MODELS FOR THREAT INTELLIGENCE

MALWARE ANALYSIS AND INCIDENT FORENSICS

M.Sc. in Cyber Security

A.Y. 2024/2025

SAPIENZA
UNIVERSITÀ DI ROMA

CIS SAPIENZA
CYBER INTELLIGENCE AND INFORMATION SECURITY

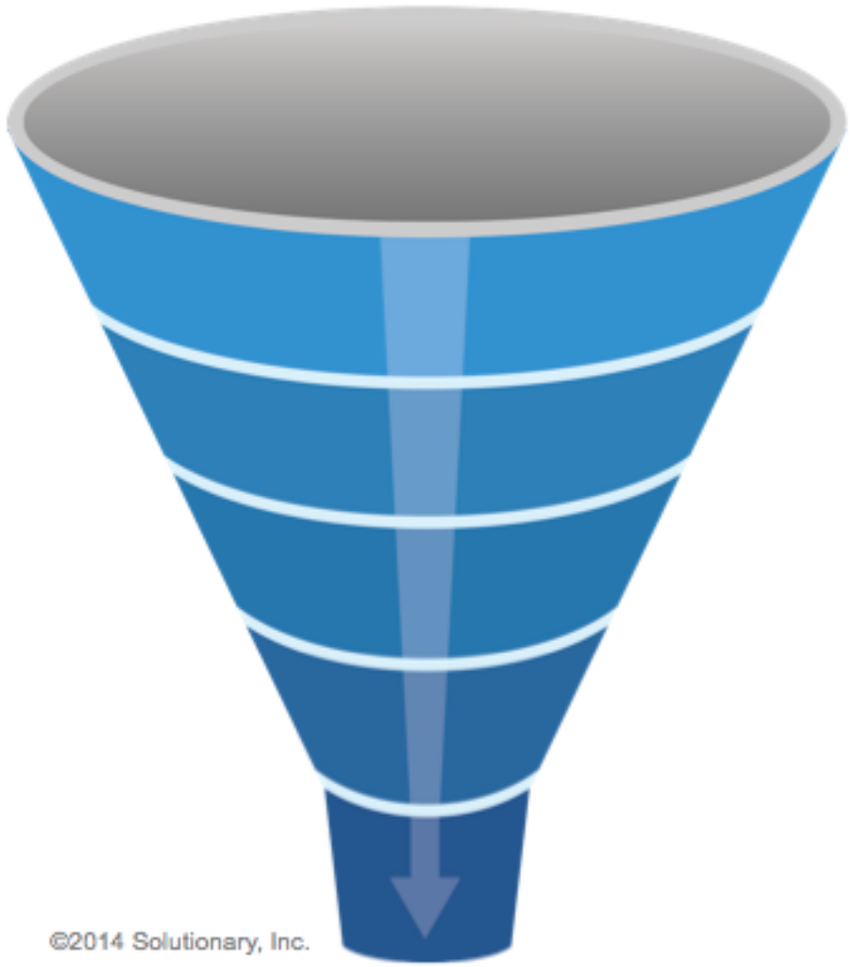
WHAT ARE WE LOOKING FOR?

Most dangerous attacks are today characterized by

- Precise targeting
- Use of advanced intrusion/obfuscation techniques
- Full control of target (RAT)
- Long-term persistence (APT)

The attacker (threat) has enough motivations/resources/skills to try to attack its target for prolonged periods of time, using several approaches.

CYBER THREAT INTELLIGENCE



NOISE
Is comprised of everything that is collected according to Priority Intelligence Requirements (PIR).

DATA
Noise is filtered out/scrubbed. Remaining artifacts are grouped according to defining characteristics.

INFORMATION
Data with a purpose. Once it is assigned a purpose, it becomes information.

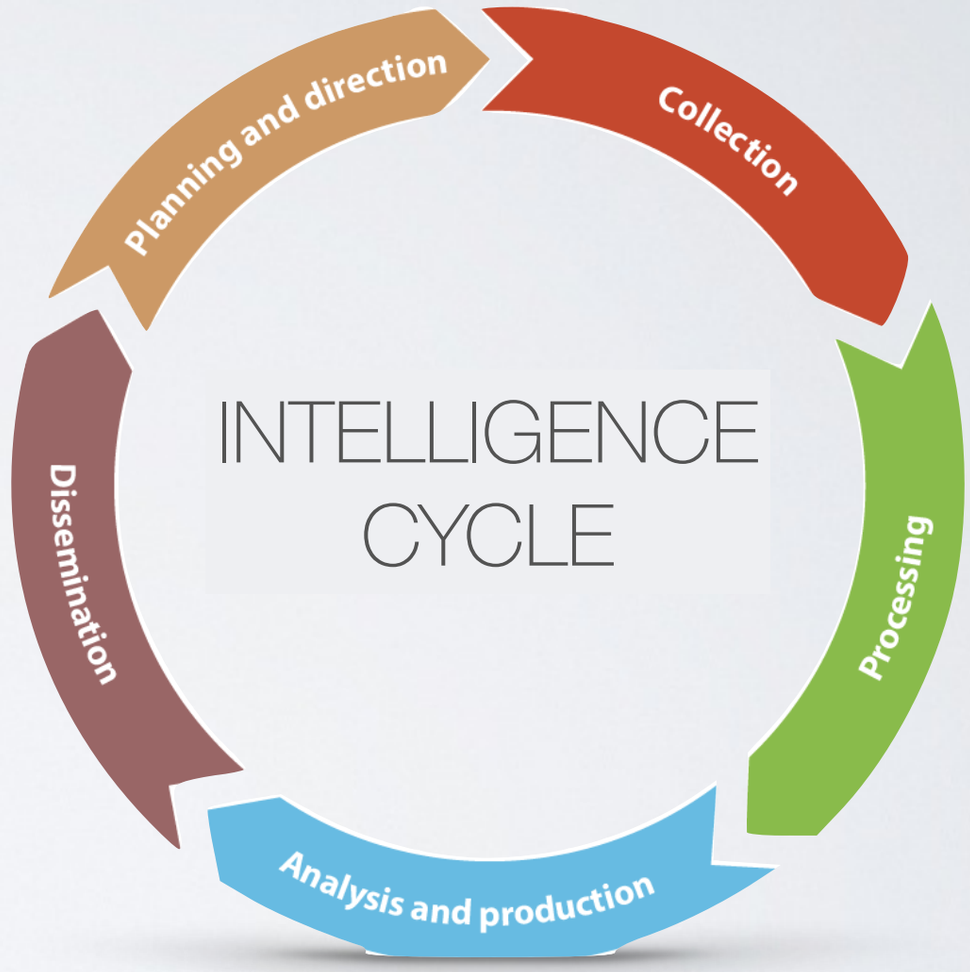
INTELLIGENCE
Information with a strategic purpose that can be used to gain advantage. This is exclusively a human centered activity.

ACTIONABLE INTELLIGENCE
Intelligence led, evidenced based assessments that can be initiated and acted upon.

CYBER THREAT INTELLIGENCE

An effective threat intelligence process has the main purpose of characterising threats in order to identify valuable preventative controls.

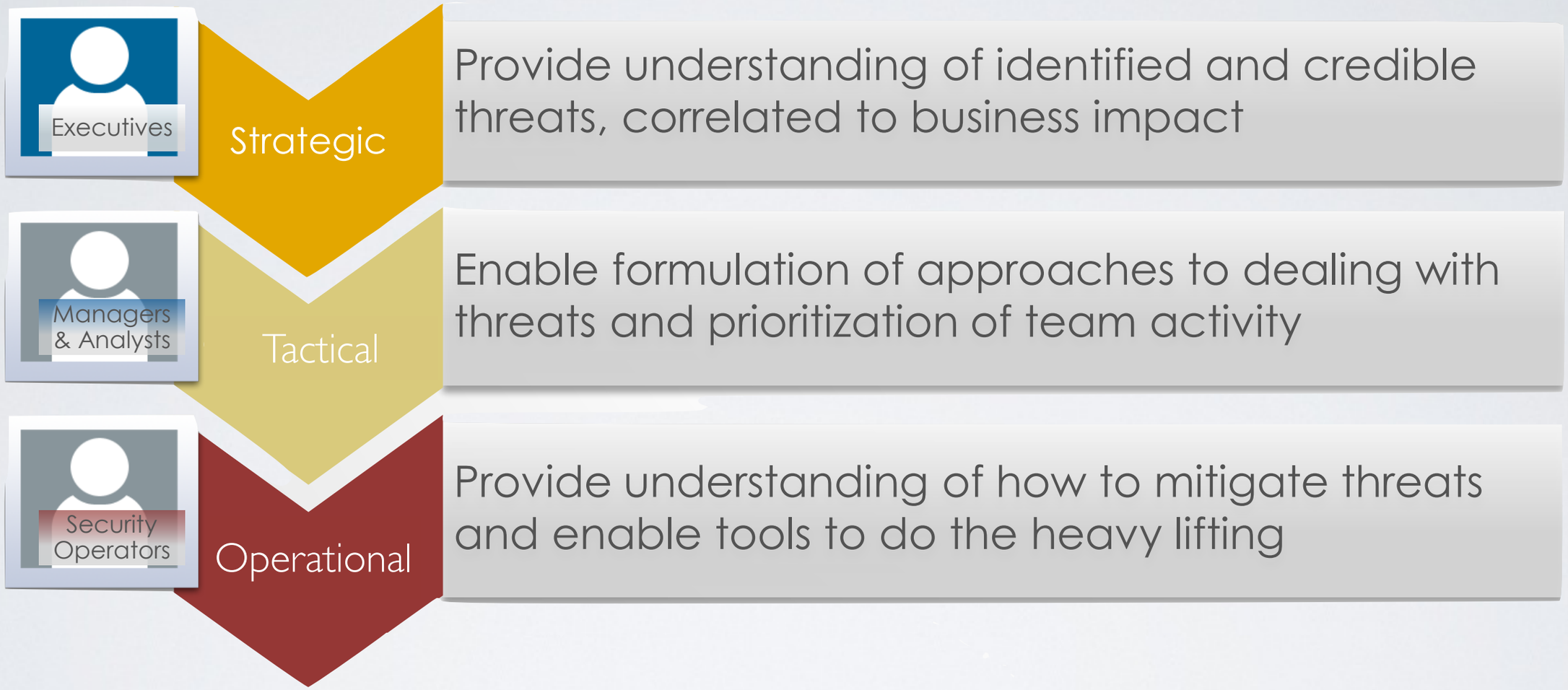
- Identify attackers
- Profile attackers
- Identify key tactics
- Rebuild playbook
- Replay playbook
- Utilize results



src: https://blog.nettitude.com/uk/eight-things-to-consider-before-deploying-cyber-threat-intelligence

CYBER THREAT INTELLIGENCE

Goals at several levels:



src: http://nova.issa.org/wp-content/uploads/2014/07/iSIGHT-Partners_ISSA-NOVA-v2.pdf

CYBER THREAT INTELLIGENCE

How do we make intelligence “actionable”?

