Cyber Kill Chain, MITRE ATT&CK, and the Diamond Model: a comparison of cyber intrusion analysis models

Francesco Maria Ferazza

Technical Report

RHUL–ISG–2022–5

11 April 2022

Information Security Group
Royal Holloway University of London
Egham, Surrey, TW20 0EX
United Kingdom

CYBER KILL CHAIN, MITRE ATT&CK, AND THE DIAMOND MODEL: A COMPARISON OF CYBER INTRUSION ANALYSIS MODELS

Francesco M. Ferazza
RHUL student number: 190339407

Supervisor: Dr. Jorge Blasco Alis

Submitted as part of the requirements for the award of the
MSc in Information Security of the University of London.

Information Security Group
Royal Holloway, University of London
Egham, Surrey, TW20 0EX
United Kingdom

ACKNOWLEDGEMENTS

Before starting this dissertation, I would like to begin with the most important part.

I want to express my most sincere gratitude to my girlfriend Alessia and my parents, Stefano and Marzia.
The past year has been incredibly packed and hectic, with this dissertation to write, massive work projects undertaken, and extra shifts as a volunteer paramedic during the COVID-19 pandemic in northern Italy.

I would have never been able to endure all of this without their support, patience, and love.

I want to thank Filippo Caprioli and Joe Beard, two incredibly talented individuals, colleagues, and friends. Their amazing work and brilliant IT expertise are an example for me to follow.

A special mention goes to Claudio Neri, for having nurtured, over many years, my interest in the topics of security and intelligence analysis.

Last, but not least, I want to thank everyone in the Royal Holloway Information Security Group, in particular Jorge Blasco Alis for supervising this project and Konstantinos Mersinas, Martin Warren, Keith M. Martin, Peter Komisarczuk, and Lubna Ali for their prompt, professional, and kind support.

*"*WISDOM CONSISTS OF KNOWING HOW TO DISTINGUISH THE NATURE OF TROUBLE, AND IN CHOOSING THE LESSER EVIL.*"*[1]

CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

ACRONYMS

ACH - Analysis of Competing Hypotheses

AD - Active Directory

APT - Advanced persistent threat

AWS - Amazon Web Services

C2 - Command and Control

CKC - Cyber Kill Chain

CTI - Cyber Threat Intelligence

CVE - Common Vulnerabilities and Exposure

CVSS - Common Vulnerabilities Scoring System

DOD - U.S. Department of Defense

DOJ - U.S. Department of Justice

F2T2EA - Find, Fix, Track, Target, Engage, Assess

FBI - Federal Bureau of Investigation

FIN - FINancially motivated advanced threat groups

GCP - Google Cloud Platform

IaaS - Infrastructure as a Service

IOC - Indicator of Compromise

LM - Lockheed Martin

NIC - Network Interface Card

OWL - Ontology Web Language

PaaS - Platform as a Service

SaaS - Software as a Service

STIX - Structured Threat Information eXpression

TTPs - Tactics, Techniques, and Procedures

VPC - Virtual Private Cloud

ABSTRACT

Every day Advanced Persistent Threats become more sophisticated and aggressive, and so do their attacks and intrusions. Many intrusion analysis models and frameworks exist to help defenders and analysts in their efforts to understand and counter advanced adversarial campaigns.

This dissertation analyses the three most renowned and widely used models, Lockheed Martin's Cyber Kill Chain, MITRE's ATT&CK framework, and the Diamond Model.

A method to compare them, based on the differences between their purposes, designs, and levels of abstraction, will be presented and used.

This comparison will highlight the ontological and eschatological differences between the models and will illustrate how the three of them can be harmoniously used in a complementary, integrated way.

KEYWORDS:

CYBER KILL CHAIN | MITRE ATT&CK | DIAMOND MODEL | APT | INTRUSION ANALYSIS | INTELLIGENCE ANALYSIS | DEFENSIVE GAP ASSESSMENT | CYBER SECURITY ONTOLOGY

1 INTRODUCTION

This project is going to showcase, analyse, and compare the three main models for cyber intrusion analysis, Lockheed Martin's Cyber Kill Chain, the Diamond Model, and the MITRE ATT&CK framework.

1.1 PROJECT MOTIVATION AND CONTRIBUTION

While the three models examined are all actively used to perform intrusion analysis, their ultimate purposes and inner workings are vastly different.

Lockheed Martin's approach is aimed at augmenting a defender's resilience by dissecting malicious activity into an extremely specific kill chain.

MITRE's framework is a burgeoning cyber ontology aimed at providing a globally shared and curated knowledge base of documented adversarial behaviour.

The Diamond Model is a vade mecum, aimed at enhancing, structuring, and streamlining analysts' cognitive processes.

The objective of this dissertation is to analyse these differences, compare the models, and understand how they can be integrated and used concurrently.

There is scarce, almost non-existent, academic literature on this.
Most of the research available focuses on one model at a time, seeing how that one can be enhanced or integrated with other processes or technologies[2][3].

This project contribution is to bridge this knowledge gap by:

1. Explaining why intrusion analysis is a necessary task.
2. Analysing each model, and highlighting the very characteristics that make it unique.
3. Proposing a methodology to compare the models.
4. Applying said methodology to appreciate the differences between the models.

The project follows a precise narrative and information flow aimed at leading the readers in a step-by-step fashion to its conclusions.

The structure of the dissertation will be described in the next section and illustrated in figure #1.

1.2 PROJECT OUTLINE

Chapter 1 - Introduction: This first chapter contains this introduction, which defines the project, its scope, structure, methodology, and objectives.

Chapter 2 - **On the importance of cyber attack analysis:** The project will first define a fundamental set of information security-related concepts to create a shared and precise lexicon, a common conceptual ground on which to build the rest of the dissertation.

Afterwards, the second chapter will broadly explain why analysing intrusions and attacks from advanced threat actors is a critical activity, part of other fundamental information security processes such as cyber threat intelligence (hereinafter "CTI") enrichment, cyber attribution, adversarial simulation, security management, and decision making.

Chapter 3 - The three intrusion analysis models: The project will introduce Lockheed Martin's Cyber Kill Chain model, the Diamond Model, and the MITRE ATT&CK framework. Each of these will be presented, described, summarised, and, most importantly, analysed, highlighting its characteristics and limits, if any.
Models cannot be compared without being thoroughly understood.

Chapter 4 - Model comparison: In this chapter the project will propose a method to compare the three models, based on their purpose, abstraction levels, and design. Afterwards, the dissertation will examine how they can be used in a complementary way, using a real-world scenario.

Chapter 5 - Conclusions: The project will end with a conclusion chapter, summarising the results of the analysis.

Reference list: The dissertation contains a full reference list using the IEEE referencing standard.

Appendices: Four appendices have been added to the document. Each appendix is a referenced and verbatim excerpt from third-party work, included in the document to enhance readability by preventing the readers from having to jump back and forth between the dissertation and external sources.

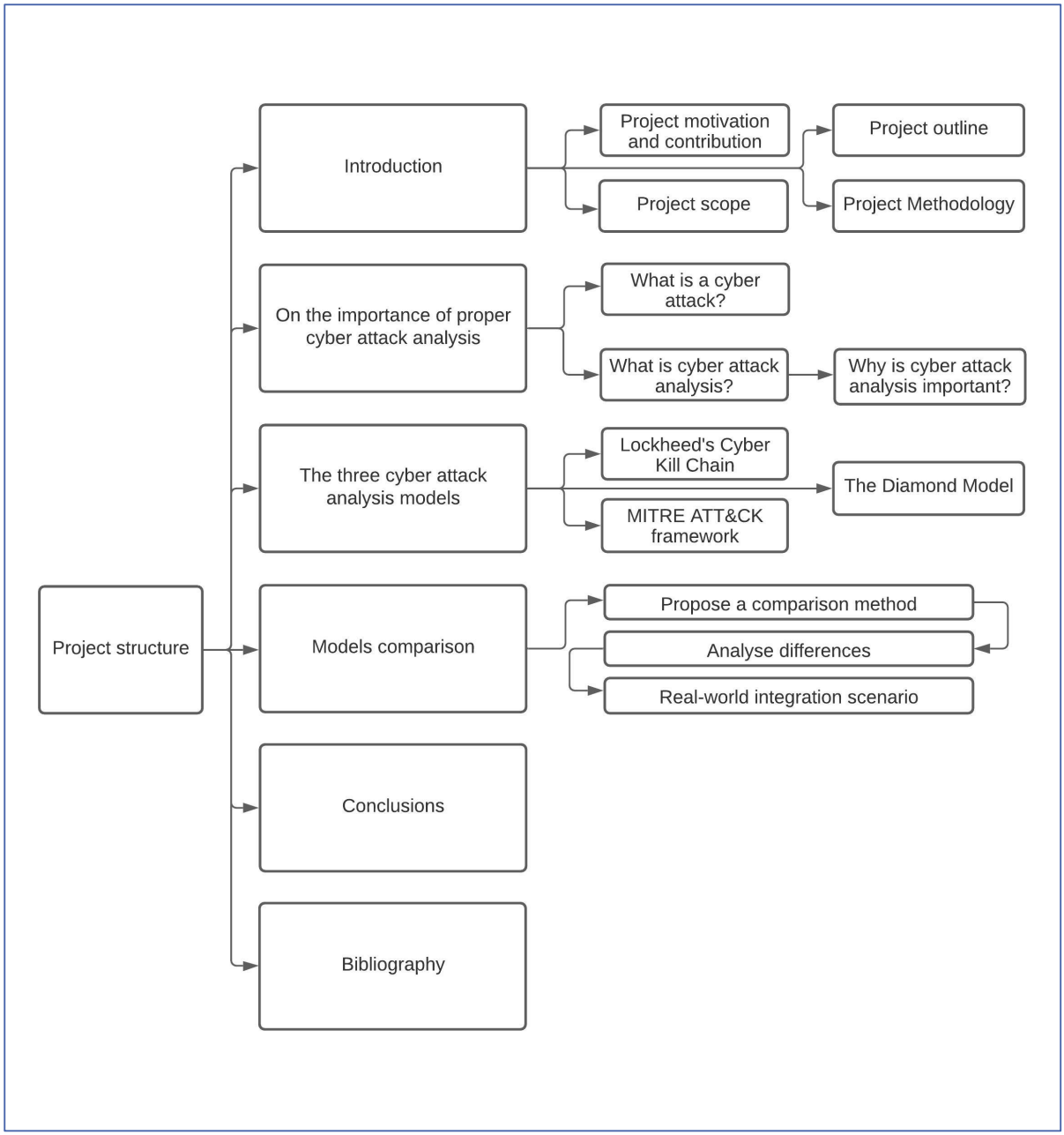

Figure 1 - Project Structure

1.2.1 DOCUMENT STYLING, FORMATTING, AND REFERENCING

Calibri font is used across the whole paper, the only exception will be code snippets, a monospace font is used for them to enhance readability.
The same fonts and colour palette are also used for all images and tables.

