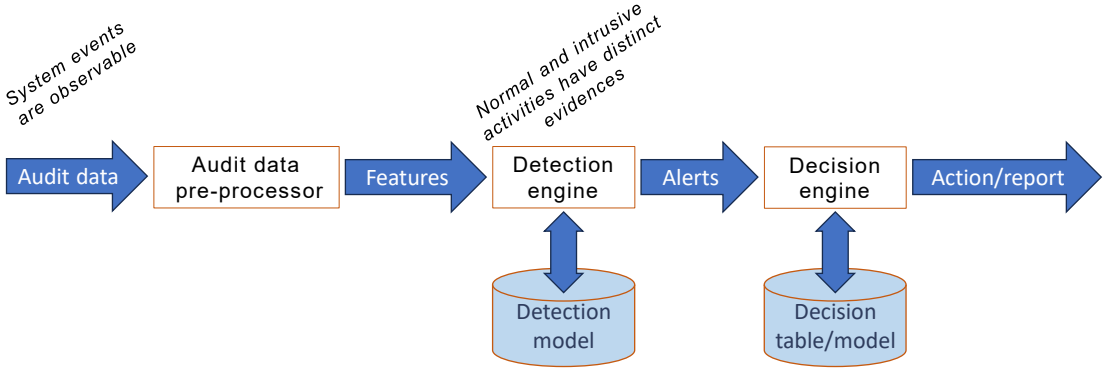
42

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Intrusion Detection System (IDS)

ARCHITECTURE AND OPERATION



43

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Misuse Intrusion Detection

INTRUSION DETECTION SYSTEM (IDS))

- Addresses a pre-defined set of known attack vectors.
- Features extracted from known intrusions.
- Attacks or weak spots of the system are described through static rules and patterns.
- Uses *pattern matching* algorithms to match and identify intrusions.

44

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Misuse Intrusion Detection

INTRUSION DETECTION SYSTEM (IDS))

- Addresses a pre-defined set of known attack vectors.
- Features extracted from known intrusions.
- Attacks or weak spots of the system are described through static rules and patterns.
- Uses *pattern matching* algorithms to match and identify intrusions.

But...

- **Cannot detect novel or unknown attacks.**
- **The *knowledge base* must be continuously updated**

45

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Anomaly Detection

INTRUSION DETECTION SYSTEM (IDS)

It is based on the *knowledge* of the normal behavior of a subject, that is learned through the analysis of audit data collected over a period of normal operation.

- Normally, training audit data does not include intrusion data.

Any action that *significantly deviates* from the normal behavior is *considered* an intrusion.

46

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

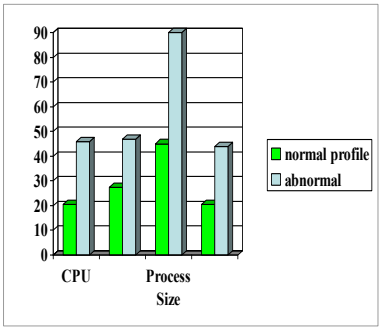
UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Anomaly Detection

INTRUSION DETECTION SYSTEM (IDS)

Possible features to take in account:

- *loginfrequency,*
- *locationfrequency,*
- *UseofCPU,*
- *UseofIO,*
- *ExecutionFrequency*
- *FileReadFails*
- *FileWriteFails*



47

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Anomaly Detection

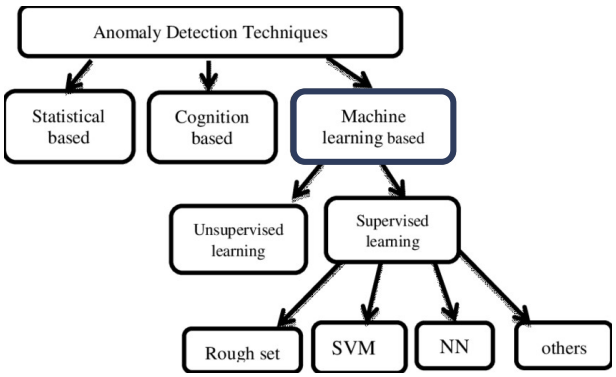
INTRUSION DETECTION SYSTEM (IDS)

Deviations can be observed and measured by means different techniques including:

Statistical analysis

(Mean and standard deviation, Markov process, Time Series…)

➡ Machine Learning



48

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Anomaly Detection issues

INTRUSION DETECTION SYSTEM (IDS)

The learning process is based on audit data collected over a period of normal operation, so:

- Knowledge always needs to be updated through periodical training sessions with *fresh* and *clean* data
- Noise data (from an intrusion) in the training datasets will make a mis-classification.