- We assume the attack is resource constrained and behaves like a rational human being
- Cost factors
  - Expertise
  - Time
  - Money
  - Resources
- Success factors
  - Target ubiquity
  - Probability
  - Access

CYBER THREAT INTELLIGENCE

How do we make intelligence “actionable”?

- Value for the attacker comes from the possibility of re-using attack methodologies/techniques/tools as much as possible

- **Repeatability**: the capability to change the target and have the attack still work with the same success rate.

- **Scalability**: the capability to launch the attack against multiple targets with minimal cost per additional target.

CYBER THREAT INTELLIGENCE

How do we make intelligence “actionable”?

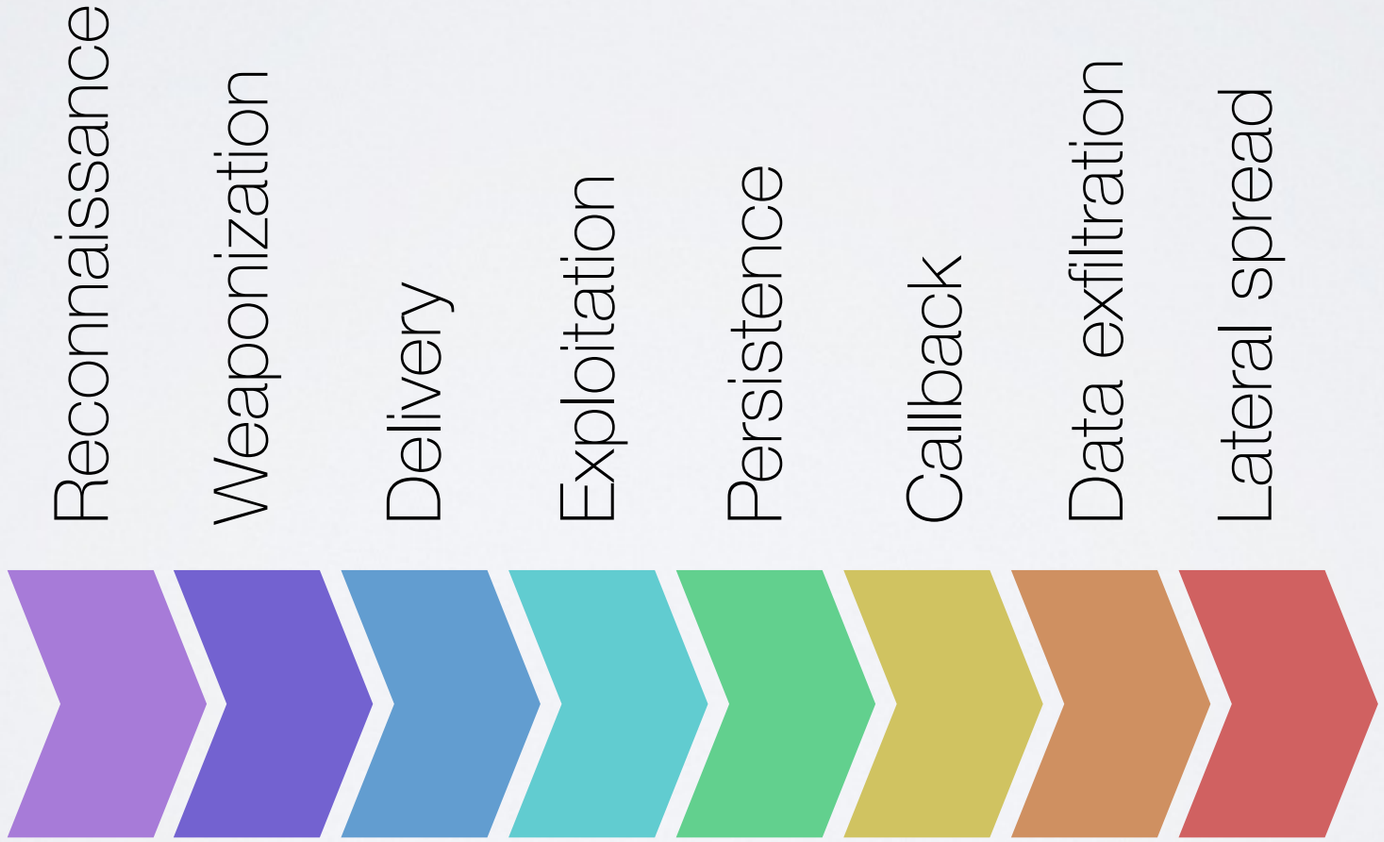
- Attackers determine the least costly and most valuable attacks based on
  - Who the targets are
  - Required success rate
  - Speed of conversion

Inexpensive, valuable, scalable, or repeatable:	**Costly, valueless, unscalable, or unrepeatable:**
Phishing	Web vulnerabilities
Credential reuse	0-day exploits
Known vulnerabilities with public exploits	Known vulnerabilities without public exploits
Office macros	Embedded devices
Spyware	Crypto weaknesses
Vendor compromise	Insider threat

INTRUSION KILL CHAIN

The kill chain is not just a way to represent attacks, but rather a methodology to analyse (tentatives of) intrusions to extract actionable information (intelligence).

Hutchins et al. “Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains”, Leading Issues in Information Warfare & Security Research 1.1 (2011)



INDICATORS

The fundamental element of intelligence in this model is the indicator

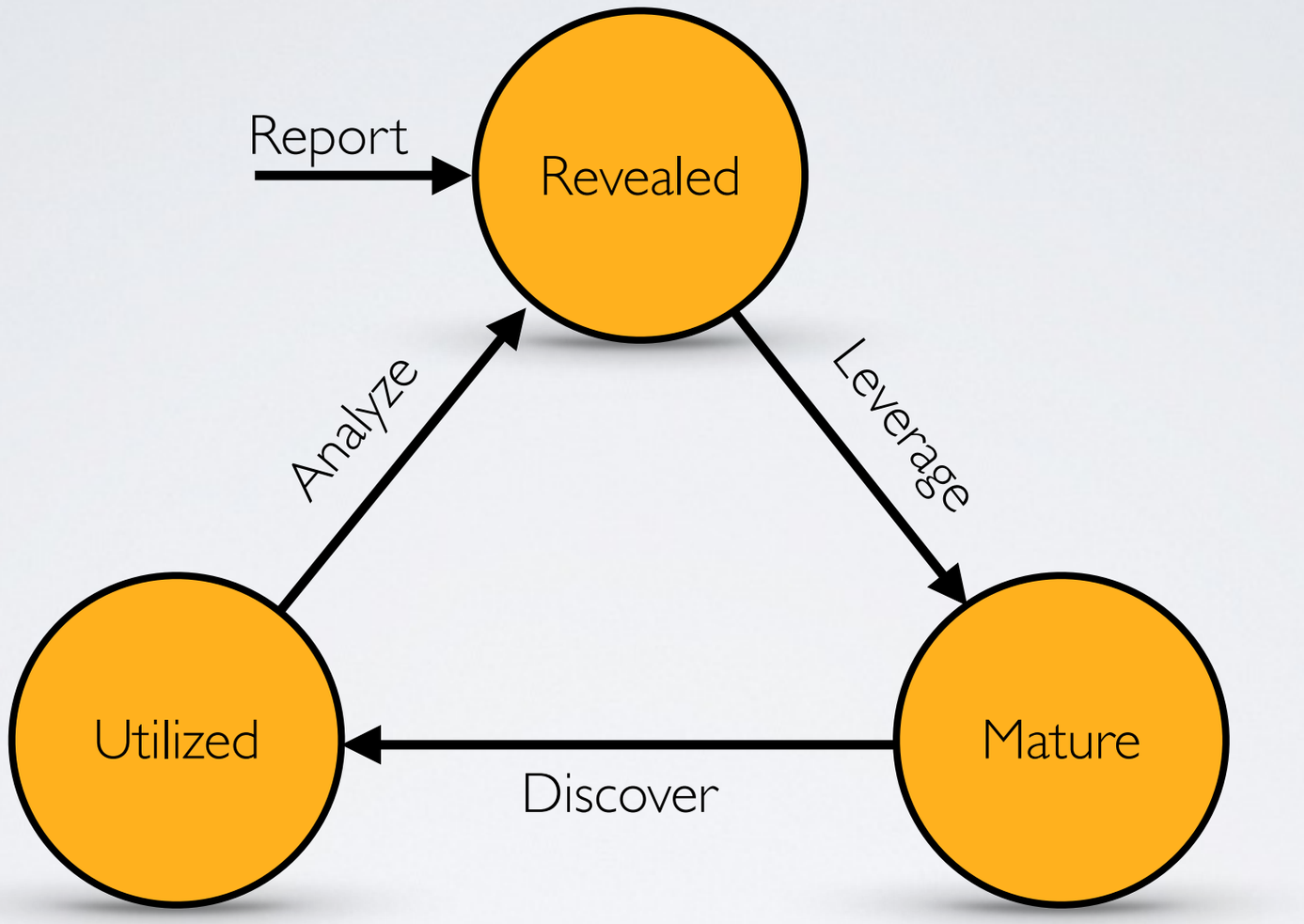
- **Indicator**: any piece of information that objectively describes an intrusion

Three types of indicators

- Atomic - those which cannot be broken down into smaller parts and retain their meaning in the context of an intrusion. (e.g. IP addresses, email addresses, vulnerability identifiers, etc.)
- Computed - those which are derived from data involved in an incident. (e.g. hash values)
- Behavioral - collections of computed and atomic indicators, often subject to qualification by quantity and possibly combinatorial logic. (e.g. “the intruder initially used a backdoor which generated network traffic matching [regular expression] at the rate of [some frequency] to [some IP address]”)

INDICATORS

Life cycle



INTRUSION KILL CHAIN

Indicators can be used to describe/identify various stages of the intrusion using the kill chain as a reference model