The following formatting rules apply to the paper:

Bold:	Used to highlight pivotal parts of the project, foundational to the full understanding of the dissertation.
Italics:	Used to highlight special terms the first time they are introduced
Underline:	Used for cross-referencing throughout the paper. Figures, tables will appear in bold.

IEEE referencing style has been chosen[6]
Short direct quotes will be enclosed in double quotation marks.
Quotes longer than four lines will be written in a monospace font and indented.
Citation numbers within square brackets will be used next to every quote, the full corresponding reference will be listed at the end of the document.
It will contain Author(s), paper/book/website title. DOI, website URL, and website last access date will be added where available.

1.3 PROJECT SCOPE

The scope of the project is limited to examining and comparing the three leading and industry standard intrusion analysis models.

This is a complex task that requires tapping into many areas of expertise within the information security knowledge domain and, often, even into entirely different domains, such as those of intelligence analysis, cognitive sciences, international relations, economics, and law.

Due to time and resource constraints, the project will not explain all the terms and concepts used in minute detail.
For this very reason, a vast reference list of external resources will be provided across the dissertation, using the IEEE referencing standard, and the author strongly advises consulting them.

The project will mainly refer to cyber attacks by advanced and persistent actors. Script kiddies[4], hacktivists[5] and common cybercriminals aren't the main concern for cyber intrusion analysis models.

1.4 PROJECT METHODOLOGY

This project mostly relies on *literature research* and *analysis*.

Different literature sources have been used.

In order to better understand and analyse the three models, the design whitepapers that originated them have been scrutinized.

Papers and reports from security vendors have been leveraged to obtain updated hard data on adversarial behaviour and activity in the current threat landscape.
These documents are often filled with more or less explicit advertisements for the vendors' products, but the dissertation will only take into consideration hard data and intelligence corroborated by evidence and reproduced by peer, highly reputable vendors.

Publications from ISO, NIST, and analogous institutions and bodies will be referenced, especially in section 2.1.1, when trying to define a few basic information security concepts.

Conference papers and websites of renowned security researchers have also been used.

1.4.1 METHODOLOGIC DIFFICULTIES

Academic publications that examine and compare these models are extremely scarce, almost non-existent.

Another challenge encountered while working on this project is that, while security vendors and researchers often publish the results of their research activity, they do not disclose in detail which model(s) they used for it, why, and with what benefits.

Overcoming this lack of literature required an additional analytic effort.

2 ON THE IMPORTANCE OF PROPER CYBER ATTACK ANALYSIS

2.1 WHAT IS A CYBER ATTACK?

The United States Committee on National Security Systems in 2015 defined a cyber attack as: “Any kind of malicious activity that attempts to collect, disrupt, deny, degrade, or destroy information system resources or the information itself”[7].

This definition encompasses both successful attacks and attempts - and will be used by this project - when discussing “cyber attacks”. Failed or thwarted attempts are still cyber attacks in every practical sense.

Cyber attacks can target networks, systems, web-applications, operational technologies (SCADA/ICS/PLCs)[8], and people [9] and they happen all the time, every minute of every day. They can range from the simple that uses commodity tools and is easily detectable to the complex, using custom tools and extremely hard to identify.

Before proceeding further, it is necessary to define some of the most important and frequent information security concepts that will be used in the following chapters.

Please note that high-level definitions will be given, for the scope of this project is not to analyse these concepts in detail, they will serve as a common ground to discuss cyber attack analysis models.

2.1.1 BASIC DEFINITIONS

As per above, in this section, the project will illustrate some core concepts to establish a well-defined lexicon to better understand the rest of the paper.

Some definitions will be verbatim excerpts or quotations from either NIST or ISO/IEC publications.
The majority will be references to the ISO 27000 series, which is a series of best practice recommendations and standards for the design and implementation of an information security management system (hereinafter ISMS)[10].

Information Security

Information security refers to the efforts, methodologies, and processes aimed at protecting any form of private, confidential, and sensitive information from unauthorised access, misuse, use, disclosure, destruction, modification, or disruption[11].

Assurance of the CIA triad or the Parkerian hexad is also an often seen, yet somewhat outdated, definition of information security[12].
Information security can be goal-based or it can be threat-based.
The former aims to achieve specific information assurance objectives for its assets.
e.g., ‘The availability of our web application must be guaranteed’.

The latter is focused on countering the specific threats that would harm the assets of an organisation.
e.g., 'We must protect our web application against DDoS attacks'.

Vulnerabilities and Controls

ISO 27001 defines a vulnerability as "a weakness of an asset or control that could potentially be exploited by one or more threats"[10].

A control is a safeguard put in place in an attempt to tackle the risk related to a specific asset and/or vulnerability.

Threat

A threat is a potential event that once realised, can adversely affect one or more assets. The adverse effect can range from the trivial to the catastrophic, threats can be intentional or unwanted.

A threat can have natural causes (e.g., hurricanes, floods) or it can have anthropic ones. The latter can be due to malicious intent but also to incompetence or negligent behaviour.

The project discusses cyber attacks and will thus only consider the former kind of threats, the malicious, hostile ones.

Threat actor

A threat actor is a person, group, or entity behind an event that impacts or can potentially impact an organisation's asset. Some of them can be particularly advanced, as illustrated in the below chapter, fully dedicated to these actors.

2.1.2 APT

A methodological foreword on this section is necessary.

This dissertation will mostly be using data provided by FireEye and Mandiant, the industry leaders in these kinds of reports.
While other security vendors and firms, such as Crowdstrike, Kaspersky, and Cisco routinely produce similar cyber intelligence products, FireEye and Mandiant reports will be mostly referenced due to being the most detailed and the most frequently updated. See section 1.4 for additional information on this.

2.1.2.1 DEFINITION OF APT

APT is the acronym of *Advanced Persistent Threat*, the term has been long used in the telecommunication industry, but was first used with this specific meaning by the US Air Force in 2006[13].

An APT is an extremely motivated and stealthy threat actor with a definite mission.

It is an **advanced** actor because it has extremely high technical expertise and is often able to craft and develop custom tools or to heavily modify commodity ones.
E.g., NSA is the very definition of an APT for its adversaries.
NSA created what is currently seen as one of the most advanced malware and hacking toolkits out in the wild, Regin (also known as QWERTY or Prax)[14][15].
Another example of an advanced, non-commodity tool crafted by an APT would be Stuxnet.
While its attribution is far from certain, it is well known to be one of the costliest malware to ever be developed, requiring programmers with expertise ranging from Windows systems to PLCs.

It is a **persistent** actor because it has a mission to accomplish, it takes orders and directives, and it will take all the time necessary to achieve its goals.
Unlike most threat actors APTs do not have a predatory behaviour looking for immediate gains (financial or socio-political[5]) but tend to stay silent over long periods of time, moving across their target networks and systems, collecting as much information as possible. Their typical objective is - but is not limited to - exfiltrating critical and sensitive data from many industry sectors, military and civil alike. Espionage.
E.g. APT1 (also known as Comment Panda, Comment Crew, Byzantine Candor) is a threat group that specialises in stealing commercial, industrial, and business data and intellectual property from US enterprises (or US-based multinationals)[16][17].
This APT started its activity back in 2006 and, as of 2020, is still operating.
The average dwell time in the networks and systems of the victims was one year, with a peak of 1,764 days in a specific case.
This malicious group perfectly exemplifies the notion of persistency in APTs.
More details on the notion of persistence can be found in section 2.1.2.3 below on *dwell time*.

It is a **threat** because it is well funded, organised, and motivated.

Based on the targets, the tools used, and the resources involved in these kinds of operations, it appears clear that most APTs belong to nation states or state-sponsored groups[18][19].

It is worth mentioning that in recent years a few financially motivated actors started behaving like APTs, with persistent, advanced, and targeted behaviours.
FireEye and Mandiant analysts and researchers labelled these *FIN groups*, as financially motivated threats, with the same capabilities of APTs, but with different missions and stakeholders[20].
The same two companies have identified over 1,800 threat groups, with hundreds of new groups identified in 2019 alone.

Of those, 23 groups were active advanced threat groups (APT + FIN).
The picture below illustrates the threat group landscape [21].

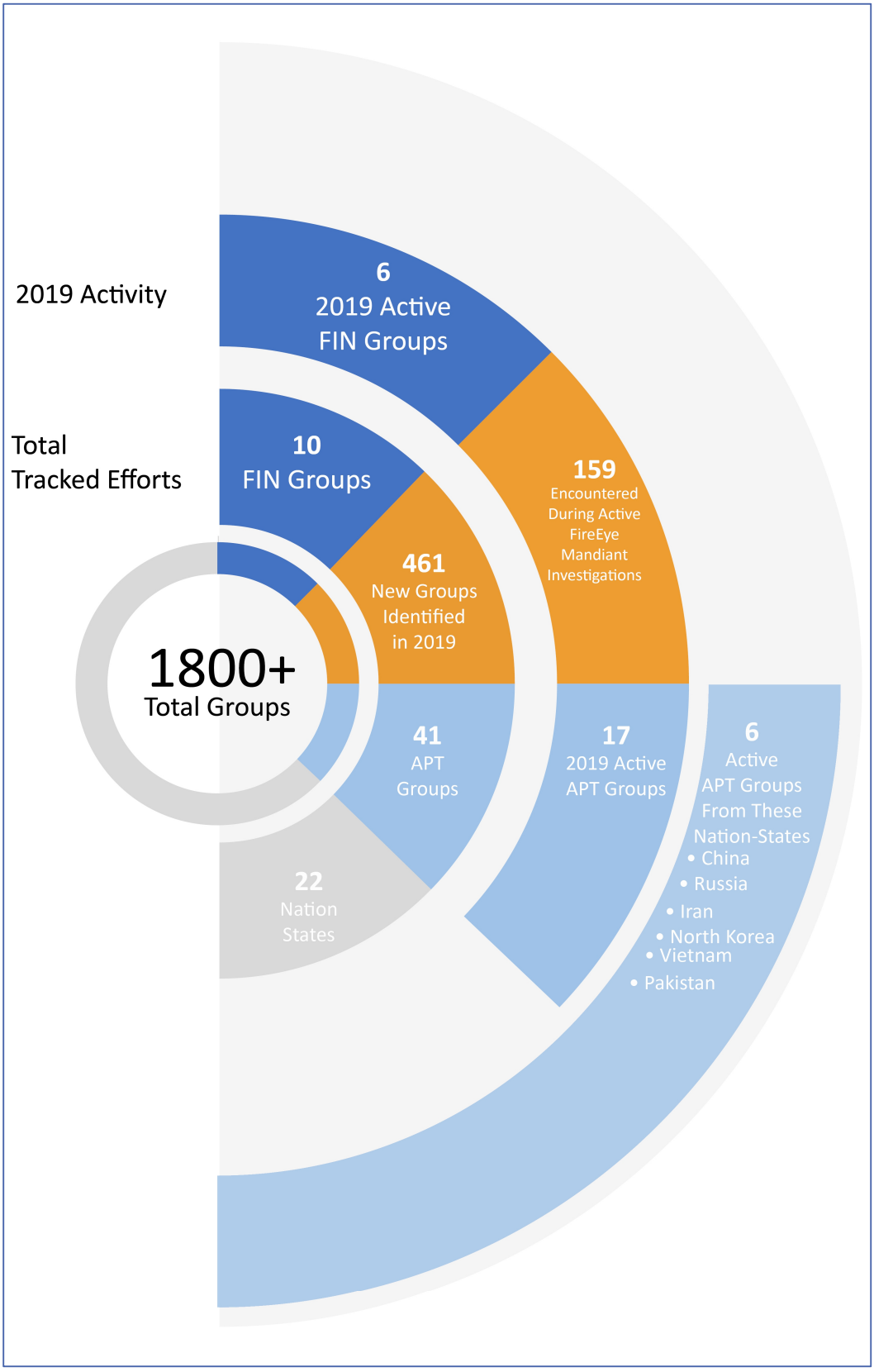

Figure 2 - Active threat groups

2.1.2.2 INDUSTRY SECTORS TARGETED

APTs have high profile targets ranging over many different industry sectors, from the DPRK-tied Lazarus group behind the SONY hack of 2014 [22] to the Russian Fancy Bear (APT 28) with its campaigns aimed at the US government and at interfering with elections [23].

