Analyzing and Attributing Cyber-Attacks

446H - Applied Network Security

Erisa Karafili

Imperial College London February 15, 2019

Agenda

- Examples of Digital Investigation Process Models
- Attacks' Modus Operandi and Motivations
- 3 A Reasoner for Attributing Cyber-Attacks
 - The Attribution Problem
 - Argumentation-Based Reasoner

1 Examples of Digital Investigation Process Models

2 Attacks' Modus Operandi and Motivations

- 3 A Reasoner for Attributing Cyber-Attacks
 - The Attribution Problem
 - Argumentation-Based Reasoner

Main Steps in Digital Investigation Process Models

- **Preparation**: Create a plan of actions
- Identification: Finding potential sources of evidence
- Preservation: Collecting and storing the evidence
- Examination and Analysis: Extracting and viewing information from the evidence, and analyzing it, by answering different questions (who, what, where, when, how and why)
- **Presentation**: Reporting the findings in a satisfiable way (legal, corporate, military, etc.)

Different Types of Evidence

The evidence can be distinguished in different categories:

- Digital Evidence e.g., email, logs, invoices, /var/log/messages;
- Network-Based Digital Evidence e.g., chat log, emails, browser activities, logs;
- Real Evidence, e.g., physical hard drive or USB device, the computer itself;
- Best Evidence, e.g., a file recovered from the hard drive, a snapshot of a network transaction;
- Direct Evidence e.g., somebody is stating "I saw him with that USB";
- Circumstantial Evidence, e.g., email signature, a file containing password hashes;
- Hearsay Evidence e.g., a personal letter, a memo, bookkeeping records;

Imperial College

OSCAR (1 of 4)

OSCAR is a Network Forensics Investigative Process Model. Its main steps are:

- Obtain
- Strategise
- Collect
- Analyse
- Report

OSCAR (2 of 4)

- Obtain information about the incident and about the environment.
 - Description of the incident, date, time, method of discovery, persons involved, systems and data involved, actions taken since discovery, time frame for the investigation/recovery/ resolution, legal issues;
 - Information about the environment (that is constantly changing, complex social and political dynamics): business model, legal issues, network topology, available sources of network evidence, organizational structure, incident response management process/procedures, communication system, resources available.

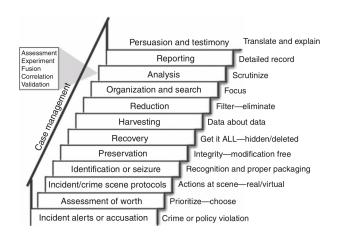
Oscar (3 of 4)

- **Strategise**: it is crucial to have a strategy where you assess your resources and plan your investigation.
 - Investigation strategy: Understand the goals and time frame of the investigation; understand your resources and identify sources of evidence, for each source of evidence estimate the cost and value of obtaining it, prioritise the evidence acquisition, plan the initial acquisition/analysis; decide upon methods and time of communications/updates, if needed, iterate.
- Collect evidence: acquire ASAP, analyze only copies, use tools that are reputable and reliable, document everything.
 - Document all the actions taken during evidence collection e.g., systems accessed log, date, time, sources, method of acquisition.
 - Capture the evidence e.g., capture the packets, copy logs, image hard drivers of web proxies or logging servers.
 - Store and transport the collected evidence.

OSCAR (4 of 4)

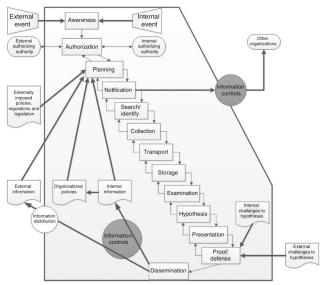
- **Analyze**: this process is usually non-linear but certain elements should be considered essential.
 - Correlation: how data are related to each other, e.g., sources of evidence with the data;
 - Timeline: create a timeline of activities by using the collected and correlated evidence;
 - Events of interest: some events will stand out more than others, especially the one related to the hypotheses;
 - Corroboration: try to eliminate the "fake evidence" by using multiple sources and by identifying inconsistencies;
 - Recovery of additional evidence: the above processes might need further evidence, so we might need to re-iterate them in order to recover new evidence;
 - Interpretation: the interpretation of events is a hypothesis, that can be proved or disproved.
- Report: document and report the investigation in order to be understandable by technical and non technical people and imperial College factual.

Staircase Model



Taken from "Digital Evidence and Computer Crime", Eoghan Casey, Third Edition, 2011.