Phase	Indicators
Reconnaissance	[Recipient List] Benign File: tcnom.pdf
Weaponization	Trivial encryption algorithm: Key 1
Delivery	dn...etto@yahoo.com Downstream IP: 60.abc.xyz.215 Subject: AIAA Technical Committees [Email body]
Exploitation	CVE-2009-0658 [shellcode]
Installation	C:\...\fssm32.exe C:\...\IEUpd.exe C:\...\IEXPLORE.hlp
C2	202.abc.xyz.7 [HTTP request]
Actions on Objectives	N/A

COURSE OF ACTION

The intrusion kill chain becomes a model for actionable intelligence when defenders align enterprise defensive capabilities to the specific processes an adversary undertakes to target that enterprise

Phase	Detect	Deny	Disrupt	Degrade	Deceive	Destroy
Reconnaissance	Web analytics	Firewall ACL				
Weaponization	NIDS	NIPS				
Delivery	Vigilant user	Proxy filter	In-line AV	Queuing		
Exploitation	HIDS	Patch	DEP			
Installation	HIDS	“chroot” jail	AV			
C2	NIDS	Firewall ACL	NIPS	Tarpit	DNS redirect	
Actions on Objectives	Audit log			Quality of Service	Honeypot	

COURSE OF ACTION

Strategies to counteract intrusions:

- **Detect**: set up detection rules of an indicator for future events. (e.g. rules in an intrusion detection system (IDS), firewall or alert on SIEM)
- **Deny**: prevent the event from taking place. (e.g. firewall block)
- **Disrupt**: make the event fail as it is occurring. (e.g. file quarantining or memory protection measures)
- **Degrade**: slow down the further actions of the attacker. (e.g. bandwidth throttling)
- **Deceive**: learn more about the intentions of the attacker by making them think the action was successful. (e.g. use an honeypot)
- **Destroy**: offensive action against the attacker. (think twice about it…)

COURSE OF ACTION

Measure effectiveness to drive defense strategies over time!

	December	March	June
Reconnaissance			
Weaponization	◇	→	◇
Delivery	◆	→	◆
Exploitation	→ ◆	→	◆
Installation	◆ → ◆	→	◆
C2	◆ → ◆	→	◆
Actions on Objectives			

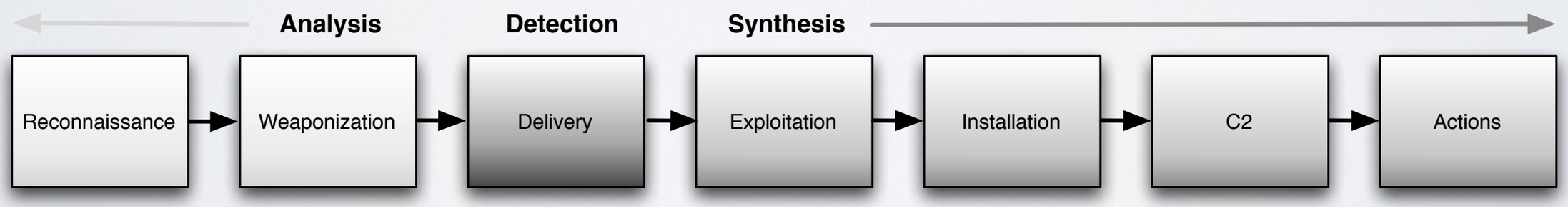
Legend ◇ *Detection* ◆ *Mitigation* → *Leverage new indicators*

RECONSTRUCT THE KILL CHAIN

Improved protection affects the course of actions

Late phase detection

Early phase detection



INVESTIGATE CAMPAIGNS

Recurring attacks from persistent threats require deeper understandings.

Phase	Indicators
Reconnaissance	[Recipient List] Benign File: tcnom.pdf
Weaponization	Trivial encryption algorithm: Key 1
Delivery	dn...etto@yahoo.com Downstream IP: 60.abc.xyz.215 Subject: AIAA Technical Committees [Email body]
Exploitation	CVE-2009-0658 [shellcode]
Installation	C:\...\fssm32.exe C:\...\IEUpd.exe C:\...\IEXPLORE.hlp
C2	202.abc.xyz.7 [HTTP request]
Actions on Objectives	N/A

INVESTIGATE CAMPAIGNS

Recurring attacks from persistent threats require deeper understandings.

Phase	Intrusion 1	Intrusion 2
Reconnaissance	[Recipient List] Benign File: tcnom.pdf	[Recipient List] Benign File: MDA_Prelim_09.pdf
Weaponization	Trivial encryption algorithm: Key 1	
Delivery	Downstream IP: 60.abc.xyz.215 Subject: AIAA Technical Committees [Email body]	Downstream IP: 216.abc.xyz.76 Subject: 7th Annual U.S. Missile Defense Conference [Email body]
	dn...etto@yahoo.com	
Exploitation	CVE-2009-0658 [shellcode]	
Installation	C:\...\fssm32.exe C:\...\IEUpd.exe C:\...\IEXPLORE.hlp	
C2	202.abc.xyz.7 [HTTP request]	
Actions on Objectives	N/A	N/A

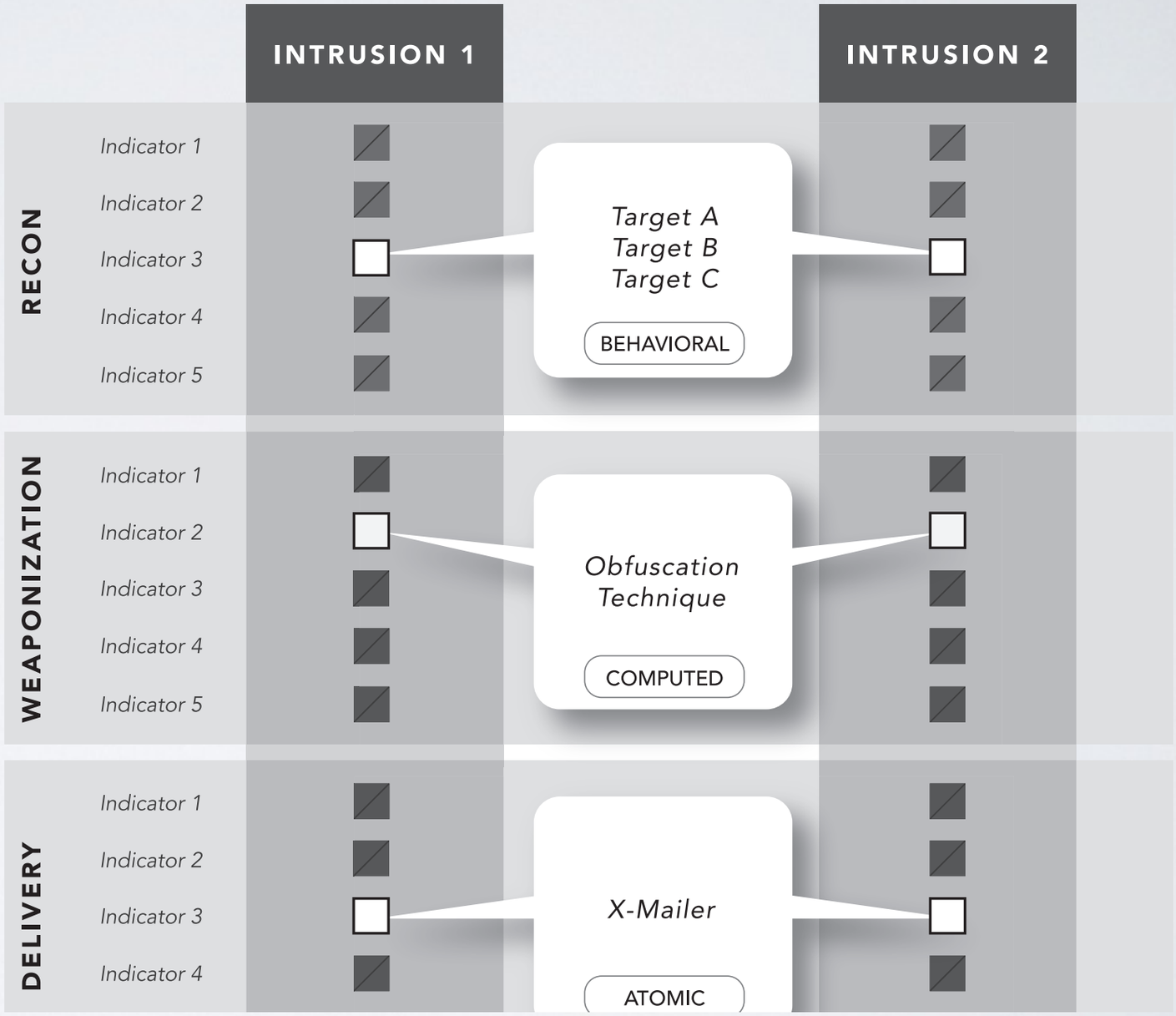
INVESTIGATE CAMPAIGNS

Recurring attacks from persistent threats require deeper understandings.