Many organisations and security vendors provide periodic reports on APT campaign targeting.

In 2019, as per figure 3, the top ten most APT-targeted industries were[21]:

01. Entertainment
02. Financial
03. Government
04. Business
05. Construction
06. High Tech
07. Telecom
08. Health
09. Energy
10. Logistics

While this targeting nexus hasn't changed much over the past five years[21], the COVID-19 pandemic of 2020 is likely to change this, with many APT groups targeting healthcare, medical research, manufacturing, and logistics[24] [25] [26].

Figure 3 - APT targeting per industry sector

2.1.2.3 DWELL TIME

APT activity is persistent and advanced, one of the most important metrics is its *dwell time*.

The dwell time is the time an intruder has been inside the victim's network before detection.

It is a pivotal metric to observe as an undetected APT group will relentlessly work towards achieving its mission, moving laterally, exfiltrating data, and/or tampering with systems without any active response from the defender.

The persistence of APTs is one of the main reasons for their perilousness and dwell time is a seminal metric to gauge that.

The median dwell time in 2019 was 56 days, FireEye reports that 12% of their clients had a dwell time higher than 2 years (700 days)[21].

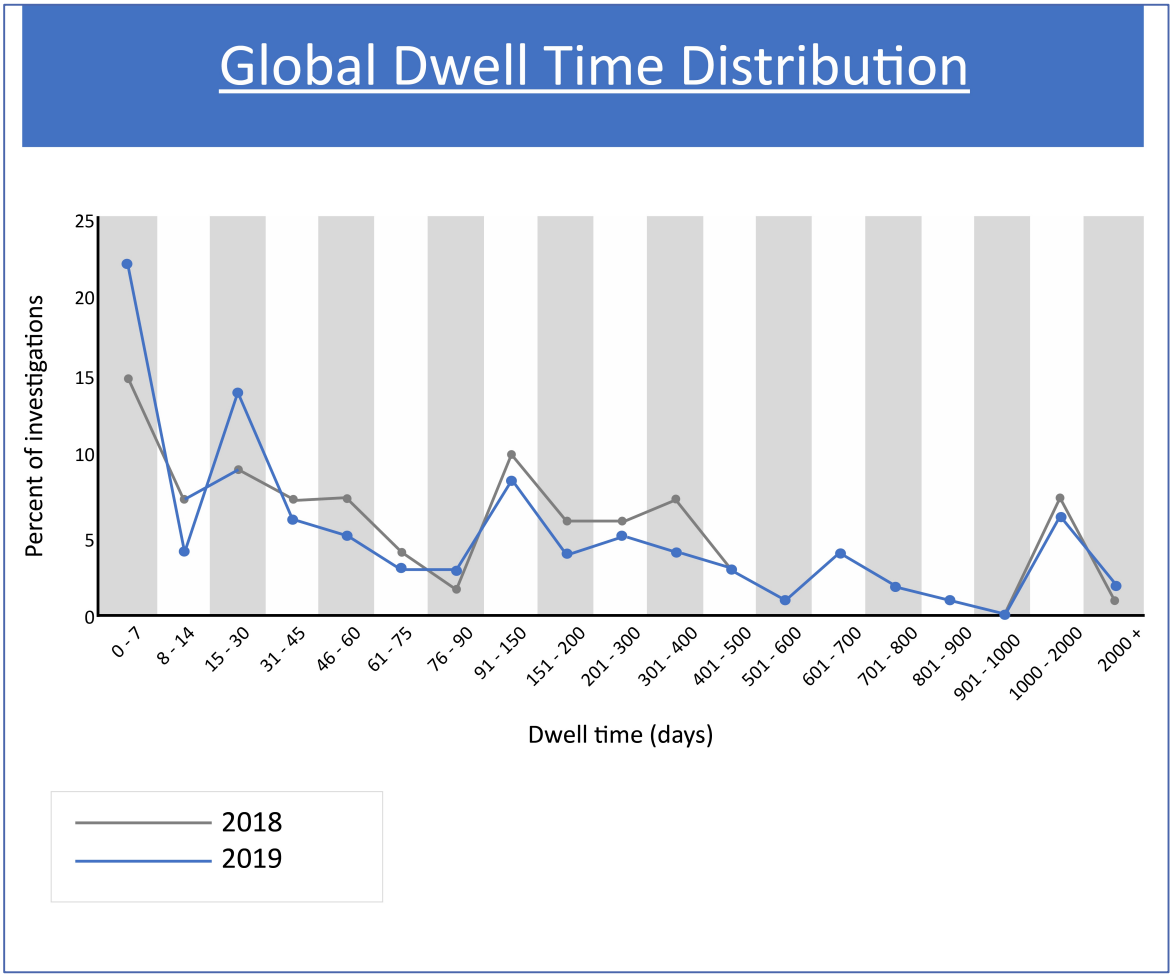

Figure 4 - Global dwell time distribution of APTs

Cyber attacks by advanced threat actors, targeting high profile objectives and being highly stealthy and difficult to tackle, are the most important to analyse and the very reason so many different analysis models and frameworks have been proposed in recent years.

2.2 CYBER ATTACK ANALYSIS

2.2.1 DEFINITION

The previous sections showed, in basic terms, how difficult it is to tackle cyber attacks from advanced actors.

Consequently, they need to be thoroughly analysed.

Cyber attack analysis is the process of understanding via the use of specific models, tools, and techniques, how and why an intrusion has occurred.

It is a process used to break down an entire attack into smaller and easier to understand activities, in order to generate actionable intelligence to be fed to other information security processes in an effort to sharpen the security posture of organisations.

During the project "cyber attack analysis" and "cyber intrusion analysis" will be used interchangeably, while there is a slight difference between the two, it is not appreciable nor useful within the scope of the project.

2.2.2 WHY IS CYBER ATTACK ANALYSIS IMPORTANT?

As discussed in section 2.1.2 cyber attacks by advanced actors have high-profile targets and the technological expertise and resources to cause remarkable damage to a plethora of organisations: enterprises, corporations, and even nation-states[27].

The economical, geopolitical, industrial, commercial, and military impact that successful intrusions by APTs can have is dramatic, hence the need to have models to accurately analyse them.

An accurate analysis of these intrusions using dedicated models will result in better protection. This because the intelligence generated can be used to perform several processes, such as those listed below.

This is not pretending to be an exhaustive list, the elements in it have been chosen as they are all directly referenced by one or more of the models examined.

2.2.2.1 CYBER THREAT INTELLIGENCE (CTI) ENRICHMENT

CTI represents the knowledge organisations have of the cyber threat landscape.

Knowledge of the behaviours, tactics, techniques, procedures, intents, and resources of potential adversaries.

The models presented, analysed, and compared in this project can all be used to enrich the cyber threat intelligence of an organisation.
They can be leveraged, in a CTI context, by defenders and analysts to better describe and

visualise attack patterns, identify similarities, and generate and test new hypotheses on the adversarial nexus.

An enhanced CTI allows organisations to perform better adversarial emulation, attribution, and defensive assessments.

2.2.2.2 ADVERSARY EMULATION

An organisation, using the models discussed in this project, can better understand how specific adversaries behave and act.

This knowledge of the attackers is then used to run adversarial emulation scenarios.

These are performed by organisations to test their security against a red team that will behave like a specific known adversary.

2.2.2.3 DEFENSIVE GAP ASSESSMENTS

The correct use of cyber attack analysis models allows organisations to better discern the attack paths and patterns more often used against them.

Such knowledge can be leveraged to perform better defensive gap assessments.

Defensive gap assessments are systematic processes that allow defenders to understand what parts of their technological domain lack controls, safeguards, or countermeasures, hence identifying and fixing any weak points and blind spots.

A similar process can be applied to assess the performance and maturity of a SOC[28].

2.2.2.4 CYBER ATTRIBUTION

Finally, analysing intrusions plays a critical role in the extremely complex activity of cyber attribution.

“Attribution describes the process of assigning a particular act to its source not necessarily in the sense of its physical perpetrator but more importantly in the sense of its mastermind. Attribution is important because it forms the basis of appropriate and effective technical, political, and legal determinations and underpins technical, political and legal action, and responsibility[29].”

Cyber attribution poses many challenges, it is a dramatically difficult task, requiring expertise from multiple disciplines and fields and has serious repercussions in the real world.

This topic is both extremely interesting and complex, its detailed analysis is beyond the scope of this dissertation, however, a brief explanation of how it works and why it is needed is required, as all of the examined models are related to, and can help with, attribution.

2.2.2.4.1 WHY IS ATTRIBUTION NEEDED?

Identifying culprits is a fundamental need for humans, there is a strong psychological need to attribute causes to events[30][31]. It is unsatisfactory to think that incidents happen by pure chance.
However, satisfying this need is not enough to justify the costs of cyber attribution.

Defining the identity of an attacker brings several advantages.

In the resource-constrained environment of information security, where diminishing marginal utility is an inescapable reality, understanding the enemy and its objectives allows a more precise definition of defensive requirements, and thus expenditure, by prioritising specific technologies and targets.

Additionally, attribution serves as a means to exert diplomatic pressure.

The US-China cyber agreement of 2015 is a perfect example of this[32].
In May 2014 a court in Pennsylvania condemned several officers of the PLA for cyber espionage with extremely strong evidence[33].
This cyber attribution effort and the resulting threats of sanctions[34] towards China for its cyber intrusions were used to pressure Xi Jinping's government to accept the cyber agreement. Before that, the Chinese government always dismissed any accusation and accord due to the lack of evidence (and hence attribution).

Lastly, if threat actors had the ability of plausible deniability, the lack of deterrence wouldn't prevent them from committing the most heinous actions.