Evidence Flow Model



1 Examples of Digital Investigation Process Models

2 Attacks' Modus Operandi and Motivations

- 3 A Reasoner for Attributing Cyber-Attacks
 - The Attribution Problem
 - Argumentation-Based Reasoner

Modus Operandi

Modus operandi (MO) means "a method of operating", it answers the *how* question

An example of MO behaviour is the following:

- Amount of planning before a crime;
- Materials used by the offender, e.g., system type, connection type, software;
- Presurveillance of the victim/target;
- Offender precautionary actions e.g., IP spoofing, aliases, anti-forensics countermeasures.

Usefulness of the Modus Operandi

Understanding the modus operandi of the attacker helps to:

- Put in act mitigation action for the current attack;
- Put in place preventive measures;
- 3 Attribute the attack to a possible culprit.

Example: Ukraine Power Grid Cyber-Attack

In 23 Dec 2015 a power grid cyber-attack took place in Ukraine¹. The hackers compromised successfully the information systems of three energy distribution companies, and disrupted the electricity supply to around 230 thousand end-users, and left them without electricity for 1-6 hours.

- The corporate networks were compromised using spear-phishing email with BlackEnergy malware;
- The attackers took control of SCADA networks and switched off the substations;
- Disabled/destroyed IT infrastructure components, e.g., modems, converters;
- They used KillDisk malware to destroy the files (that could be used as evidence) stored on servers and workstation;
- Denial-of-service attack on the call-center.

Imperial College

https://ics.sans.org/media/E-ISAC_SANS_Ukraine_DUC_5.pdf ondon

Example of MO: Ukraine Power Grid Attack Dec 2015

- Planning: Spear-phishing campaign to corporate networks, Manipulation of Microsoft Office documents that contained BlackEnergy malware;
- Tool used: BlackEnergy malware with phishing emails, KillDisk malware;
- Results of the actions were: Get control of the SCADA network, Destroyed IT infrastructures of 3 Energy Distribution Companies, Denial of Service Attack on Call-Center, Destruction of files on servers using KillDisk.

Example of MO: Actions

- Spear-phishing to gain access to the companies business networks;
- Theft of credentials from the business networks;
- Used virtual private networks (VPNs) to enter the ICS network;
- Used existing remote access tools within the environment;
- Reconfigured the Uninterruptable Power Supplies (UPS) that provide backup power to two of the control centres;
- Replaced the firmware on serial-to-ethernet converts with a malicious one written by the attackers;
- At 3.30pm 23 December the attackers entered the SCADA network using the hijacked VPNs, and disabled the UPS system they had already reconfigured;
- Launched a telephone denial-of-service attack on the call-center;
- Open the breakers for the shut down;
- Used a modified KillDisk to erase the master boot records of impacted organization systems and the targeted deletion of logs; imperial College
 The KillDisk was launched using a logic bomb.

Example of MO: Ukraine Power Grid Attack (Cont.)

- Pre-surveillance: the spear-phishing campaign took place months before the attack, and the attacker were observing the victim.
- Precaution actions: KillDisk destroyed all the files on the servers, including all the left evidence. When the systems were rebooted they had an error "Operating system not found".

Motives of the Attack

- Answer the question "Why";
- Understanding the motives of an attack helps to attribute the attack to a possible culprit;
- Some of the motives that can push the offenders to commit a cyber-crime are:
 - Power Reassurance, Power Assertive, Anger Retaliatory,
 Sadistic, where usually the attacker knows the victims, e.g.,
 cyberstalking;
 - Opportunistic and Profit Oriented e.g., monetary profits (ransomware), service disruption that will cause the victim economic lost, political motives.

Ukraine Power Grid Attack Dec 2015: Motives (I)

Given the high level of organization of the attack, we expect the possible culprit to be an organized, well-trained, and well-funded group of attackers, or a collaboration between different group of attackers.

- The attack maybe had economical motives, because there were economical losses due to the service disruption and the costs of restore the service (in some cases it took months).
 - The attack did not had economical motives because:
 - The economical losses were not as high as the impact it had on the people affected.
 - The attacker could have done more damages.
- The attack maybe had political motives due to the geopolitical conflict between Ukraine and Russia, with Russian annexation of Crimea in 2014.

Ukraine Power Grid Attack Dec 2015: Motives (II)

- The attack maybe had political motives as a response to the physical attack pro-Ukrainian activists made to the substations feeding power to Crimea, leaving 2 million Crimean residents without power, as well as a Russian naval base, right before Dec 2015. The physical attack was a reaction to Crimean authorities nationalising Ukrainian-owned energy companies.
 - The preparation of the attack started in Spring 2015, thus it cannot be considered as a retaliation for the attack on the Crimean substations. It could have been a catalyst but not the original motivation.
- The attack had political and somehow also economical and power assertion motives because:
 - The Ukrainian parliament was preparing a bill for nationalize privately owned power companies in Ukraine (most of them were owned by a Russian oligarch). Thus, the attack can be seen as a message sent to the Ukrainian authorities.
 - The attackers limited the damaged.