Phase	Intrusion 1	Intrusion 2	Intrusion 3
Reconnaissance	[Recipient List] Benign PDF	[Recipient List] Benign PDF	[Recipient List] Benign PPT
Weaponization	Trivial encryption algorithm		
	Key 1		Key 2
Delivery	[Email subject] [Email body]	[Email subject] [Email body]	[Email subject] [Email body]
	dn...etto@yahoo.com		ginette.c...@yahoo.com
	60.abc.xyz.215	216.abc.xyz.76	
Exploitation	CVE-2009-0658 [shellcode]		[PPT 0-day] [shellcode]
Installation	C:\...\fssm32.exe C:\...\IEUpd.exe C:\...\IEXPLORE.hlp		
C2	202.abc.xyz.7 [HTTP request]		
Actions on Objectives	N/A	N/A	N/A

INVESTIGATE CAMPAIGNS

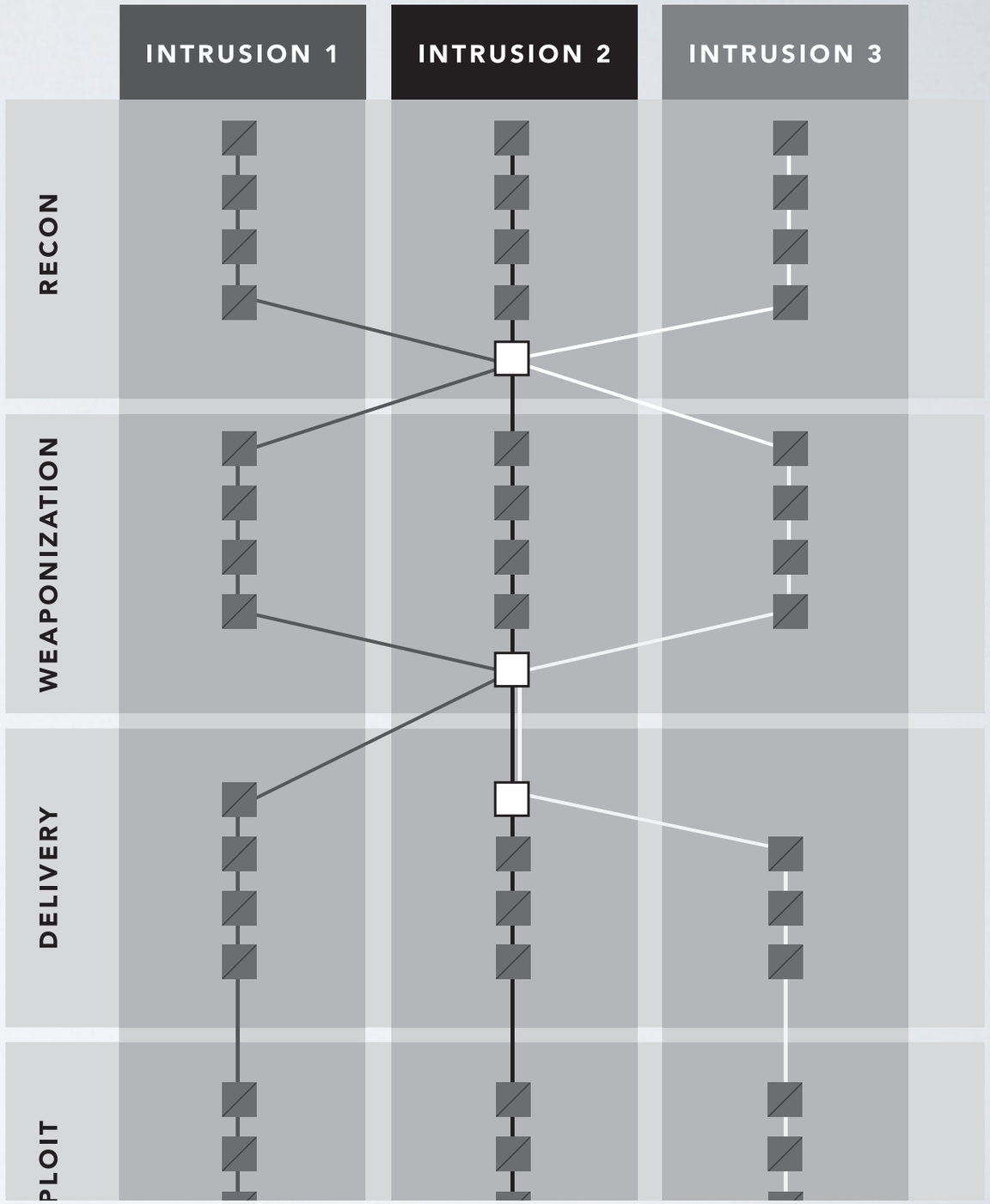
Compare intrusion indicators to find commonalities



INVESTIGATE CAMPAIGNS

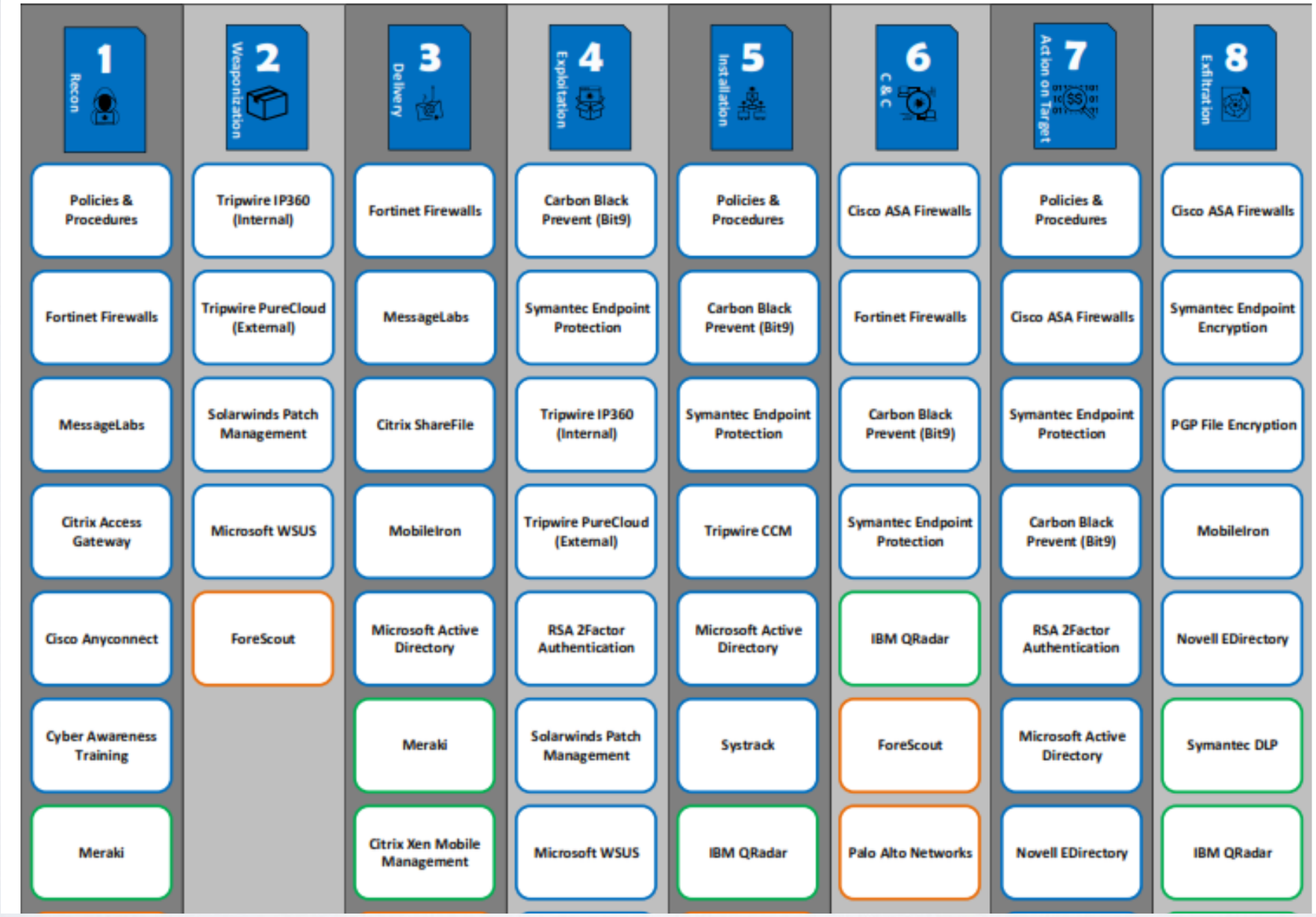
Inflection points among diverse intrusions may represent common patterns in a long-term campaign