Attribution allows deterrence, both via pre-emptive or retaliative action, be it political, commercial, technological, or military.

2.2.2.4.2 HOW DOES ATTRIBUTION WORK?

Attribution (especially when APTs are involved) is an instance of abductive reasoning based on a vast set of premises[35] such as, but not limited to, these:

- APT activity is usually carried out by nation states or contractors working for them. See section 2.1.2.
- Attacks done by the same actor will have some degree of similarity. See section 3.1.4.2.
- The more shared code parts with specific functionalities that two pieces of malware have in common, the more chances there is a common developer.

Dr Timo Steffens, a scholar involved in the analysis of many high profile malicious activities in Germany and an author often referenced on this topic, in his research[35] divides the attribution process into six phases:

1. *Data collection* - This is a routine procedure for all researchers and security vendors; collecting as much data as possible about exploits, malware, IOCs, compromised domains, malicious IP ranges, and any other important information on hostile activity.
2. *Clustering* - Similarities in collected data are used to cluster activities that share some common traits into intrusion sets. This process can be automated with machine learning, manually done by human analysts, or a combination of the two[36].
3. *State-sponsored vs criminal activity* - Analysts in this phase must understand whether the campaign being observed belongs to state-sponsored or criminal activity. Targets, types of data exfiltrated, infrastructures involved, and scale of operations are typically good metrics to use for this. e.g. A campaign aimed at stealing non-easily monetisable information or a campaign requiring hundreds of specialists will hardly be performed by common cybercriminals.
4. *Attribution to a specific country* - Based on many indicators analysts can guess, with various degrees of confidence, where an attack originated. The target of the malicious activity might be of special geopolitical or economic interest for a specific nation state. The original IP addresses that started an attack might belong to a specific area. The hours in which an attack occurred might indicate the office hours of attackers. Language in the code might indicate the country of origin.
5. *Attribution to specific organisations or persons* - This is by far the most detailed attribution step achievable, defining the very group or person behind a campaign. It seldomly succeeds, and even then it is hardly legally binding as it's more a set of circumstantial evidence and technically defined parallels. Not enough to overcome the *in dubio pro reo*, which clearly also applies to cyber attribution when it is brought to court[37][38].
6. *Confidence check and communication* - The findings of all previous five phases are given a confidence level, often using estimative language[39] mutated from intelligence agencies, and then communicated to the relevant stakeholders and decision maker.

Now that a shared lexicon has been established and the importance of using cyber intrusion analysis models has been delineated, the dissertation will start examining the three models separately.

3 THE THREE INTRUSION ANALYSIS MODELS

This chapter will illustrate the three seminal cyber attack analysis models, Lockheed Martin's Cyber Kill Chain, the MITRE ATT&CK framework, the Diamond Model.

A brief introduction on the context that generated them will be given, then their logic and workflow will be analysed.

As mentioned in section 1.3 these three models have been chosen as they are the most used, referenced, and respected.

3.1 LOCKHEED'S CYBER KILL CHAIN

3.1.1 INTRODUCTION

In 2011 the US DoD officially recognised cyberspace as the fifth domain of warfare, in addition to sea, air, land, and space[40].

Coincidentally, in that very year, Lockheed Martin Corporation released a whitepaper named '*Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains*', which effectively explained how to apply some common military concepts to this new domain.

The whitepaper proposed a new approach to the analysis of cyber attacks.

The authors started from the premise that in those days conventional incident response methods failed to protect from advanced threats, such as APTs, due to two major, faulty, assumptions.

The first one being that response should happen after a compromise, the second one that the compromise was the result of a fixable flaw[41].

The proposed solution was an intelligence-driven computer network defence, built on the analysis of attacks and intrusions, conducted by applying a Cyber Kill Chain model.

It is necessary to explain what a kill chain is, before exploring Lockheed's Cyber Kill Chain.

3.1.2 WHAT IS A KILL CHAIN?

"A kill chain is a systematic process to target and engage an adversary to create desired effects[42]."

The authors of the paper take the US military doctrine kill chain as a reference, and it is represented by the acronym F2T2EA[43].

The acronym stands for the following steps:

- Find: Identifying a target through means of intelligence, surveillance, reconnaissance.
- Fix: Obtaining a precise location of the target, such as coordinates.
- Track: Continuously monitoring the target until an engagement decision has been made.
- Target: Applying C2 to assess the value of the target and the availability of the means and weapons to engage it.
- Engage: Using the chosen asset or weapon against the target.
- Assess: Evaluating the outcome and effects of the attack.

It is fundamental to understand that, in the military doctrine, a kill chain:

- Describes the steps of an attack, sequentially.
- Represents an end-to-end integrated process where a deficiency at any step will nullify the entire process.
- That a defender can place controls and safeguards at any level of the kill chain to thwart and disrupt the attacker's efforts.

The Cyber Kill Chain model proposed in the Lockheed Martin whitepaper tries to adapt the more kinetic warfare oriented F2T2EA paradigm to the cyber domain.

3.1.3 CYBER KILL CHAIN PHASES

The LM whitepaper expands on the traditional kill chain concept to create a Cyber Kill Chain model to be used to analyse cyber attacks.

The proposed model is made up of seven kill chain phases:

- *Reconnaissance* - This phase is where the attacker researches, identifies, and gathers information on the target. Information can be obtained by any means, active and passive. e.g., OSINT tools such as Maltego/Shodan, SIGINT, HUMINT...
- *Weaponisation* - In this phase the attacker produces a weaponised payload, via a weaponiser or similar tool. Usually, a backdoor trojan and an exploit are coupled and embedded in a weaponised deliverable (such as a PDF file, or a similar document).
- *Delivery* - In this phase either the attacker or the target initiates and conducts the transmission of the weaponised deliverable. The former most often happens via SQL injections or compromised network services, the latter, far more frequent, usually happens via drive-by-download, watering hole attacks, USB removable media.
- *Exploitation* - The deliverable now resides on the victim's machine, the exploit contained in it will compromise the machine, usually leveraging vulnerabilities at the application or system level.

- *Installation* - In this phase, a backdoor is installed in the compromised machine to guarantee future and persistent access to the attacker.
- *Command and Control (C2)* - C2 is another military term, it represents the exercise of authority and direction by a properly designated commander over assigned and attached forces in the accomplishment of the mission[44].
  In this context, it means that the compromised machine will open an outbound channel to an external, adversary-owned server that will be used by the malicious actor to impart orders to the installed malware. These channels often also grant "hands-on-keyboard" access to the victim system. e.g. Bind shells, reverse-shells[45].
- *Actions on Objectives* - In this phase the attacker, having obtained full and permanent access to the machine, will start to achieve their original goals, most likely by collecting, encrypting and thus exfiltrating data, but also by pivoting on the machine to start a lateral movement in the network.

The project will now elaborate on how the CKC can help an organisation understand a threat and defend against it.

3.1.4 USES

LM's CKC model is part of the broader "Intelligence-Driven Computer Network Defense" approach.

The model becomes actionable intelligence when defenders base their security decisions and measurements on a thorough understanding of the adversary kill chain.

This understanding allows them to better identify defence gaps, rectify them, and more carefully plan security investments.

3.1.4.1 ACTIONABLE INTELLIGENCE

Hutchins, Cloppert, and Hamin in the LM whitepaper also propose a basic template to counter adversarial actions using the CKC.

It is based on using the actions of *detect*, *deny*, *disrupt*, *degrade*, *deceive*, and *destroy* in each and every phase of the adversary CKC. The above actions are those suggested by the US DoD Information Operations Doctrine[46].

Table 1 is a sample matrix illustrating the wide range of countermeasures (actions) a defender can deploy at each level of the kill chain, it is not intended to be a comprehensive list. Ø represents within a cell represents an empty set.

Phase	Detect	Deny	Disrupt	Degrade	Deceive	Destroy
Reconnaissance	Website logs, web analytics	Firewalls	Ø	Tarpits	Ø	Ø
Weaponisation	NIDS, HIDS	NIPS, HIPS	Ø	Ø	Ø	Ø
Delivery	Aware user	Proxy Filter	In-line AV, Firewalls	Queueing	Ø	Ø
Exploit	HIDS	Application and system patching	ASLR,DEP, Canaries	Ø	Ø	Ø
Installation	HIDS	Jails such as chroot ones	AV	Ø	Ø	Ø
C2	NIDS	Firewalls	NIPS	Tarpits	DNS Redirects	Ø
Actions on objectives	Log auditing	Ø	Ø	Ø	Honeypot	Ø

Table 1 - Cyber Kill Chain Matrix | Resilience

The more complete the above matrix, the more the defender is resilient.

Resiliency is the primary goal against advanced and persistent attackers, although full coverage of the matrix is not feasible in terms of costs and benefits. No defender can plan and implement detection, denial, disruption, degradation, deceit, and destruction for every kill chain phase. For this very reason, some empty cells (Ø), as seen in table #1, are expected.

It is extremely important to stress that the CKC is used, within this approach, to produce actionable, intelligence-driven, defence across the entirety of the chain.

It is worth noting that under this intelligence-driven perspective if an attacker with a known CKC adopts a single new technology or action, such as exploiting an unknown zero-day vulnerability, this will not be a major issue for resilient defenders, for they will be able to disrupt the attack at the other, well known, and already managed phases of the chain.

This is an extremely important concept to understand, as it shows how focusing on specific threats, such as zero-days, is a myopic attitude compared to the broader kill chain approach.

The project has shown how the CKC model can become actionable intelligence and now will delve into how it can be used to reconstruct intrusions.

3.1.4.2 INTRUSION RECONSTRUCTION

In the days when the LM's CKC whitepaper was released, it was common for incident response processes to start only after the *exploit* phase of the kill chain. This was relegating defenders to a disadvantageous and tardive position. See *section 3.1.1*.

As shown in the previous chapter, being able to respond to every adversarial attack phase is the very definition of a defender's resilience.

Hence, one of the main purposes of LM's CKC was to give response teams a model to analyse and reconstruct adversarial actions before the time of detection.

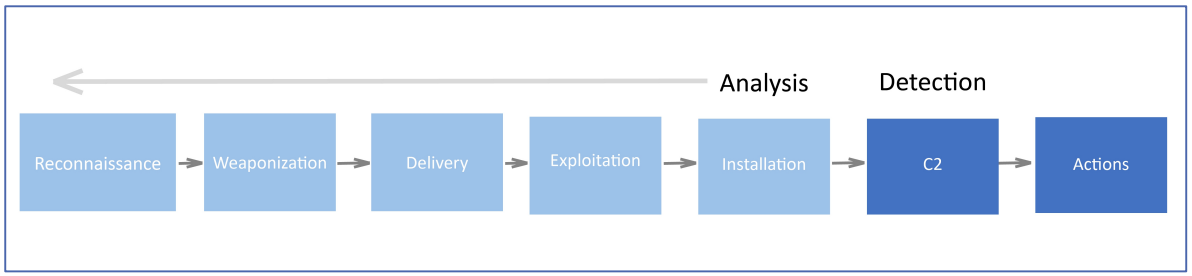

Figure 5 - Cyber Kill Chain phase reconstruction

Cataloguing the attacker's action in kill chain phases is necessary to put in place controls, safeguards and countermeasures at every stage to cope with future attacks.