50

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Anomaly Detection issues

INTRUSION DETECTION SYSTEM (IDS)

How to decide the features to be considered?

- Features are usually decided by domain experts. Choices might be not enough effective…

Performance

- Relatively high false positive rate

51

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Ransomware detection

A CASE STUDY

52

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Ransomware at a glance

A *malware* which locks (using cryptographic algorithms) the data of its victims and demands the payment of a *ransom* to give back the access to such data.

Nowadays *ransomware* increasingly threaten «traditional» malware prevention systems such *as signature-based detection systems (e.g* off-the-shelf antiviruses) and *anomaly detection systems:*

53

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Ransomware at a glance

A *malware* which locks (using cryptographic algorithms) the data of its victims and demands the payment of a *ransom* to give back the access to such data.

Nowadays *ransomware* increasingly threaten «traditional» malware prevention systems such *as signature-based detection systems (e.g* off-the-shelf antiviruses) and *anomaly detection systems:*

- New variants appear very quickly

54

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Ransomware at a glance

A *malware* which locks (using cryptographic algorithms) the data of its victims and demands the payment of a *ransom* to give back the access to such data.

Nowadays *ransomware* increasingly threaten «traditional» malware prevention systems such *as signature-based detection systems (e.g* off-the-shelf antiviruses) and *anomaly detection systems:*

- New variants appear very quickly
- The number of *polimorphic* and *randomized* viruses constantly grows

55

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Ransomware at a glance

A *malware* which locks (using cryptographic algorithms) the data of its victims and demands the payment of a *ransom* to give back the access to such data.

Nowadays *ransomware* increasingly threaten «traditional» malware prevention systems such *as signature-based detection systems* (e.g off-the-shelf antiviruses) and *anomaly detection systems:*

- New variants appear very quickly
- The number of *polimorphic* and *randomized* viruses constantly grows
- Ransomware *mimic* legitimate processes, so that their detection by ADSes gets slower (and harder)

56

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Ransomware is different from “generic” malware

RANSOMWARE AT A GLANCE

- «Generic» Malware present a wide variety of behaviors and objectives. These objectives may require a certain time to be achieved. In the meantime, it may remain dormant…

57

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Ransomware is different from “generic” malware

RANSOMWARE AT A GLANCE

«Generic» Malware present a wide variety of behaviors and objectives. These objectives may require a certain time to be achieved. In the meantime, it may remain dormant…

Ransomware pursue a sole plan: encrypting as much data as possible in the shortest possible time. Ransomware are immediately dangerous since the process starts.

58

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Ransomware is different from “generic” malware

RANSOMWARE AT A GLANCE

«Generic» Malware present a wide variety of behaviors and objectives. These objectives may require a certain time to be achieved. In the meantime, it may remain dormant…

Ransomware pursue a sole plan: encrypting as much data as possible in the shortest possible time. Ransomware are immediately dangerous since the process starts.

ADSes need to gather information about suspect processes for a while, before making a decision. For generic malware, this might happen *in time* to avoid damages

59

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Ransomware is different from "generic" malware

RANSOMWARE AT A GLANCE

«Generic» Malware present a wide variety of behaviors and objectives. These objectives may require a certain time to be achieved. In the meantime, it may remain dormant…

Ransomware pursue a sole plan: encrypting as much data as possible in the shortest possible time. Ransomware are immediately dangerous since the process starts.

ADSes need to gather information about suspect processes for a while, before making a decision. For generic malware, this might happen *in time* to avoid damages

For ransomware, this always happens too late, as every single file it accessed in the meantime, is gone.

60

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Ransomware is different from "generic" malware

RANSOMWARE AT A GLANCE

Ransomware can be thwarted adopting *ad hoc* strategies that take advantage from two specific characteristics:

At very end, ransomware do nothing more than encrypting files and presenting the bill.