Such patterns can be used to uniquely identify threat actors



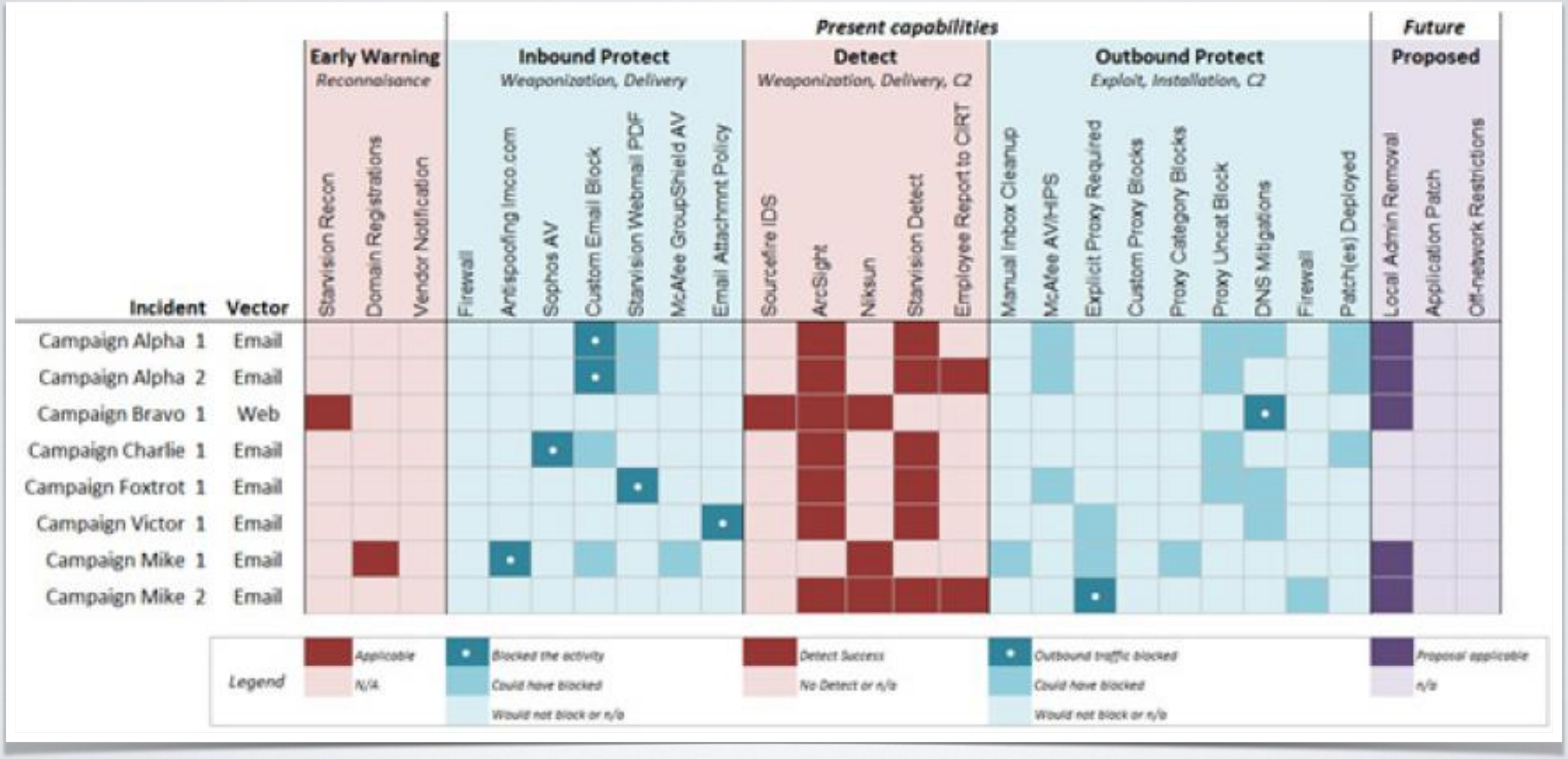
CKC USAGE SCENARIOS

Assess completeness of controls & countermeasures



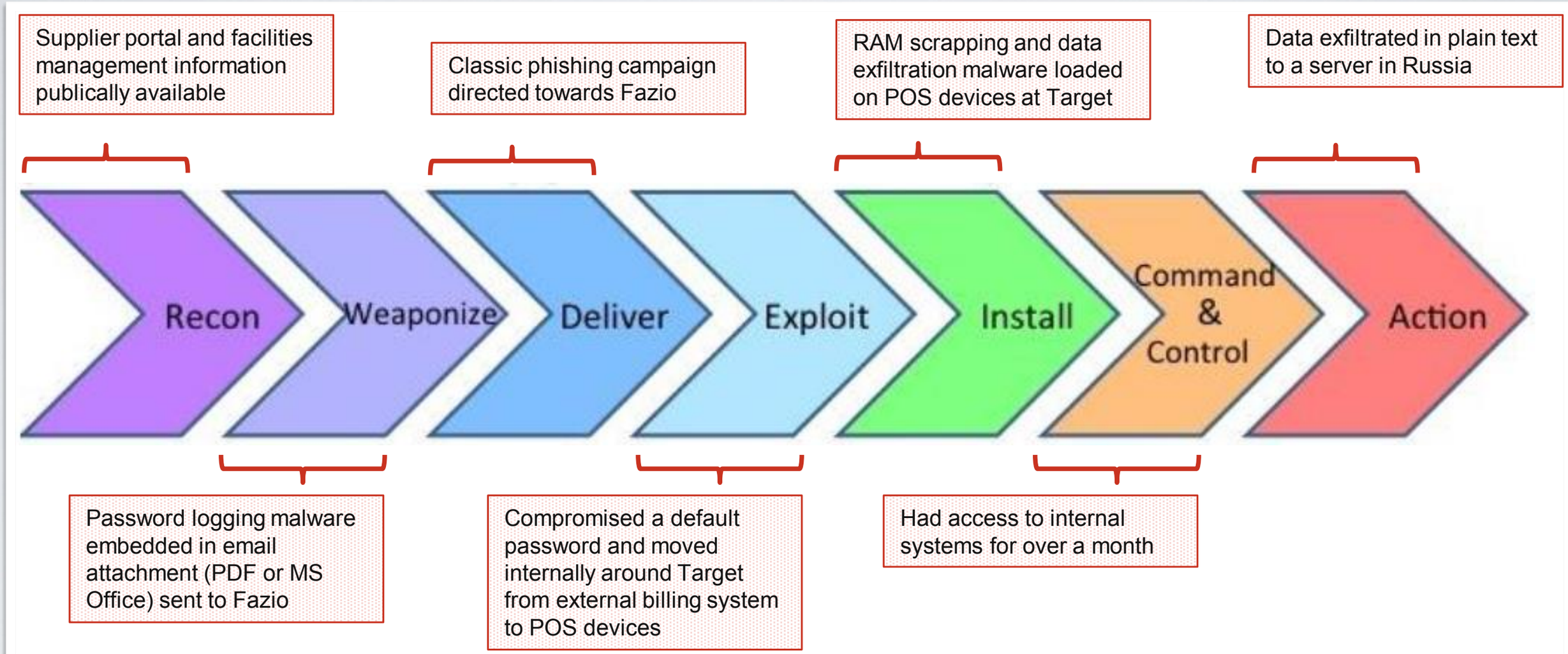
CKC USAGE SCENARIOS

Assess their effectiveness



CKC USAGE SCENARIOS

Post-mortem analysis…

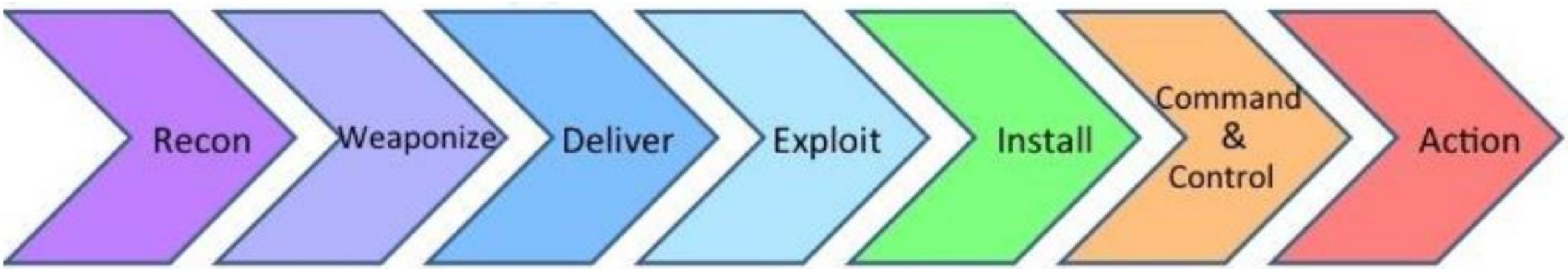


CKC USAGE SCENARIOS

…and lessons learned

Attackers took advantage of weak security at a Target vendor, gaining a foothold in Target's inner network.

Attackers took advantage of weak controls within Target's network and successfully maneuvered into the network's most sensitive areas.



Target missed warnings from its anti-intrusion software that attackers were installing malware in its network.

Target missed information provided by its anti-intrusion software about the attackers' escape plan, allowing attackers to steal as many as 110 million customer records.

THE PYRAMID OF PAIN

Types of indicators

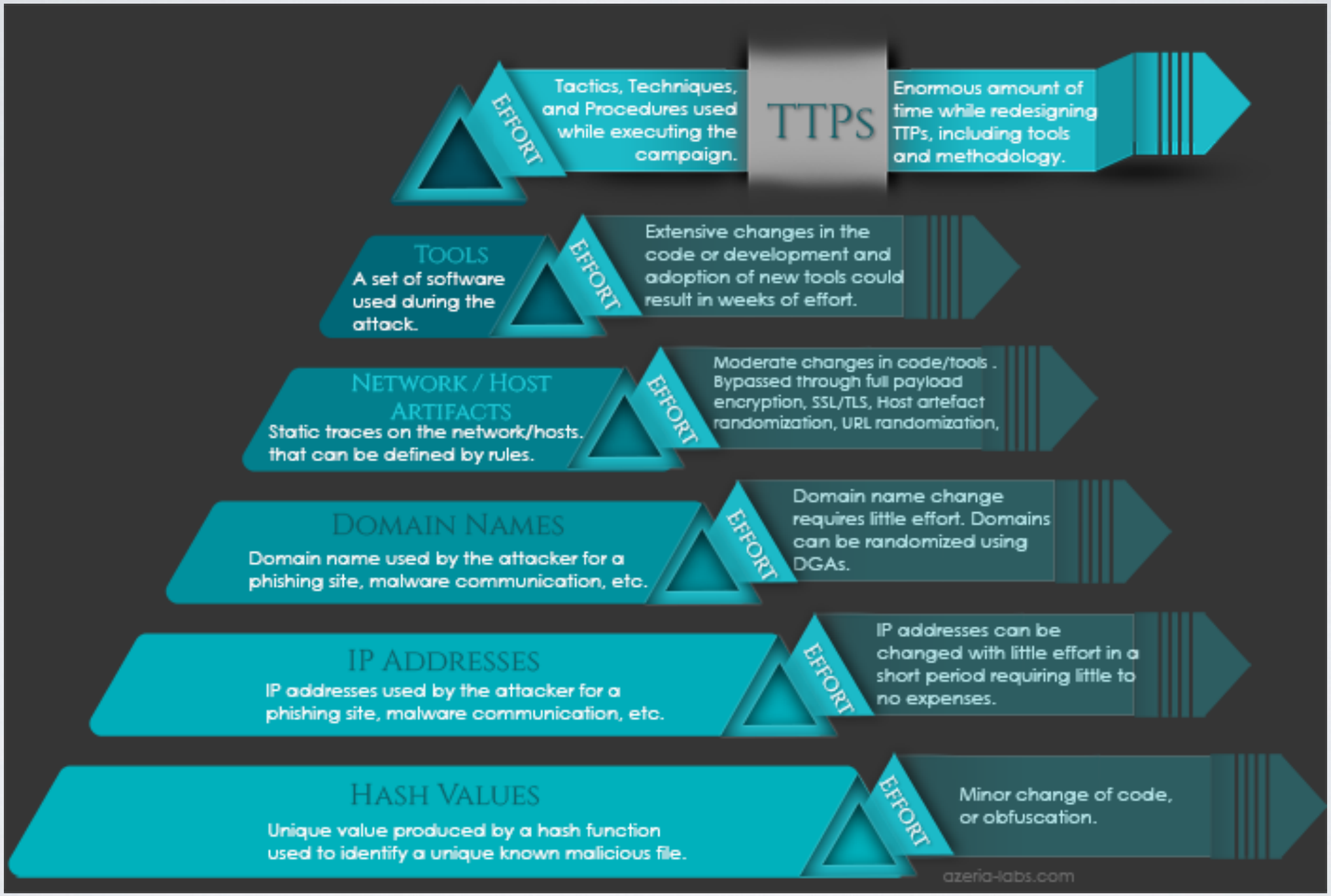
- **Hash Values**: SHA1, MD5 or other similar hashes that correspond to specific suspicious or malicious files. Often used to provide unique references to specific samples of malware or to files involved in an intrusion.
- **IP Addresses**: IP address or netblocks referring to the attacker.
- **Domain Names**: this could be either a domain name itself (e.g., "evil.net") or maybe even a sub- or sub-sub-domain (e.g., "this.is.sooooo.evil.net")
- **Network Artifacts**: observables caused by adversary activities on your network. In practice this really means those pieces of the activity that might tend to distinguish malicious activity from that of legitimate users. Typical examples might be URI patterns, C2 information embedded in network protocols, distinctive HTTP User-Agent or SMTP Mailer values, etc.