Advanced persistent threats re-use tools and infrastructures to make their attacks economical, they must effectively make use of economies of scale.

If defenders can fully reconstruct intrusion steps, they can leverage this intelligence to force adversaries into changing parts of their CKC.
This means that defenders can effectively reach a new level of resilience by harming the adversarial - necessary - economies of scale.

The repeated use of tools and procedures to keep costs low becomes a liability for the persistent attacker.

3.1.4.3 CAMPAIGN ANALYSIS

The CKC model also has a remarkable use at the strategic level, in what the authors define as *campaign analysis*.

The idea is that by analysing the kill chain of multiple intrusions and identifying commonalities amongst them, defenders can define intrusion campaigns, achieving two major objectives.

Figure 6 illustrates this process, it shows how two intrusions can be clustered into a single attack campaign, due to the numerous similarities in their *Indicators of Compromise* (IoC).

e.g., The hash values of the exploit are the same, the same C2 protocol is used, the installed files have the same names, et cetera.

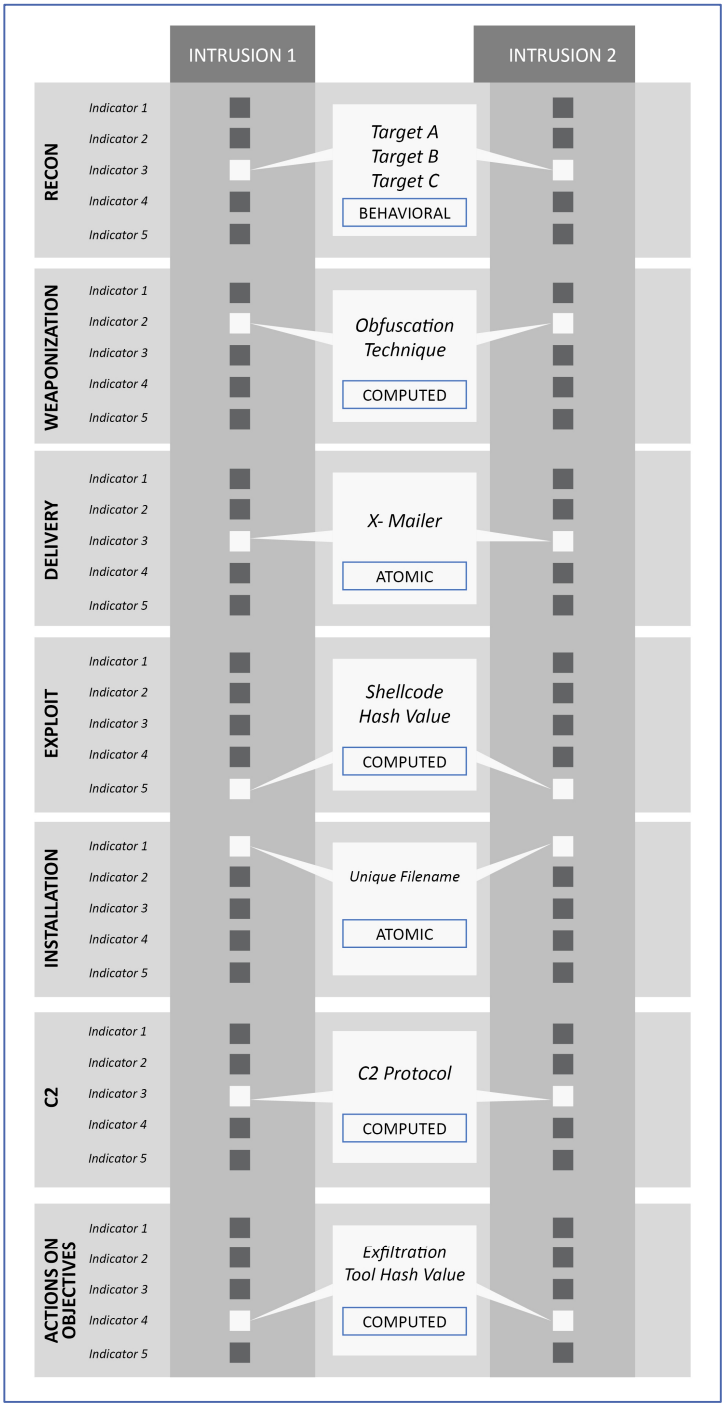

Figure 6 - Cyber Kill Chain campaign analysis

The primary objective of campaign analysis is to understand how attackers operate by defining their sets of techniques, tactics, and procedures. These are defined *intrusion sets*. Understanding how the threat actor behind a campaign conducts its operations provides useful intelligence that can be used to strengthen defence gaps, allowing defenders to better assess their security posture on a per campaign basis.

The key deliverable here is to understand the capabilities, modus operandi, limitations, and the doctrine of the threat actor behind a campaign, not identifying it.
Attribution is not the objective of LM's CKC-based campaign analysis but can be a useful side product of it.

The second objective of campaign analysis is to understand the attacker's *intent*.
Achieving this allows defenders to further sharpen their security posture in the light of more accurate threat modelling and threat landscaping.
The idea is that knowing the technologies or the individuals of interest for the attacker will make defence easier.
It is worth noting that threat actors and intents are well separated notions. A defender can understand and defend against adversarial intent without knowing who the attacker is, and vice versa.

3.1.5 KEY POINTS AND ADDITIONAL CONSIDERATIONS

Lockheed Martin's Cyber Kill Chain model is an adaptation of the well-established military doctrine concept of a kill chain, as seen in section 3.1.1.

It is proposed by the authors as a tool to analyse intrusions and attacks to produce the actionable operative, tactical, and strategic intelligence necessary to achieve intelligence-driven cyber security.
Only by carefully mapping adversarial attacks to Cyber Kill Chain phases and by identifying intrusion patterns, can a defender effectively refine his security posture, measure the defensive performance, and plan future investments (see section 3.1.4)**.**

LM's approach has received some criticism in recent years from security practitioners and academics, it is, in fact, the most criticised of the three models.
The criticism stems from the CKC being heavily influenced by outdated, perimeter-focused paradigms and being too rigid and sequential[47].
According to the critics, this causes several problems.

The first problem is that the CKC is hardly capable of describing and providing actionable intelligence on insider attacks.
Trying to use it to analyse and reconstruct an attack started by an insider threat can prove to be problematic, if doable at all.
If the culprit is an insider, a rigid and perimeter-focused kill chain cannot be effective, as the threat is already within the perimeter.

The FBI and the US DOJ provided a variation on the Cyber Kill Chain to tackle insider threats, as illustrated in figure 7[48].

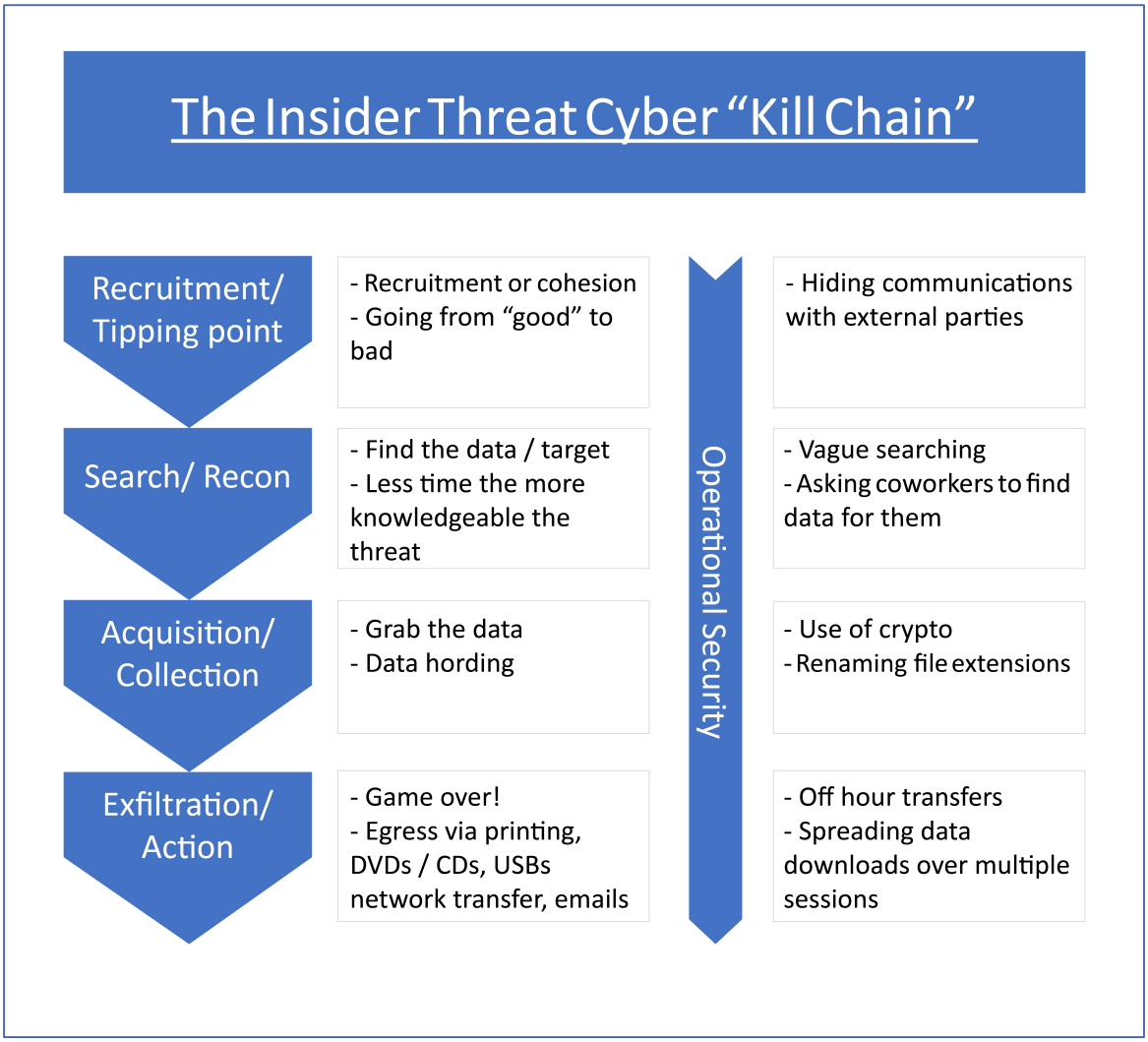

Figure 7 - FBI "Insider" Cyber Kill Chain

The second problem resides in how the phases are structured.
Most of the criticism revolves around them being disproportionate when compared to real, modern attacks.
Phase one to six take relatively little time while phase seven, especially when advanced threats are involved, can last many years, as many as five[49].