This is done by following few strategies that are quite similar each others

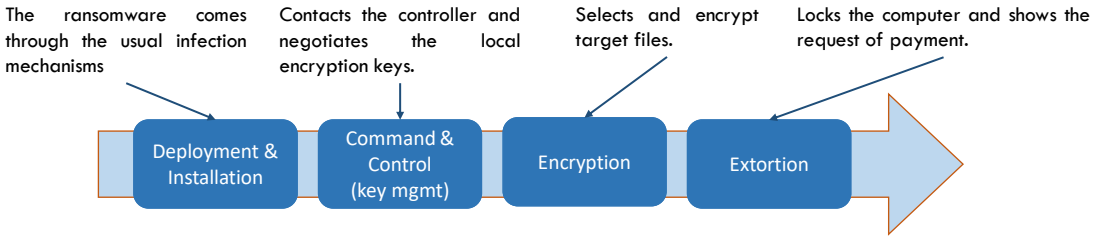
61

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

What Ransomware do…

RANSOMWARE AT A GLANCE



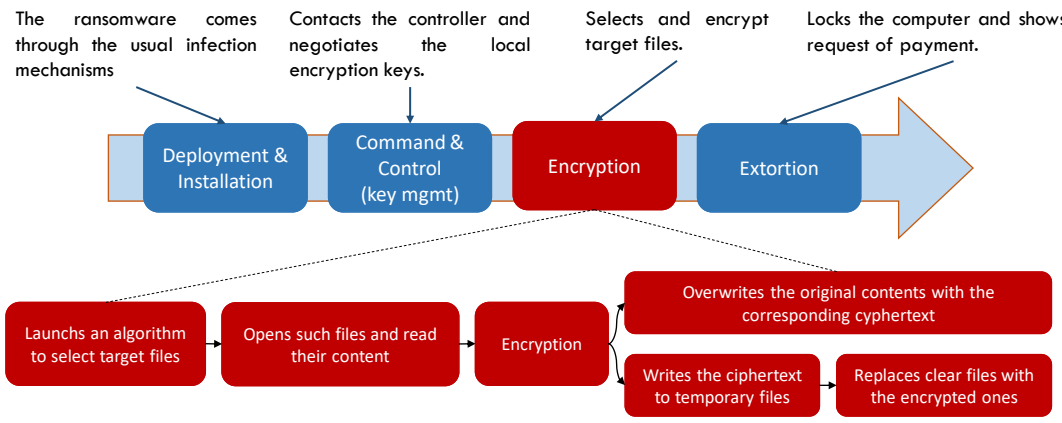
62

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

What Ransomware do… and how

RANSOMWARE AT A GLANCE



63

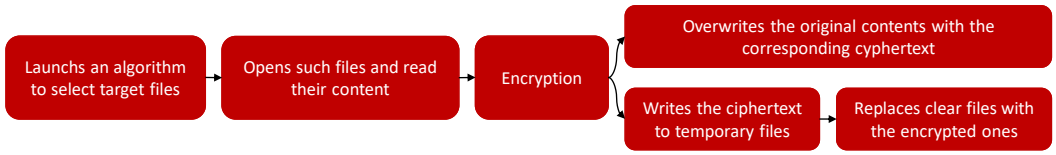
MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

What Ransomware do… and how

RANSOMWARE AT A GLANCE

Ransomware can not avoid to stress the use of certain specific file system functionalities and to leave some characteristic «tracks» on the affected files.



Extensive file system traversals and directory listing;
Massive files attributes reading;

Massive files attributes reading;
Highly frequent read and write calls on overlapping file segments (overwrites)

Frequent file type changes;
Significant changes of content-related file properties (e.g. *entropy*)

Frequent file deletion, moving and renaming;

64

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Tracking Ransomware on the file system

RANSOMWARE DETECTION

65

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Tracking Ransomware on the file system

A recent approach to ransomware detection and mitigation consists in seeking evidence of ransomware activity by tracking the victim's file system operation.

- The *monitor* measures the amount and frequency of the invocation of «sensible» file system calls by each process (along with other metrics related to the files content)
- Whenever values that might be likely related to malicious activities are found, the *decision maker* identifies the suspect process and carries out the appropriate actions.

66

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Tracking Ransomware on the file system

Differencies with the «*traditional*» ADS approach:

- The system looks for those processes whose activity converges to a limited set of expected «malicious behaviors» instead of analyzing every deviation from the presumptive «normal» behavioral model.
- The detection process focuses on the well-known effects of the malware rather than identifying its executable and its running process.