THE PYRAMID OF PAIN

Types of indicators

- **Host Artifacts**: observables caused by adversary activities on one or more of your hosts. They could be registry keys or values known to be created by specific pieces of malware, files or directories dropped in certain places or using certain names, names or descriptions or malicious services…
- **Tools**: software used by the adversary to accomplish their mission. Mostly this will be things they bring with them, rather than software or commands that may already be installed on the computer.
- **Tactics, Techniques and Procedures (TTPs)**: how the adversary goes about accomplishing their mission, from reconnaissance all the way through data exfiltration and at every step in between. "Dumping cached authentication credentials and reusing them in Pass-the-Hash attacks" would be a TTP.

THE PYRAMID OF PAIN



LIMITS

A tool that is not perfectly suited to all uses:

- User centric
- Low-level description of observables (events)
- Too focussed on malware-based attacks
- Several variants
- Attackers do not always follow a playbook

THE DIAMOND MODEL

This model guides the analysis process through several predefined steps that starting from indicators, strive to identify high-level facts about the security incidents that are analysed

- Describes security **events** (incidents) through few basic aspects
- Connects events that took place in a single intrusion through **activity threads**
- Identifies commonalities among activity threads forming **activity groups** that can then be used for strategic planning

Sergio Caltagirone, Andrew Pendergast, and Christopher Betz, "Diamond Model of Intrusion Analysis," Center for Cyber Threat Intelligence and Threat Research, Hanover, MD, Technical Report ADA586960, 05 July 2013

THE DIAMOND MODEL

The model represent "events" (intrusion) as a set of **core features** pertaining to few fundamental aspects

For every intrusion event there exists an adversary taking a step towards an intended goal by using a capability over infrastructure against a victim to produce a result.

An event defines discrete time-bound activity restricted to a specific phase

The adversary may require access to further external resources to reach its goal (**meta features**).

THE DIAMOND MODEL

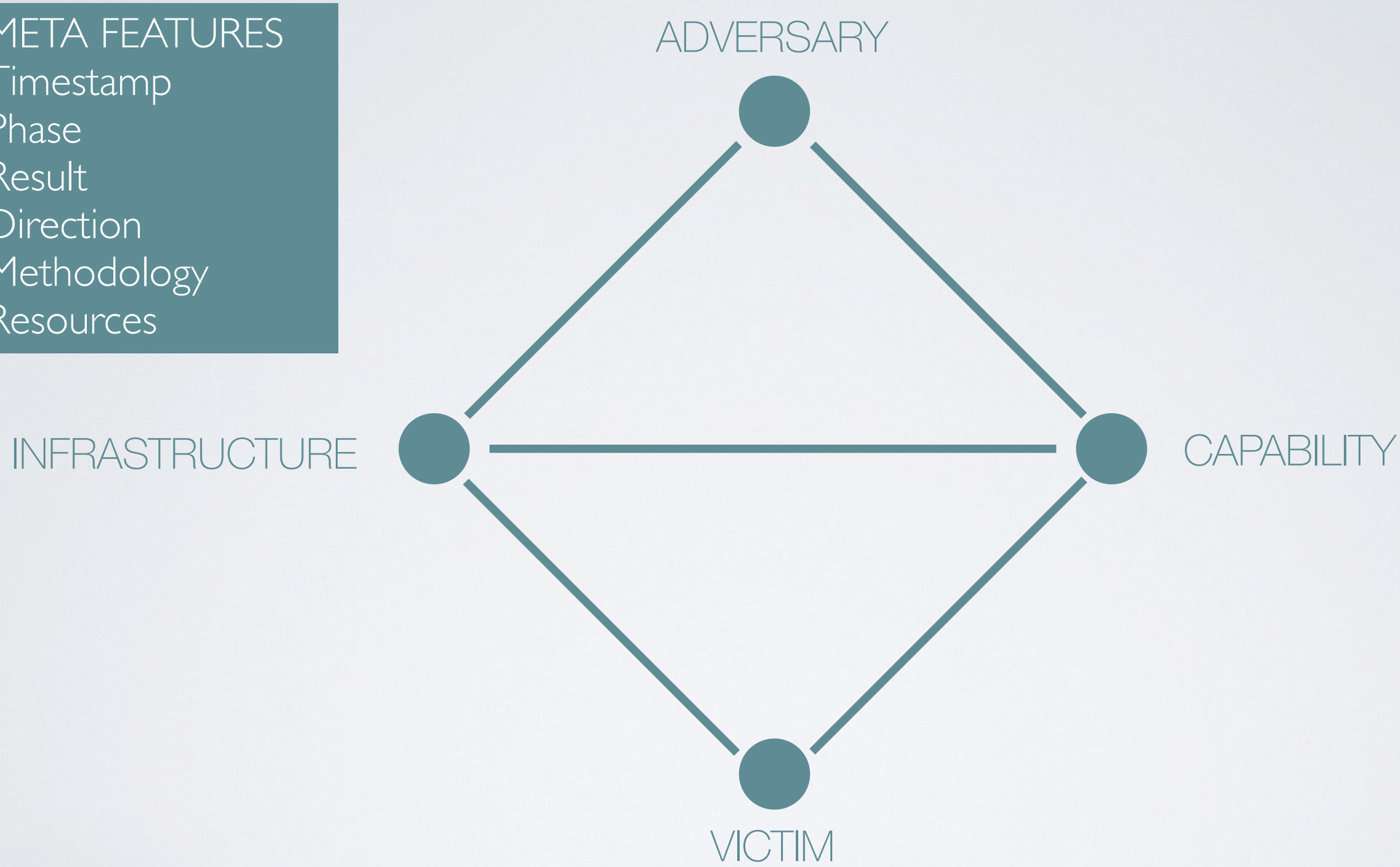
Elements can be further break down to detail them

$$
\begin{aligned}
\boldsymbol{E} = \langle & \langle Adversary, Confidence_{adversary} \rangle, \\
& \langle Capability, Confidence_{capability} \rangle, \\
& \langle Infrastructure, Confidence_{infrastructure} \rangle, \\
& \langle Victim, Confidence_{victim} \rangle, \\
& \langle Timestamp_{start}, Confidence_{timestamp_{start}} \rangle, \\
& \langle Timestamp_{end}, Confidence_{timestamp_{end}} \rangle, \\
& \langle Phase, Confidence_{phase} \rangle, \\
& \langle Result, Confidence_{result} \rangle, \\
& \langle Direction, Confidence_{direction} \rangle, \\
& \langle Methodology, Confidence_{methodology} \rangle, \\
& \langle Resources, Confidence_{resources} \rangle \rangle
\end{aligned}
\longrightarrow =
$$

$$
\begin{aligned}
& \langle \langle Organization, Confidence_{organization} \rangle, \\
& \langle HostIPAddress, Confidence_{IP} \rangle, \\
& \langle Hostname, Confidence_{Hostname} \rangle, \\
& \langle Application, Confidence_{Application} \rangle, \\
& \langle TCPPort, Confidence_{TCPPort} \rangle \rangle
\end{aligned}
$$

THE DIAMOND MODEL

META FEATURES
Timestamp
Phase
Result
Direction
Methodology
Resources



THE DIAMOND MODEL

ADVERSARY: actor/organisation responsible for utilising a capability against the victim to achieve their intent.

- **Adversary Operator** - the actual “hacker” or person(s) conducting the intrusion activity.
- **Adversary Customer** - the entity that stands to benefit from the activity conducted in the intrusion. It may be the same as the adversary operator, or it may be a separate person or group.

THE DIAMOND MODEL

CAPABILITY: tools and/or techniques of the adversary used in the event

- **Capability Capacity** - all of the vulnerabilities and exposures that can be utilised by the individual capability regardless of victim are considered its capacity.
- **Adversary Arsenal** - the adversary's complete set of capabilities, and therefore the combined capacities of their individual capabilities.

The C&C typically represent a specific kind of capability used by the adversary.

THE DIAMOND MODEL

INFRASTRUCTURE: describes the physical and/or logical communication structures the adversary uses to deliver a capability, maintain control of capabilities and effect results from the victim.

- **Directly controlled** infrastructure
- Infrastructure controlled by an **intermediary**
- **Public infrastructure** providers

THE DIAMOND MODEL

VICTIM: target of the adversary and against whom vulnerabilities and exposures are exploited and capabilities used.

- **Persona**: and individual or organization being targeted
- **Asset**: the attack surface and consist of the set of networks, systems, hosts, email addresses, IP addresses, social networking accounts, etc. against which the adversary directs their capabilities.

The set of vulnerabilities and exposures of a victim susceptible to exploitation is referred to as the victim susceptibilities

- They are sub-features of the victim

THE DIAMOND MODEL

META FEATURES

- Timestamp: date and/or time when the event occurred. May include start/end timestamps
- Phase: represents the specific logical role played by the event in an activity thread
  - You may want to stick with the cyber kill chain jargon for phases
- Results: post-conditions of an adversary's operations
  - Qualitative: <Success,Failure,Unknown>
  - Informative: Confidentiality|Integrity|Availability-compromised
  - Details: refer to the specific compromised assets