1 Examples of Digital Investigation Process Models

2 Attacks' Modus Operandi and Motivations

- 3 A Reasoner for Attributing Cyber-Attacks
 - The Attribution Problem
 - Argumentation-Based Reasoner

The Attribution Problem

Attribution is the process of assigning a cyber-attack to an entity

 The growing of connectivity increases the security challenges and the need for efficient countermeasures

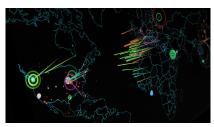


Attribution permits efficient attacker-oriented countermeasures

- The attribution process is a difficult one
- Attribution is mainly human based

Digital Forensics Techniques

- Digital Forensics techniques help during the attribution process by collecting, storing and analyzing the evidence
- Digital Forensics techniques suffer from the quantity and quality problem;
- These techniques are not able to deal with:
 - incomplete information,
 - conflicting information,
 - and social evidence.



An Argumentation-Based Solution

Solution

An automatic reasoner (ABR) based on argumentation and abductive reasoning.

- Given evidence of the attack, ABR helps the forensics analyst during the analysis and attribution process;
- It works with incomplete and conflicting pieces of data;
- ABR works with technical and social evidence;
- It categorises the evidence using a social model;
- ABR provides an explainable attribution.

Preference-Based Argumentation Framework

ABR uses a preference-based argumentation framework

Definition

An argumentation theory is a pair $(\mathcal{T}, \mathcal{P})$ of argument rules \mathcal{T} and preference rules \mathcal{P} .

The argument rules ${\mathcal T}$ are a set of labelled formulas of the form:

$$rule_i: L \leftarrow L_1, \ldots, L_n$$
.

The preference rules are a set of labelled formulas of the form:

$$p: rule_1 > rule_2$$

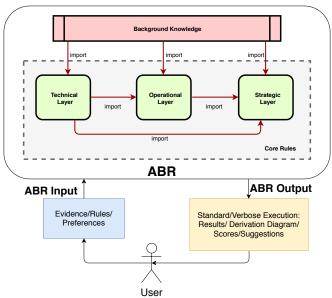
where $rule_1$, $rule_2$ are labels of rules in \mathcal{T} , and > is higher priority relation between the rules.

Social Model used by ABR

- ABR is based on the Q-Model
- The Q-Model represents how the analysts perform the attribution process of cyber-attacks
- The pieces of evidence and the reasoning rules are divided in three layers



Argumentation-Based Reasoner for Attribution



ABR Example: Ukraine Power Grid Attack Dec 2015 (1/4)

- The attack had as target Power Supplies companies in Ukraine.
- The attack (dec) was a high level skill attack.
- The attack was constructed to have a specific target.

```
target(ukraine_power, dec).
industry(electricity, ukraine_power).
highLevelSkill(dec).
targetCountry(ukraine, dec).
malwareUsedInAttack(killdisk, dec).
malwareUsedInAttack(blackenergy, dec).
attackPeriod(dec, [2015, 12]).
specificTarget(dec).
```

ABR Example: Ukraine Power Grid Attack Dec 2015 (2/4)

• Russia had economical motives to perform the attack.

```
target(ukraine_power, dec).
industry(electricity, ukraine_power).
highLevelSkill(dec).
targetCountry(ukraine, dec).
malwareUsedInAttack(killdisk, dec).
malwareUsedInAttack(blackenergy, dec).
attackPeriod(dec, [2015, 12]).
specificTarget(dec).
hasEconomicMotive(russian_federation, ukraine_power).
```

ABR Example: Ukraine Power Grid Attack Dec 2015 (2/4)

• Russia had economical motives to perform the attack.

```
target(ukraine_power, dec).
industry(electricity, ukraine_power).
highLevelSkill(dec).
targetCountry(ukraine, dec).
malwareUsedInAttack(killdisk, dec).
malwareUsedInAttack(blackenergy, dec).
attackPeriod(dec, [2015, 12]).
specificTarget(dec).
hasEconomicMotive(russian_federation, ukraine_power).
industry(ukraine_power).
```

ABR Example: Ukraine Power Grid Attack Dec 2015 (3/4)