The third problem is that the CKC logic is too rigid and sequential, it lacks the flexibility to adapt to modern APTs that often rely heavily on pivoting and lateral movement within networks.
To shorten this gap, security researchers propose an amendment to the CKC to incorporate an additional step, before the final one, dedicated to adversarial lateral movement[50].

These critiques will be expanded and illustrated in greater detail when actively comparing the models, in section 4.2.3.3.

3.2 MITRE ATT&CK

3.2.1 INTRODUCTION

The MITRE ATT&CK framework was born in 2013 as a spin-off of MITRE's Fort Meade eXperiment (FMX).

FMX was a research environment used to emulate both adversarial and defender behaviours in an effort to use telemetry and behavioural analysis to improve post-compromise detection of threats[51].

To measure the effectiveness of the detection of documented adversarial behaviours, a tool to scientifically organise such behaviours was deemed necessary.

This tool is the MITRE ATT&CK framework.

The MITRE ATT&CK framework is meant to be a curated and globally shared knowledge base of adversarial tactics, techniques, and procedures, providing a common taxonomy of both attacks and defences to enable better threat modelling, cyber threat intelligence, adversarial emulation, and red teaming.

At a high-level, ATT&CK is a behavioural model that consists of the following

core components: tactics, techniques, catalogued adversary usage of such techniques and their procedures, and related mitigations[52].

It was originally only created for Windows systems but has since been expanded to Linux, Mac OS, Mobile, Cloud, and even ICS systems.
MITRE created three technology domains to accommodate this change, these domains group different platforms, as shown in table #2.

Technology Domain	Platform(s)
Enterprise	Windows, Linux, Mac Os, Azure AD, AWS, GCP
Mobile	Android, iOS
ICS	SCADA systems, PLCs, HMIs

Table 2 - ATT&CK: Technology Domains

The project will now illustrate the core entities of this model and will then proceed to explain how they are used.

Unless specified otherwise, the project will mostly refer to the MITRE ATT&CK enterprise technology domain, as other domains are less widespread and not as well-known. Most high-level concepts apply to all technology domains anyway.

3.2.2 STRUCTURE AND LOGIC

This section will illustrate the overall layout of the ATT&CK model and explain, at first, its three most important entities: *techniques, tactics, procedures.*

These main entities, as well as other minor ones, are referred to as *objects* in the whitepaper that first introduced this framework.

ATT&CK Matrix

Before illustrating TTPs, it is worth mentioning the ATT&CK Matrix.
It is a visual representation of the relationship between tactics, techniques, and sub-techniques.
There is a Matrix for each technology domain.

Appendix A of the project is the Enterprise Matrix, the author strongly recommends examining it at this point of the project.

Appendix B of the project is the Cloud Matrix, it will be referenced in section 4.2.3.2.

3.2.2.1 TACTICS

Tactics are the tactical achievements an attacker aims for when performing an action. The tactic is the reason behind that action.
In other words, when malicious actors use techniques, the tactics are what they want to achieve.
e.g., *Drive-by downloads* are a technique to reach the tactical objective of *initial access* to a system.

The current tactics in the MITRE ATT&CK for the enterprise domain are:

- *Reconnaissance*: This phase is where the attacker researches, identifies, and gathers information on the target. Information can be obtained by any means, active and passive. e.g., OSINT tools such as Maltego/Shodan, SIGINT, HUMINT...
- *Resource Development*: designing, gathering, and implementing the resources to support operations. e.g., setting up a botnet or a C2 infrastructure.
- *Initial Access*: gaining a foothold in the target network. e.g., drive-by downloads, or phishing.
- *Execution*: running malicious code on the accessed machine. e.g., Trying to get a remote access tool running.
- *Persistence*: making sure that the presence in the network is maintained for future re-use. e.g., changing system configurations.
- *Privilege Escalation*: trying to get higher permissions. e.g., Getting a root or pseudo user on a Linux machine or a System Administrator role in Windows.
- *Defence Evasion*: avoid the defensive tools set in place by the defender. e.g., Camouflaging the malware as legit software.
- *Credential Access*: getting credentials of more accounts and users. e.g., Keylogging, Credential Dumping.
- *Discovery*: mapping and understanding the target environment. e.g., discovering the hosts on the network or the processes running on the victim machine.
- *Lateral Movement*: moving inside the target environment. e.g., Pivoting from a machine to another one inside the same (or a peered) network.
- *Collection*: putting together as much mission-relevant data as possible. e.g., Scraping content of an S3 AWS bucket.
- *Command and Control*: establishing a C2 bridge between the victim systems and the command and control servers. e.g., Communicating with an anonymously named C2 website on the internet via HTTP GET requests.
- *Exfiltration*: transferring data across the target network boundary. E.g. Physical exfiltration (via USB device), or scheduled transfer.
- *Impact*: Damage data, interrupt services, and mission-critical processes on the target system. e.g., Disk wipes, ransomware.

This is an intentionally unordered list as its elements do not necessarily represent the sequential narrative of an attack. See section 4.2.3.3.

3.2.2.2 TECHNIQUES

Where tactics represent "why" an action is performed, techniques represent the "how".

For example, *exploiting a public-facing application* would be a technique used to obtain the tactical goal (the *tactic*) of *initial access*.

Since techniques are just tools for an end and there can be multiple tools to achieve the same end, many techniques can be grouped inside a single tactic category.

Table 3 is a brief example of this: the *lateral movement* tactic can be performed using nine different techniques.

Techniques for the Lateral Movement Tactic
Exploitation of remote services
Internal spearphishing
Lateral tool transfer
Remote server session hijacking
Remote services
Replication through removable media
Software development tools
Taint shared content
Use alternate authentication material

Table 3 - ATT&CK: the techniques for the Lateral Movement Tactic

Each technique can be made up of one or more sub-techniques.

3.2.2.3 SUB-TECHNIQUES

Techniques can be split into more specific behaviours used to obtain a tactical goal; these children behaviours are called sub-techniques.
From a taxonomic perspective, a technique can be the parent of one or more sub-technique.

They were added in 2020 in an attempt by MITRE to solve several abstraction issues that occurred within the model due to its rapid growth.

Some techniques were too broad, some were too detailed, and the overall number was getting out of hand, to the point that it made the ATT&CK Matrix hard to read, visualise, understand, and maintain.

The addition of sub-techniques fixed most of those issues:

- The overall abstraction level of techniques was normalised.
- The number of techniques got reduced to a level that was easier to read, maintain, and update.
- It showed that one technique could be performed, and thus observed and mitigated, in many ways.

Table 3 showed how techniques are grouped under a tactic.

Table 4 will expand it to include all the sub-techniques of the *remote services* technique, to show the full taxonomic logic.

Before that, for easier understanding, this is the MITRE ATT&CK definition of the *remote services* technique:

"Adversaries may use Valid Accounts to log into a service specifically designed to accept remote connections, such as telnet, SSH, and VNC. The adversary may then perform actions as the logged-on user [...][53]."

Techniques for the Lateral Movement Tactic
Exploitation of remote services
Internal spearphishing
Lateral tool transfer
Remote server session hijacking
Remote services
Sub-techniques for the Remote service Technique
Remote Desktop Protocol
SMB/Windows Admin Shares
Distributed Component Object Model
SSH
VNC
Windows Remote Management
Replication through removable media
Software development tools
Taint shared content
Use alternate authentication material

Table 4 - ATT&CK: sub-techniques for the Remote Services Technique

It can be appreciated that leveraging RDP, SMB, SSH, VNC, Windows remote management and Windows COM model are all possible sub-techniques children of the *Remote Services* technique, used to achieve the *Lateral Movement* tactic.

3.2.2.4 PROCEDURES

Procedures represent how specific adversaries implement the different techniques; their modus operandi and their tools are listed.

This concept is more easily explained with an example such as the one below.

To achieve the *C2* tactical goal many techniques and sub-techniques can be used. One of the technique groups relies on using the application layer protocol(s).
Basically, this is the instance where malicious actors try establishing a C2 bridge by blending in the existing and authorised traffic at the application layer.

Different advanced threat groups use different procedures to achieve this technique's tactical goal:

- Cobalt Strike[54] uses P2P over Windows named pipes within the SMB protocol[55].
- Rocke[56] is used to make wget requests to reach the C2 server[57].
- Magic Hound[58] malware families used IRC to establish a C2 channel[59].
- APT18 (also known as Dynamite Panda or Threat Group 0416)[60] is well known for using DNS TXT queries to communicate with its C2 server[61].

The project has now covered the three main elements of the MITRE ATT&CK framework, however, other objects also constitute it, these will be analysed in the following section.

3.2.2.5 OTHER IMPORTANT OBJECTS

In addition to TTPs, the core components of the framework, three more objects will now be briefly illustrated: *groups*, *software*, and *mitigations*.

Groups

When public or private threat intelligence reports document adversarial activities, the threat groups behind these activities are tracked under the Groups object of the framework.

Groups are either threat groups, threat actors, intrusion sets, or malware families performing targeted, advanced, and persistent activity.
e.g., APT29 is a group, but also RYUK is a group.

As seen in section 2.1.2 this means mostly APTs but also a few financially motivated actors.

Software

This object is a collection of all the software that can be used to perform specific techniques to achieve MITRE ATT&CK-documented tactical goals.
It is divided into two macro subgroups: tools and malware.

The former subgroup contains both commercial, commodity, or built-in software that can be used by red, blue, and purple teams, attackers and defenders alike. It might be found by default on most enterprise systems.
e.g., PS Exec, Windows utilities, wget.

The latter subgroup contains software, again commodity or custom, that is intended for malicious purposes only by malicious actors. e.g., FakeInstaller, DroidKungFu.

This division into malicious and non-malicious software is intended and aimed at illustrating how legitimate software can be leveraged for intrusions as much as non-legitimate software can.

Mitigations

This object is a collection of all known controls, safeguards, and countermeasures that can be put in place to prevent techniques (and sub-techniques) from being carried out successfully.

It is completely vendor agnostic, classes of mitigations (concepts) are listed, not specific products.

Currently, the Enterprise technology domain of ATT&CK lists 41 mitigations, here are a few examples, with a short description, to appreciate how high-level and product agnostic they are:

- *Application Developer Guidance*: Developers trained to code defensively to prevent the introduction of vulnerabilities in their applications[62].
- *Network Segmentation*: Isolating critical systems and resources, logically, physically, or both[63]. e.g., DMZs, VPCs
- *User Training*: Security awareness training programs[64].

This ends the description of the objects that make up the MITRE ATT&CK framework.

3.2.2.6 HOW MITRE ATT&CK OBJECTS INTERACT

The high-level objects analysed so far all interact with each other.

Understanding how tactics, techniques, procedures, mitigations, groups, and software correlate with each other is crucial for the correct use of the framework.

Figure 8 will describe the interaction at a general level.