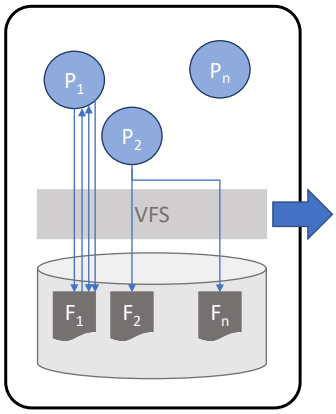
67

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Gathering information…

TRACKING RANSOMWARE ON THE FILE SYSTEM



Detection

Taking *per process* measurements wrt several aspects of the file system activity

$v_{i,j}$ *denotes the value of* M_i *withh respect to process* P_j

	P_1	P_2	…	P_n
M_1	$v_{1,1}$	$v_{1,2}$	…	$v_{1,n}$
M_2	$v_{2,1}$	$v_{2,2}$	…	…
M_3	$v_{3,1}$	…	…	…
…				

Where measurements M_i may include:

- the count or frequency of a certain system call invocations;
- quantitative values related to certain system calls parameters;
- quantitative value related to the effect such calls has on the file (e.g entropy variations, attributes changes…)

68

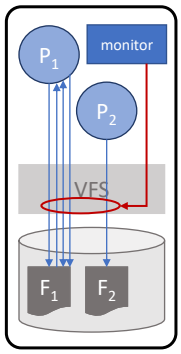
MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

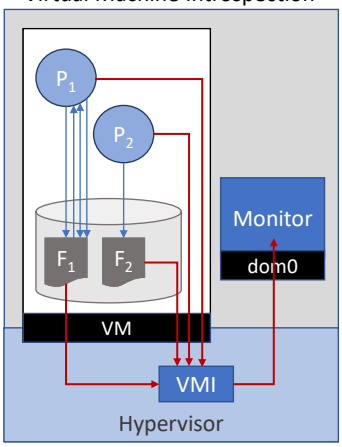
Gathering information…

TRACKING RANSOMWARE ON THE FILE SYSTEM

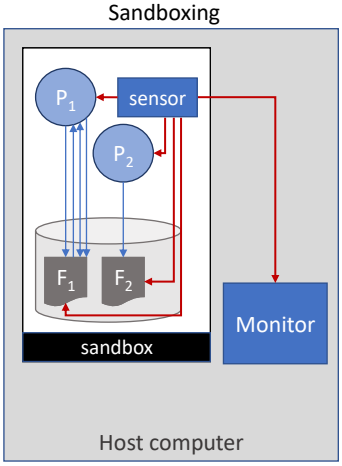
File System API hooking



Virtual Machine Introspection



Sandboxing



69

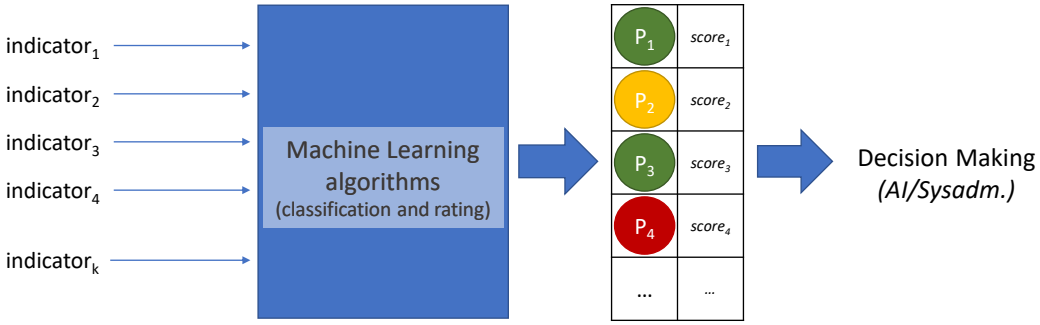
MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Malice indicators

TRACKING RANSOMWARE ON THE FILE SYSTEM

Measurements are aggregated into malice indicators whose value may reveal that a given process is potentially performing malicious activity



70

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Malice indicators

TRACKING RANSOMWARE ON THE FILE SYSTEM

Statistics about the file system API functions invocation

Some examples:

read/write and «overwrites» stats	**file types modification rate**	**funneling**
frequency of dir-based calls	**frequency of files move /removal**	**file access stats**

71

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Malice indicators

TRACKING RANSOMWARE ON THE FILE SYSTEM

Statistics about the file system API functions invocation

Statistics about how suspected processes change the files content

For example:

Divergence measures in overwritten buffers
Significant variations in the statistical distribution of the symbols contained in a file, that might reveal an in-progress encryption.