 Russia had political motives (a response to the earlier physical attack occurred in Crimea) to perform the attack.

```
target(ukraine_power, dec).
industry(electricity, ukraine_power).
highLevelSkill(dec).
targetCountry(ukraine, dec).
malwareUsedInAttack(killdisk, dec).
malwareUsedInAttack(blackenergy, dec).
attackPeriod(dec, [2015, 12]).
specificTarget(dec).
hasPoliticalMotive(russian_federation, ukraine, [2015, 11]).
```

ABR Example: Ukraine Power Grid Attack Dec 2015 (4/4)

• Russia had motives to perform the Ukraine Power Grid Attack.

```
target(ukraine_power, dec).
industry(electricity, ukraine_power).
highLevelSkill(dec).
targetCountry(ukraine, dec).
malwareUsedInAttack(killdisk, dec).
malwareUsedInAttack(blackenergy, dec).
attackPeriod(dec, [2015, 12]).
specificTarget(dec).
hasMotive(russian_federation, dec).
```

ABR Execution US Bank Hack Example (1 of 4)

US bank hack occurred in 2012, where US banks faced denial of service (Dos) attacks.

ABR Execution US Bank Hack Example (1 of 4)

US bank hack occurred in 2012, where US banks faced denial of service (Dos) attacks.

```
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
```

ABR Execution US Bank Hack Example (1 of 4)

US bank hack occurred in 2012, where US banks faced denial of service (Dos) attacks.

 The banks' web hosting services were infected by a malware called <u>Itsoknoproblembro</u>

```
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
malwareUsed(itsoknoproblembro, usbankhack).
```

US bank hack occurred in 2012, where US banks faced denial of service (Dos) attacks.

- The banks' web hosting services were infected by a malware called *ltsoknoproblembro*
- Itsoknoproblembro hijacked the corporate clouds

```
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
malwareUsed(itsoknoproblembro, usbankhack).
hijackCorporateClouds(usbankhack).
```

US bank hack occurred in 2012, where US banks faced denial of service (Dos) attacks.

- The banks' web hosting services were infected by a malware called *ltsoknoproblembro*
- Itsoknoproblembro hijacked the corporate clouds
- US placed economic sanctions against Iran in February 2012

```
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
malwareUsed(itsoknoproblembro, usbankhack).
hijackCorporateClouds(usbankhack).
imposedSanctions(usa, iran, [2012, 2]).
```

```
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
malwareUsed(itsoknoproblembro, usbankhack).
hijackCorporateClouds(usbankhack).
imposedSanctions(usa, iran, [2012, 2]).
```

```
t_{-}1: highLevelSkill(Att) \leftarrow hijackCorporateClouds(Att).
```

```
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
malwareUsed(itsoknoproblembro, usbankhack).
hijackCorporateClouds(usbankhack).
imposedSanctions(usa, iran, [2012, 2]).
```

 $t_1: \textit{highLevelSkill}(\textit{usbankhack}) \leftarrow \textit{hijackCorporateClouds}(\textit{usbankhack}).$

```
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
malwareUsed(itsoknoproblembro, usbankhack).
hijackCorporateClouds(usbankhack).
imposedSanctions(usa, iran, [2012, 2]).
highLevelSkill(usbankhack).
```

```
t_2: reqHighRes(Att) \leftarrow highLevelSkill(Att).
```

```
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
malwareUsed(itsoknoproblembro, usbankhack).
hijackCorporateClouds(usbankhack).
imposedSanctions(usa, iran, [2012, 2]).
highLevelSkill(usbankhack).
```

```
t_2 : reqHighRes(usbankhack) \leftarrow highLevelSkill(usbankhack).
```

```
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
malwareUsed(itsoknoproblembro, usbankhack).
hijackCorporateClouds(usbankhack).
imposedSanctions(usa, iran, [2012, 2]).
highLevelSkill(usbankhack).
reqHighRes(usbankhack).
```

```
op_{-}1 : hasPolMotive(C, T, Date) \leftarrow imposedSanctions(T, C, Date).
```

```
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
malwareUsed(itsoknoproblembro, usbankhack).
hijackCorporateClouds(usbankhack).
imposedSanctions(usa, iran, [2012, 2]).
highLevelSkill(usbankhack).
reqHighRes(usbankhack).
```

```
op\_1: hasPolMotive(iran, usa, [2012, 2]) \leftarrow imposedSanctions(usa, iran, [2012, 2]).
```

```
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
malwareUsed(itsoknoproblembro, usbankhack).
hijackCorporateClouds(usbankhack).
imposedSanctions(usa, iran, [2012, 2]).
highLevelSkill(usbankhack).
reqHighRes(usbankhack).
hasPolMotive(iran, usa, [2012, 2]).
```

From the background knowledge, we have that:

cybersuperpower(iran).