THE DIAMOND MODEL

META FEATURES

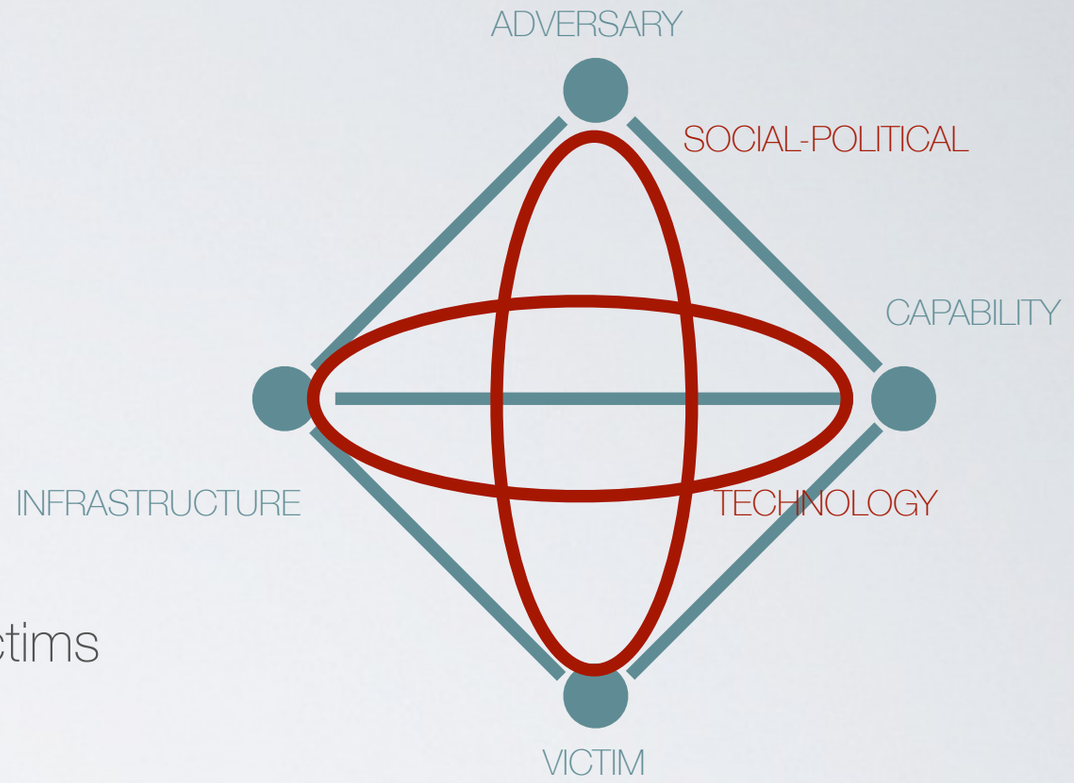
- Direction: describes qualitatively the main direction for actions in the event
  - Victim-to-Infrastructure, Infrastructure-to-Victim, Infrastructure-to-Infrastructure, Adversary-to-Infrastructure, Infrastructure-to-Adversary, Bidirectional, or Unknown
- Methodology: description of the general class of activity
  - E.g.: spear-phish email, content-delivery attack, syn flood, etc.
  - Better use an existing taxonomy
- Resources: all supporting elements on which the event, and therefore each core- and meta-feature, depends
  - E.g.: software, knowledge, information, hardware, funds, facilities, access

Meta features can be expanded

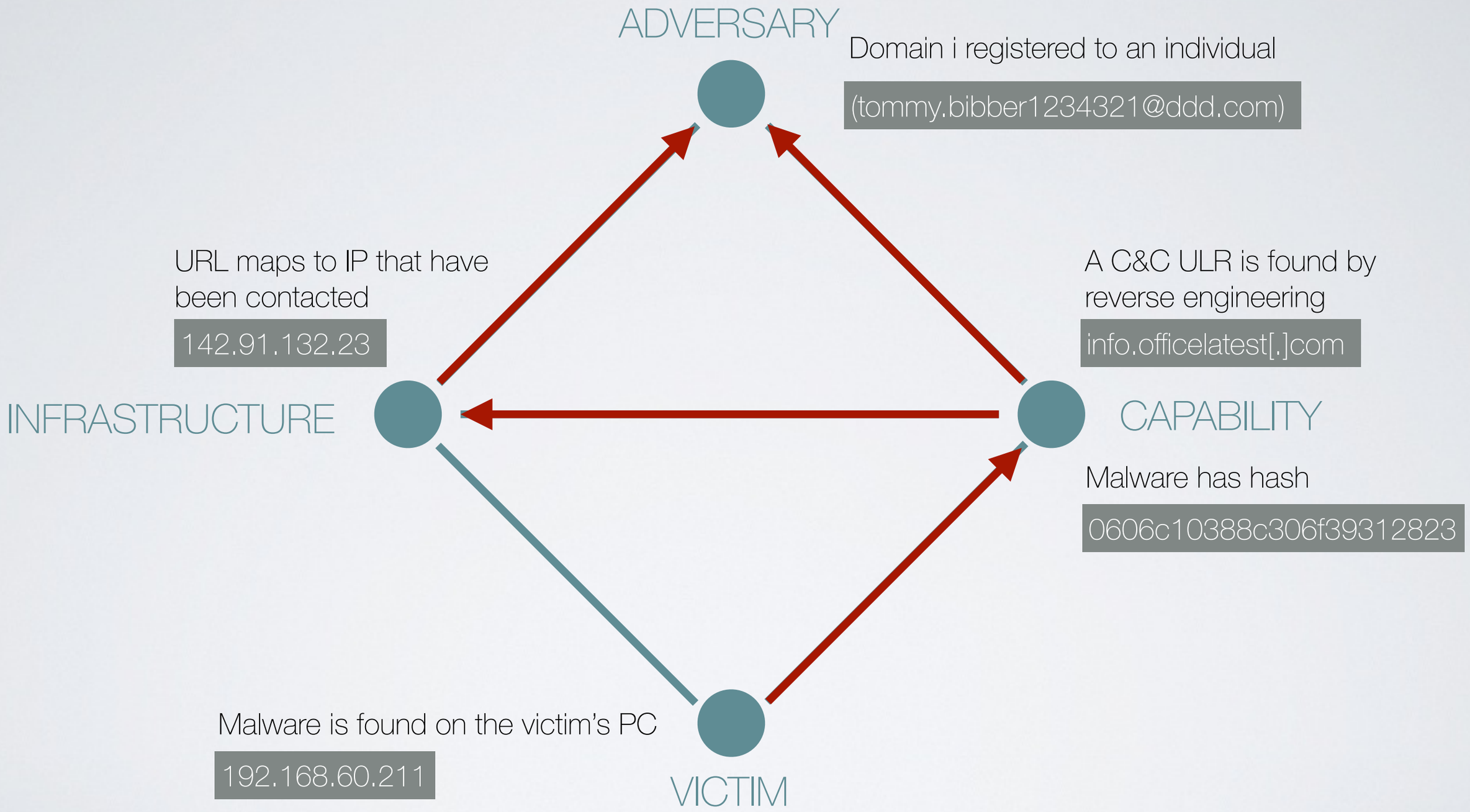
THE DIAMOND MODEL

Two useful feature extensions:

- SOCIAL-POLITICAL
  - Describes the relationship existing between the Adversary and its Victims
    - Intent
    - Persistence
  - Useful to defined a shared threat-space with other victims
- TECHNOLOGY
  - Describes the technology connecting and enabling the infrastructure and the capability to operate and communicate.



ANALYTICAL PIVOTING

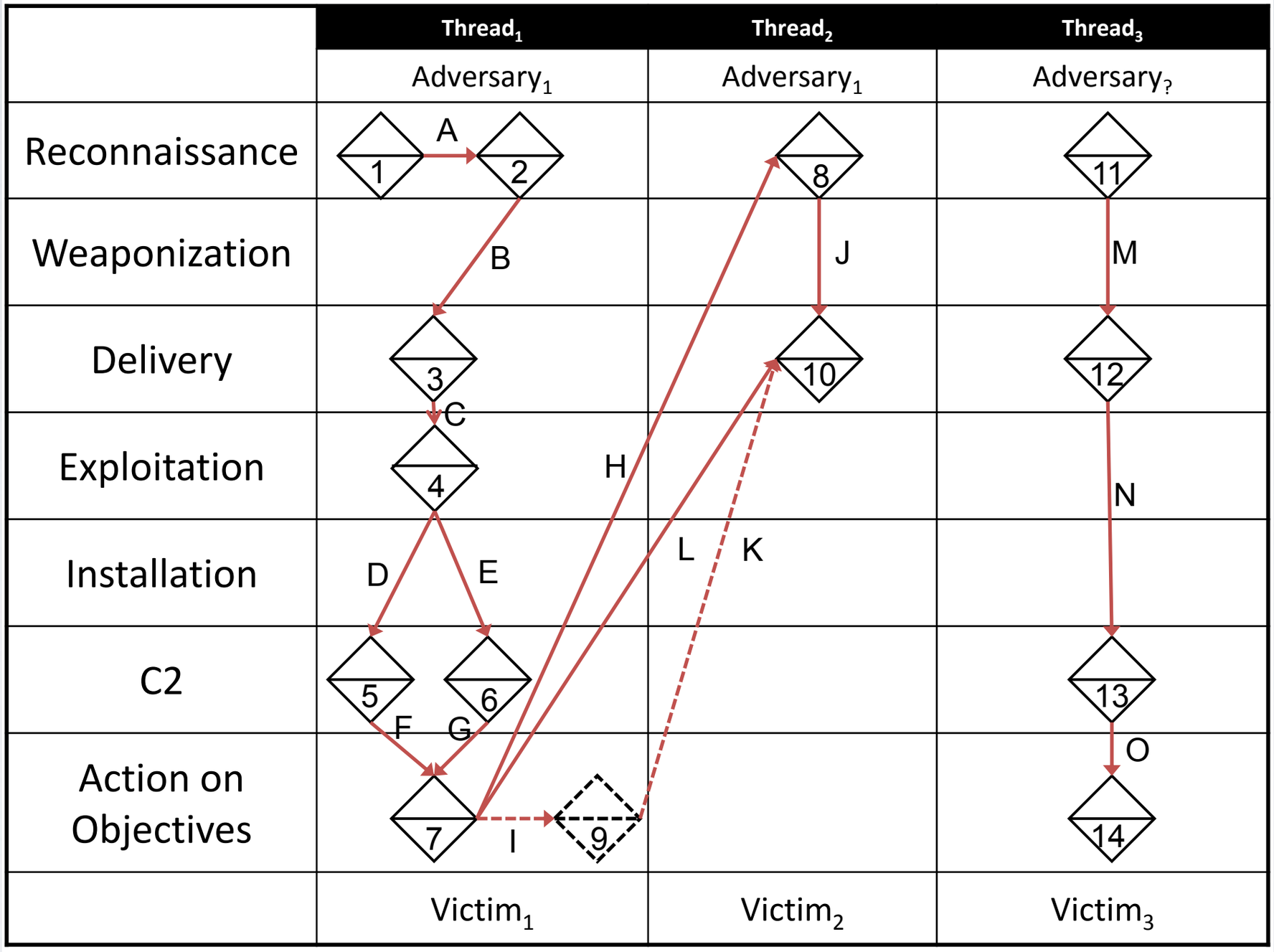


ANALYTICAL PIVOTING

Different approaches:

- **Victim-centered**: analyzing data related to a potential victim reveals the other related (and Diamond-connected) elements: malicious capabilities and infrastructure.
- **Capability-centered**: exploits features of a capability to discover those other related elements like (i) victims whom that capability is used against, (ii) infrastructure supporting the capability, (iii) technology enabling the capability, (iv) clues to other related capabilities, (v) and (possible) clues to the adversary.
- **Infrastructure-centered**: focuses on the malicious infrastructure of the adversary. From this element other related elements can be discovered. Difficult access.
- **Adversary-centered**: mostly reserved to law-enforcement agencies.