72

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Malice indicators: issues

TRACKING RANSOMWARE ON THE FILE SYSTEM

False positives	Some applications behave similarly to ransomware (if considering only certain indicators). E.g. zip archivers may trigger several indicators: FS traversal, funneling, etc. In general, applications and libraries which handle compressed file feature noticeable entropy variations between read and written data, ….
Indicator evasion	Ransomware might modify its strategy in order to not trigger certain indicators. For example, frequency based indicators can be circumvented by slowing down the pace of file accesses.
Multi-process ransomware	The encryption process might be splitten among several processes, each having in charge only one phase. So that single sub-processes are less likely to trigger any indicator (on time).

73

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Malice indicators: issues

TRACKING RANSOMWARE ON THE FILE SYSTEM

False positives
Indicator evasion

The malice score of each process can be computed taking in account multiple indicators at the same time. Indeed, benign processes rarely trigger multiple malice indicators as well as «slow» ransomware may deceive some indicators but could not operate without triggering others.

Multi-process ransomware

The system should be able to identify groups of processes (somewhat related) that collectively raise a set of malicious indicators potentially related to a malicious activity.

74

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Building training datasets

TRACKING RANSOMWARE ON THE FILE SYSTEM

Detectors training and performance evaluation require datasets composed of both ransomware samples and system logs from an infected system.

Traditionally, analysts build their dataset by placing online some honeypots in order to capture ransomware «in the wild».

Any sample is then installed and run into a *sandbox* system so that its activity is extracted from the system operation logs/stats and tagged as *malicious*.

75

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Building training datasets: issues

TRACKING RANSOMWARE ON THE FILE SYSTEM

Ransomware attack strategies are significantly driven by the victim's file system layout.

For the sake of effectiveness, training datasets should be built starting from measurements taken when samples run within a sandbox provided of a realistic FS environment.

- *In facts, running a sample on an almost empty FS gives poor information about its behavior with respect to when it is run on a computer whose HD is full of working documents and folders.*

76

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Building training datasets: issues

TRACKING RANSOMWARE ON THE FILE SYSTEM

Traditionally, analysts build their dataset by placing on line some honeypots in order to capture samples «in the wild».

Unfortunately, this is a time-consuming task that often brings rather unsatisfactory results.

Nowadays on-line cooperative sample repositories such as *VirusShare*, *VirusTotal* etc. offer an invaluable alternative to have rich and well categorizied datasets.

77

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Building training datasets: issues

TRACKING RANSOMWARE ON THE FILE SYSTEM

However, two points should be taken in account:

Many samples may result inactive due to, their C&C network infrastructure is no longer available, or because their capability to realize they are running in a sandobox.

There is the concrete risk that (coalitions of) adversaries submit misleading datasets in order to affect the detectors training and testing process *(adversarial training)*

78

MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Conclusion

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

79

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Conclusion

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Security technologies using AI and ML are a reality today.

Some people believe AI powered security products are designed to be a "*set-it and forget-it*" solution, replacing the human operator with some sort of robot.

80

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Conclusion

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

Security technologies using AI and ML are a reality today.

Some people believe AI powered security products are designed to be a "*set-it and forget-it*" solution, replacing the human operator with some sort of robot.

That's not true!

81

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Conclusion

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

AI and ML in cybersecurity *do not eliminate humans* from the equation.

Such solutions *equip the humans* with the tools they need to better defend their organization.

When implemented correctly, AI and ML can be a *force multiplier*.

The goal is to teach a cybersecurity technology to automatically *recognize the threats,* reduce *false positives* and doing that *much faster* than humans can do.

82

MEIM MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

Conclusion

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR CYBERSECURITY

In cybersecurity, ML is used to create models that often contain a large number of good and malicious pieces of data.

- These could be real-time pieces of data captured "*in the wild*" or data that was collected from known samples for the sake of training the system.
- As an ML engine runs a model, it makes assumptions about what is good data, what is malicious data and what is still unclear.
- Humans train the ML engine, telling it what assumptions were correct, what mistakes were made and what still needs to be rerun.

83

MEIM | MASTER IN ENTREPRENEURSHIP INNOVATION MANAGEMENT IN COLLABORATION WITH MIT SLOAN

IN COLLABORATION WITH MIT MANAGEMENT SLOAN SCHOOL

UNIVERSITÀ DEGLI STUDI DI NAPOLI PARTHENOPE

MASTER MEIM 2022-2023

Thank you!