From the background knowledge, we have that:

cybersuperpower(iran).

 t_3 : $hasResources(X) \leftarrow cybersuperpower(X)$.

From the background knowledge, we have that:

cybersuperpower(iran).

 $t_3: hasResources(iran) \leftarrow cybersuperpower(iran).$

```
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
malwareUsed(itsoknoproblembro, usbankhack).
hijackCorporateClouds(usbankhack).
imposedSanctions(usa, iran, [2012, 2]).
highLevelSkill(usbankhack).
reqHighRes(usbankhack).
hasPolMotive(iran, usa, [2012, 2]).
hasResources(iran).
```

```
hasResources(X).
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
malwareUsed(itsoknoproblembro, usbankhack).
hijackCorporateClouds(usbankhack).
imposedSanctions(usa, iran, [2012, 2]).
highLevelSkill(usbankhack).
reqHighRes(usbankhack).
hasPolMotive(iran, usa, [2012, 2]).
hasResources(iran).
```

 $op_2: hasCapability(X, Att) \leftarrow regHighRes(Att),$

```
op\_2: hasCapability(iran, usbankhack) \leftarrow reqHighRes(usbankhack), hasResources(iran).
```

```
target(us_banks, usbankhack).
targetCountry(usa, usbankhack).
attackPeriod(usbankhack, [2012, 9]).
malwareUsed(itsoknoproblembro, usbankhack).
hijackCorporateClouds(usbankhack).
imposedSanctions(usa, iran, [2012, 2]).
highLevelSkill(usbankhack).
hasPolMotive(iran, usa, [2012, 2]).
hasResources(iran).
hasCapability(iran, usbankhack).
```

```
op\_3: hasMotive(C, Att) \leftarrow targetCountry(T, Att), \\ attackPeriod(Att, Date1), \\ hasPolMotive(C, T, Date2), \\ dateApplicable(Date1, Date2), \\ specificTarget(Att).
```

 $op_3: \textit{hasMotive(iran, usbankhack)} \leftarrow targetCountry(usa, usbankhack), \\ attackPeriod(usbankhack, [2012, 9]), \\ hasPolMotive(iran, usa, [2012, 2]), \\ dateApplicable([2012, 9], [2012, 2]), \\ specificTarget(usa).$

```
target(us_banks, usbankhack). ...
highLevelSkill(usbankhack).
reqHighRes(usbankhack).
hasPolMotive(iran, usa, [2012, 2]).
hasResources(iran).
hasCapability(iran, usbankhack).
hasMotive(iran, usbankhack).
```

```
str\_1: isCulprit(X, Att) \leftarrow hasMotive(X, Att), hasCapability(X, Att).
target(us\_banks, usbankhack).
\cdots
highLevelSkill(usbankhack).
reqHighRes(usbankhack).
hasPolMotive(iran, usa, [2012, 2]).
hasResources(iran).
hasCapability(iran, usbankhack).
```

hasMotive(iran, usbankhack).

 $str_1 : isCulprit(iran, usbankhack) \leftarrow$

```
hasCapability(iran, usbankhack).
target(us_banks, usbankhack).
highLevelSkill(usbankhack).
regHighRes(usbankhack).
hasPolMotive(iran, usa, [2012, 2]).
hasResources(iran).
hasCapability(iran, usbankhack).
hasMotive(iran, usbankhack).
```

Imperial College London

isCulprit(iran, usbankhack).

hasMotive(iran, usbankhack),

References

- Sherri Davidoff, Jonathan Ham. Network Forensics: Tracking Hackers Through Cyberspace. 2012.
- ② Eoghan Casey. Digital Evidence and Computer Crime: Forensic Science, Computers, and the Internet. 3rd Edition, 2011.
- Emmanuel S. Pilli, Ramesh C. Joshi, Rajdeep Niyogi. Network forensic frameworks: Survey and research challenges. Digital Investigation 7(1-2): 14-27 (2010).
- Nicole Beebe. Digital Forensic Research: The Good, the Bad and the Unaddressed. IFIP Int. Conf. Digital Forensics 2009: 17-36.
- Erisa Karafili, Linna Wang, Antonis Kakas, Emil C. Lupu. Helping Forensics Analysts to Attribute Cyber-Attacks: An Argumentation-Based Reasoner. PRIMA, 510-518 (2018).



e.karafili@imperial.ac.uk
http://www.imperial.ac.uk/people/e.karafili