ACTIVITY THREAD



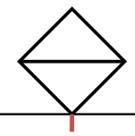
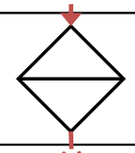
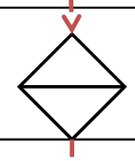
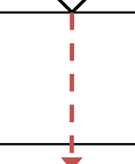
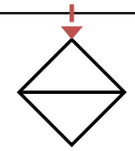
ACTIVITY THREAD

- **Vertical correlation**
  - Identify knowledge gaps
  - Fill those gaps with new knowledge
  - Establish causal relationships among events
- **Horizontal correlation**
  - The analytic process of causally linking events between vertical threads across adversary-victim pairs
  - identify common knowledge gaps between threads
  - use knowledge from one thread to fill knowledge gap in another
  - Identify common features across victims which can lead to the creation of an activity group

ACTIVITY THREAD

Identify Adversary processes

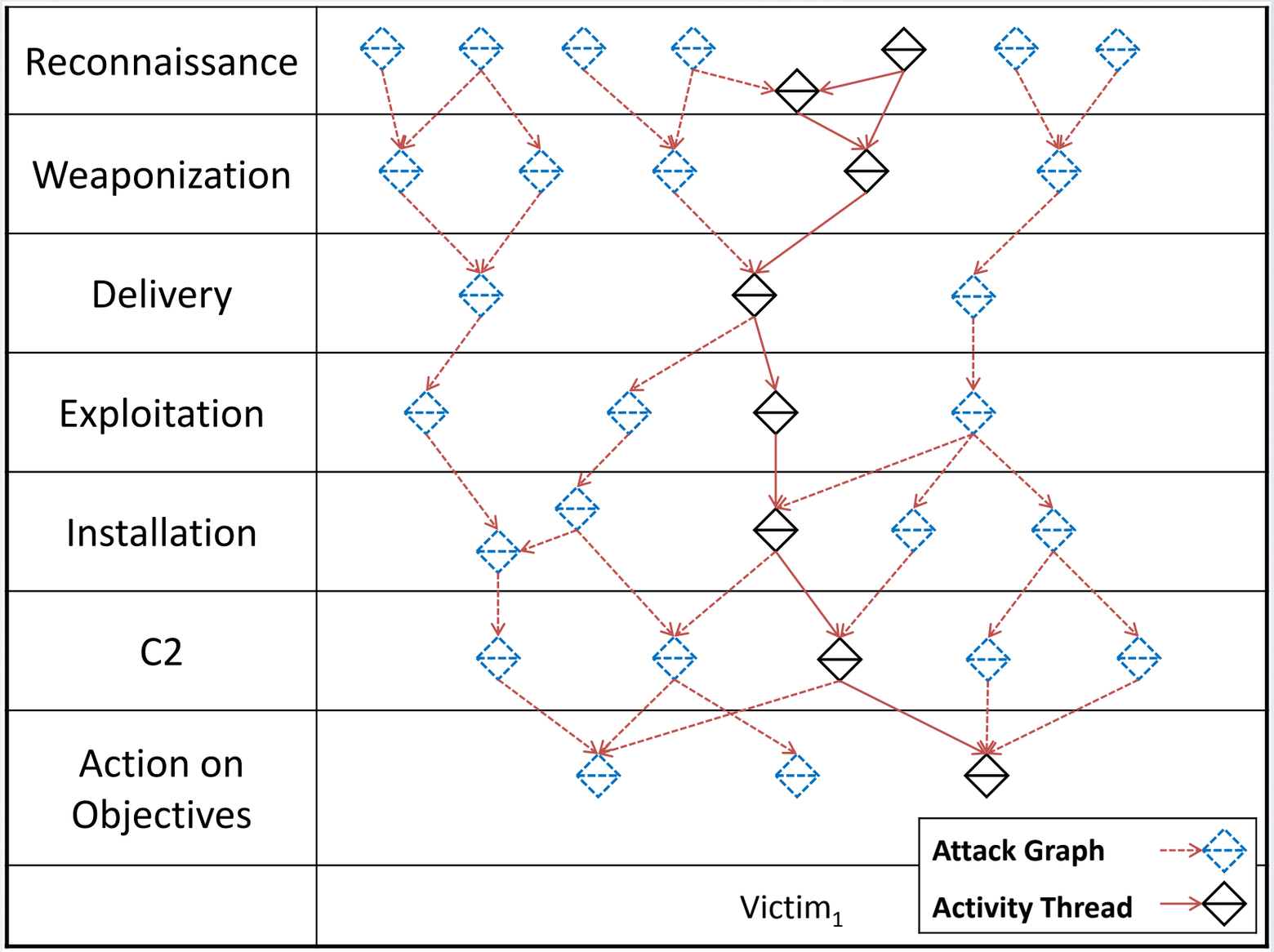
- Remove the details
- Focus on common recurring elements
- Check among known activity threads

		Process Features
Reconnaissance		Web search for “network administrator” [derived from event 2]
Weaponization		
Delivery		Email with trojanized attachment delivered [derived from event 3]
Exploitation		Specific local exploit (e.g., CVE-YYYY-XXX) [derived from event 4]
Installation		
C2		HTTP Post from victim [derived from event 6]
Action on Objectives		

ACTIVITY THREAD

Support for for analytical hypothesis checking

- Support through other approaches
- E.g. Attack graphs



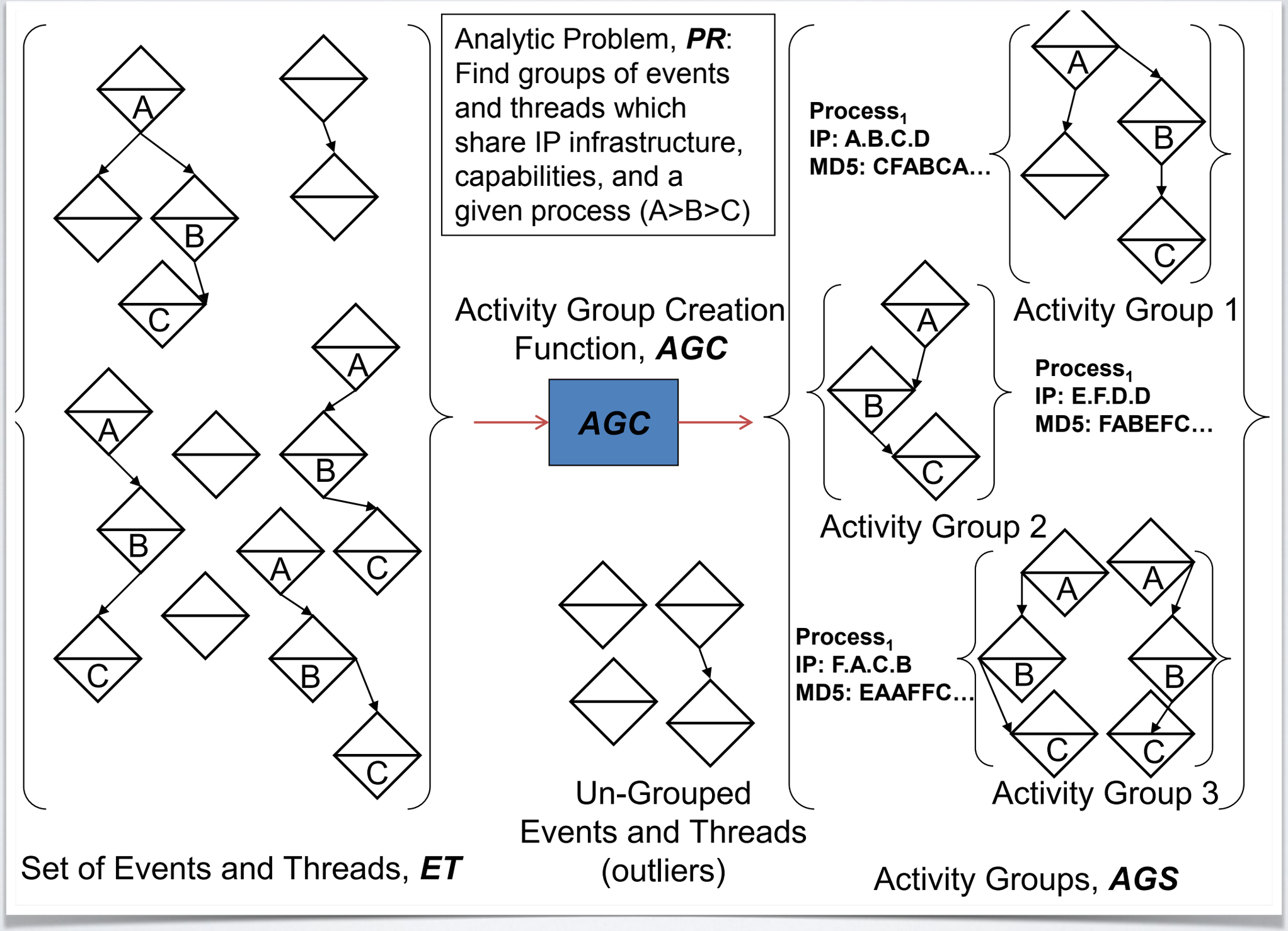
ACTIVITY GROUPS

Extract knowledge related to activity groups through a process:

- **Step 1: Analytic Problem** The particular analytic problem to be solved through grouping
- **Step 2: Feature Selection** The event features and adversary processes used to form the basis of classification and clustering are selected
- **Step 3: Creation** Activity groups are created from the set of events and threads

  Step 4: Growth As new events flow into the model, they are classified into the Activity Groups
- **Step 5: Analysis** Activity groups are analyzed to address the analytic problem(s) defined
- **Step 6: Redefinition** Activity groups need to be redefined from time-to-time to maintain their accuracy

ACTIVITY GROUPS



ACTIVITY GROUPS