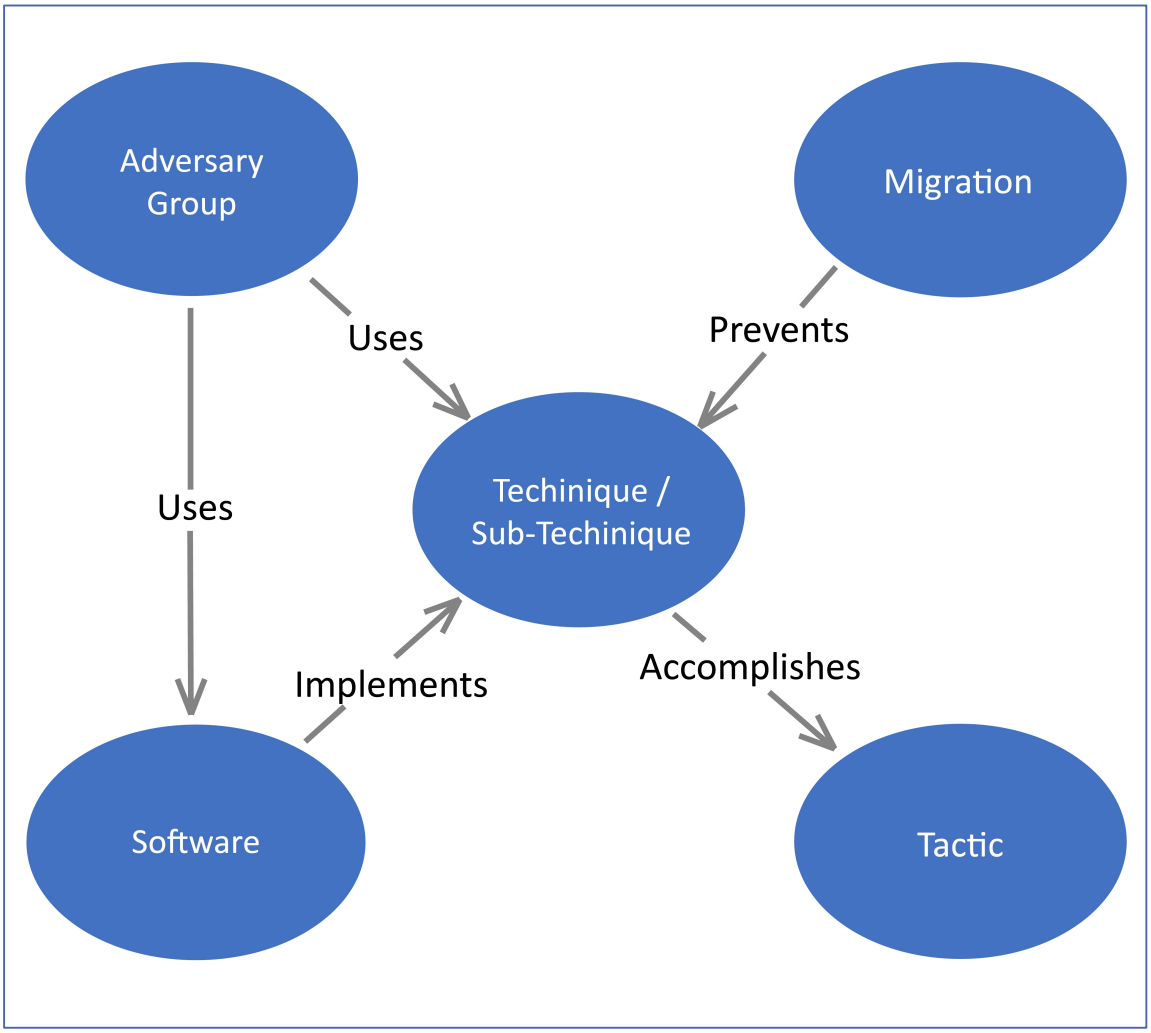

Figure 8 - Generic MITRE ATT&CK object relationships

Figure 9 will make a real-life example of the previous one.

APT 28 (Fancy Bear) behaviour is analysed, showing how they use credential dumping via software (Mimikatz) to achieve credential access[65].

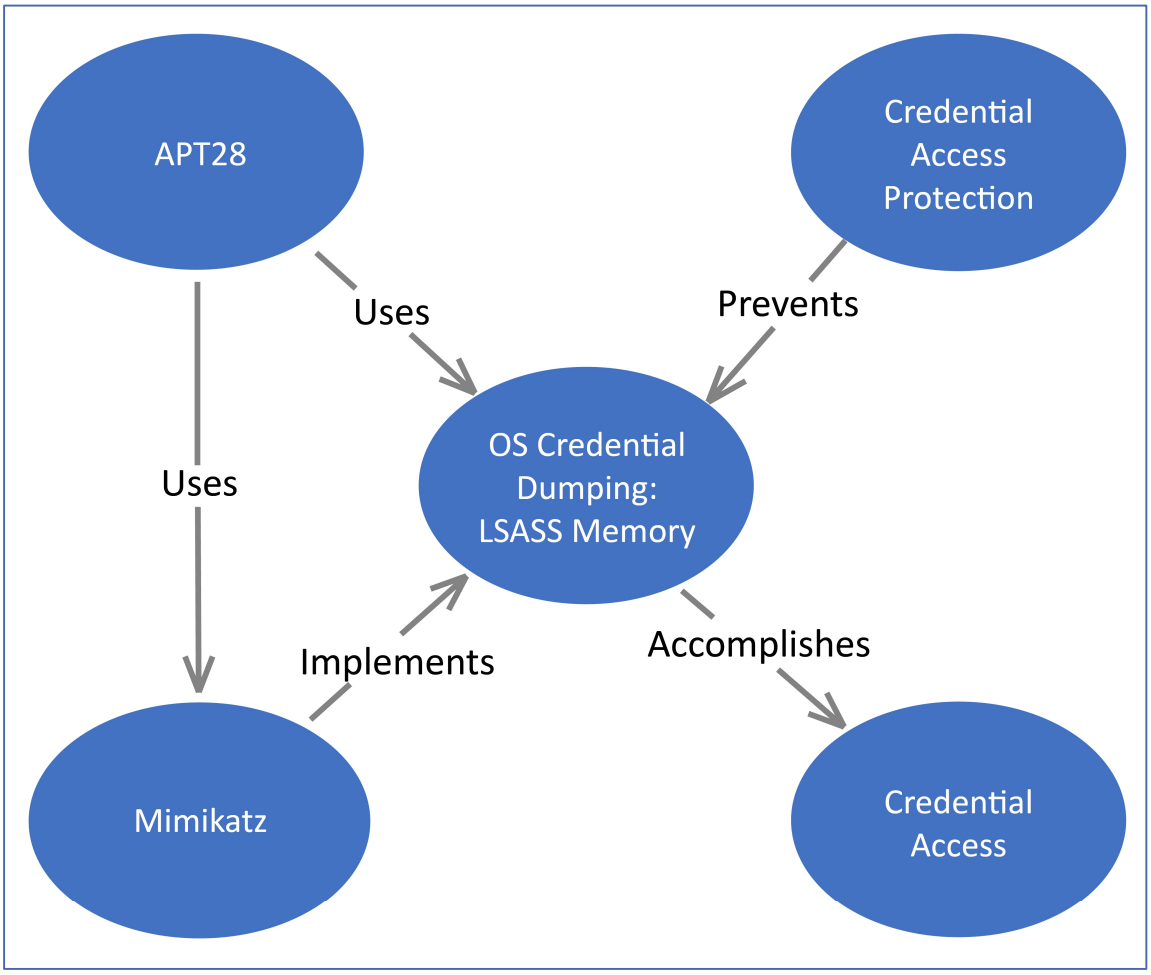

Figure 9 - Real-world MITRE ATT&CK object relationships

3.2.2.7 OBJECT STRUCTURE

All the above-mentioned objects have a precise taxonomy and contain a conspicuous number of metadata stored in an extremely strict object model.

The enforcement of these precise rules to order data and metadata is required to make the ATT&CK framework coherent, easy to maintain, update, and to consult. The framework is, after all, intended to be a globally shared and updated knowledge resource mainly based on community contributions, hence the need for these rigid consistency rules to keep entropy at the bare minimum.

All objects have a similar structure with many different metadata fields and tags to add descriptions of the objects and to define their attributes and relations to other objects.

Describing in detail every object structure of the MITRE ATT&CK framework would be outside the scope of the project, nevertheless, at least one example is needed to fully grasp the amount of information stored in the objects.
Table 5 shows how a *technique* object is structured, items with "*" are mandatory.

Data Item	Type	Description
Name*	Field	Name of the technique
ID*	Tag	Unique identifier.
Sub-techniques*	Field	IDs of sub-techs of this tech.
Tactic*	Tag	The tactic accomplished by this tech.
Description*	Field	General description of this technique, typical uses, reference to external articles analysing it.
Platform*	Tag	The target system, e.g.Windows
System Requirements	Field	Additional system requirement such as patch level or service pack.
Permissions Required*	Tag	Lowest permissions required.
Effective Permissions*	Tag	The permissions the attacker will obtain after using the technique.
Data Source*	Tag	Set of data sources that can be used to analyse this technique or its results. e.g. Process monitoring
Supports Remote	Tag	Can be executed remotely?
Defence Bypassed*	Tag	Can this evade any specific controls/safeguards?
CAPEC ID	Field	Link to related CAPEC page.
Version*	Field	Version of the technique.
Impact Type*	Tag	Illustrates if this technique can be used to impact availability or integrity.

Contributor	Tag	List of contributors outside MITRE who helped define the technique.
Procedure Examples	Relationship/Field	Short description of how this tech is used.
Detection*	Field	If detection of this technique is possible, this field will explain which sensors, data, and detective controls can be used on this technique.
Mitigation*	Relationship/Field	Any preventative or mitigating control that can be put in place against this technique.

Table 5 - ATT&CK: Sample object structure

Now that the main components of the ATT&CK framework have been analysed, the following section will illustrate the advantages and key features of the model.

3.2.2.8 UPDATES AND CONTRIBUTIONS

As stated in section 3.2.1, this framework is globally curated and shared, external contributions are always sought after and are pivotal to its growth.

In fact, MITRE's ATT&CK philosophy paper[52] contains a specific chapter dedicated to external updates and MITRE's own website has an entire page with additional information on how to contribute[66].

Both updates to existing elements and submissions of new ones are welcomed by MITRE. Analysts, defenders, and red/purple teamers are all encouraged to contribute to the framework as each of them has a different type of knowledge of adversarial actions and related TTPs, mitigations, and groups.

Every contribution must be submitted in a precise format and is then peer reviewed by MITRE.
The format follows the object structure of the element being submitted, for instance if a technique is being submitted, the format must follow the example shown in table 5.

MITRE also accepts contributions in terms of endpoint and / or network log data that will be used to test, polish, and refine existing knowledge.

3.2.3 KEY POINTS AND ADDITIONAL CONSIDERATIONS

This framework can be used for all the processes listed in 2.2.2 aside from the attribution one, at least by itself.

By using it, organisations will be able to enrich their CTI, enhance their detective and analytical capabilities, run more realistic penetration testing, red teaming, and adversarial emulation scenarios[67].

One of the most important aspects of this framework relies on its modularity, strict taxonomy, and relational database-like structure.
This has a handful of extremely useful consequences:

- The ATT&CK Matrix encompasses every known and documented tactic, technique, and sub-technique.
  Since no organisation will ever need full coverage of the whole ATT&CK Matrix, defenders will be able to pick only the tactics, techniques, and sub-techniques relevant to them based on their Cyber Threat Intelligence.
- If the CTI of an organisation is somewhat lacking, they can still get some useful intelligence insights from the framework itself to start planning their defences.
- However, if the CTI of an organisation is at a high level they will be able to map their own intelligence findings to the ATT&CK framework[67].

It is also extremely interesting to point out that the MITRE ATT&CK framework is not just a taxonomy, it might be defined as an ontology of cyber adversarial behaviour.

A taxonomy hierarchically classifies elements within a category, it prescribes a structure; an ontology is more sophisticated, it distinguishes and identifies its contents and their relationships, just as MITRE ATT&CK does with its core elements[68].

This is further corroborated by the fact that the framework has recently started supporting STIX 2.0[69][70], a JSON based language and serialisation standard, specific for cyber threat intelligence, similar to OWL[71]; Ontology Web Standard and its semantic reasoners are precisely used to detect errors in ontologies and draw inferences from rules and contents.

This recent support of STIX 2.0 not only backs up the ontology nature of ATT&CK but makes its automation and the exchange of its structured threat intelligence much simpler.

This nature of the framework also has an interesting corollary: while it was always intended to be a globally shared and curated knowledge base, its ontology-like schema can be used by companies to create their own, private, knowledge bases and matrices. Intuitively, these will be fully compatible with the public ATT&CK matrices, as they all pertain to the same ontology.

Many leading companies in the information security sector have been adopting this framework as their go-to solution for cyber attack behaviour modelling and threat hunting. It can be found, for instance, in FireEye[72], Crowdstrike[73], and Kaspersky[74] products and reports.

To summarise, the **MITRE ATT&CK framework**:

- **Is meant to be a globally shared and curated knowledge base of adversarial behaviours for intrusion modelling** (see section 3.2.1)**.**
- **Is constantly updated with new and documented tactics, techniques, procedures, software, groups, and mitigations reported by reputable researchers, both private and public** (see section 3.2.2.8)**.**
- **Is easy to consult, import, and update in a coherent fashion thanks to its precise taxonomy, rules, and support for a serialised and ontology-scoped language such as STIX 2.0.**
- **Is modular and divided into different technology domains so that every organisation can apply what they need most** (see section 3.2.1)**.**

The project will now analyse the last of the seminal intrusion models, the Diamond Model.

3.3 DIAMOND MODEL

The Diamond Model (hereinafter DM) was released in 2013, after seven years of research by A. Pendergast, S. Caltagirone, C. Betz.

They were all intrusion analysts and they wanted a model that could answer the question: "How do we do our work?"[75].

The DM is an extremely formal mathematical framework that leverages game[76]**, graph**[28]**, and clustering theories**[77] **to enhance defenders analytical and decision making capabilities.**

This strict formality is intended and aimed at having accuracy and repeatability of results, testable analytic processes, simpler hypothesis generation, better forecasting, classification, and correlation of events.

While being extremely formalised, the model is also generic enough to be expandable, versatile, and scalable. It was designed to easily grow and incorporate new concepts.

The next section will briefly explain how the model works.

3.3.1 STRUCTURE AND LOGIC

This model aims to be a tool for analysts to enhance and streamline their analysis of adversarial activity.

To do so the DM presents one atomic element, the *event*, which can be described and characterised by its core features and meta-features.

On top of that element and its characteristics, the authors of the model also propose a set of analytic processes.

Axioms, a foreword

The Diamond Model is built around *seven axioms***,** understanding a select few of them is necessary to fully appreciate this project; those will be reported verbatim and within quotes in their relevant sections.

3.3.1.1 DIAMOND EVENT

"Axiom 1: 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."[78]

The DM has one atomic element, the event.

An event is a distinct and time-bound activity performed by an *adversary* that employs a *capability* over an *infrastructure* against a *victim*.
The core features of an event are adversary, capability, infrastructure, and victim.

The model is also capable of supporting additional features, called meta-features. Six of them are proposed by the authors, but the model can have an unlimited number of meta-features. They will be thoroughly explained in this chapter.

Figure 10 depicts the atomic element of the model, the Diamond Event.

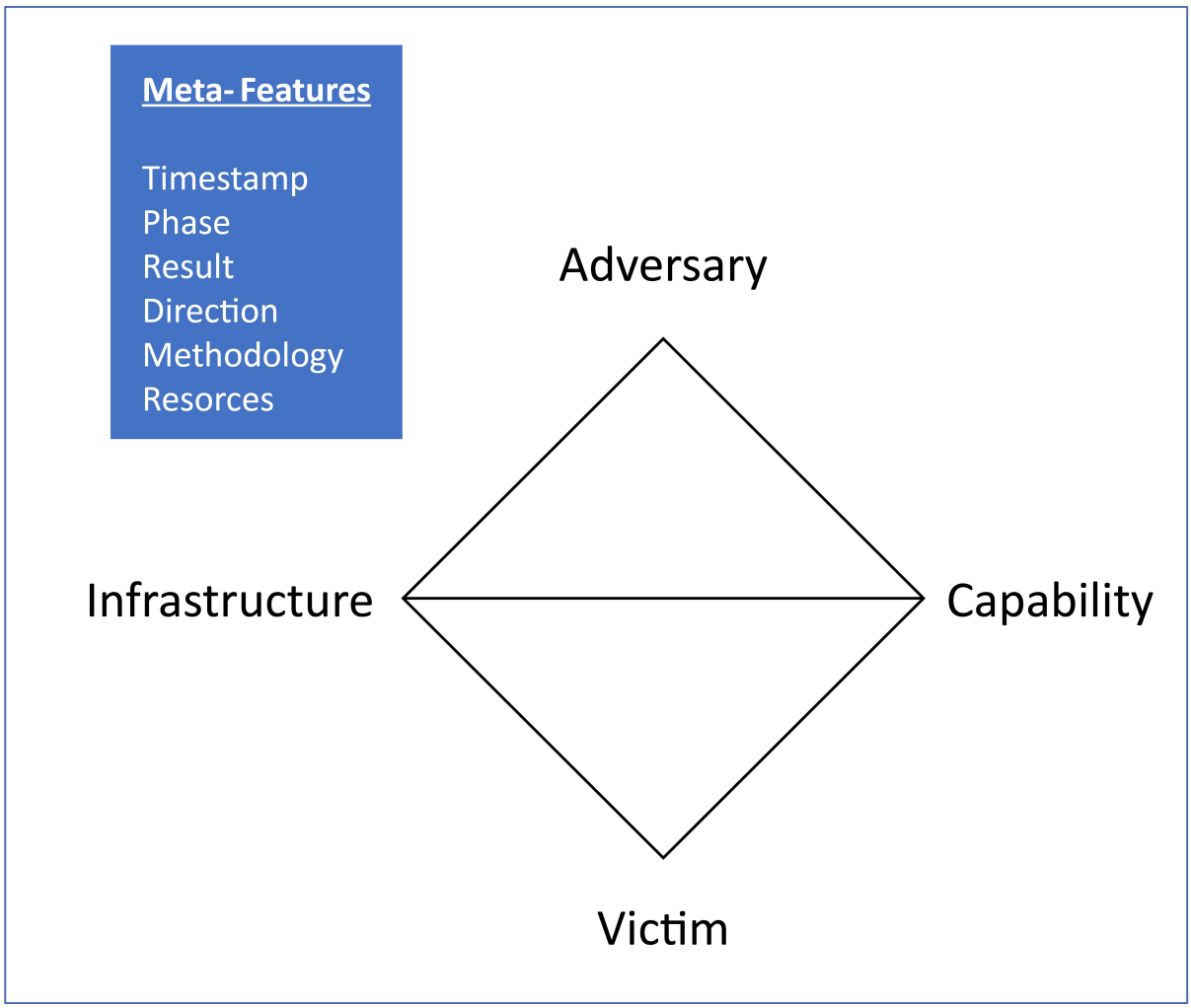

Figure 10 - The Diamond Event

Both core features and meta-features have a *confidence value*, however, the model does not provide a value for it, as different implementations of the DM might have different interpretations of the concept of confidence. Since confidence can be a function of multiple values, the model supports its representation with tuples and sub-tuples.

In terms of mathematical representation, an event E can be defined as a variable-sized n-tuple where every element of the tuple is knowledge of a feature combined with its confidence value, as per the below formula.

$$\begin{aligned} Event = (&(Adversary,\ Confidence_{adversary}),\\ &(Capability,\ Confidence_{capability}\ ),\\ &(Infrastructure,\ Confidence_{infrastructure}),\\ &(Victim,\ Confidence_{victim}),\\ &(Timestamp,\ Confidence_{timestamp}),\\ &(Phase,\ Confidence_{phase}),\\ &(Result,\ Confidence_{result}),\\ &(Direction,\ Confidence_{direction}),\\ &(Methodology,\ Confidence_{methodology}),\\ &(Resources,\ Confidence_{resources})) \end{aligned}$$

Elements of the n-tuple can be divided into further sub-tuples as deemed necessary, to present and work on a more granular and detailed knowledge base.

For example, the *Victim* element can be divided into multiple sub-tuples to represent all the information analysts have on it, as shown in the below formula.

$$\begin{aligned} (Victim, Confidence_{victim}) = &((Organization, Confidence_{organization}),\\ &(HostIP, Confidence_{hostIP}),\\ &(Hostname, Confidence_{hostname})) \end{aligned}$$

As seen in figure 10 an event E can be represented as an undirected, simple graph, with the following relations.

$$E_{vertices} = \{Adversary, Capability, Infrastructure, Victim\}$$

$$\begin{aligned} E_{edges} = \{&\{ Adversary, Capability\},\\ &\{Adversary, Infrastructure\},\\ &\{Infrastructure, Capability\},\\ &\{Infrastructure, Victim\},\\ &\{Capability, Victim\}\} \end{aligned}$$

3.3.1.1.1 CORE FEATURES

As mentioned above, four core features define an event object.

These are listed and explained below.

Adversary

"**Axiom 2** There exists a set of adversaries (insiders, outsiders, individuals, groups, and organisations) which seek to compromise computer systems or networks to further their intent and satisfy their needs."[78]

The general definition of a threat actor given by the project in section 2.1.1 can be used to describe this core feature.

Pendergast, Caltagirone and Betz, however, make a further distinction.

They split the threat actor into two separate entities:

- *Adversary operator*: The actual person performing the intrusion activity.
- *Adversary customer*: The ultimate stakeholder benefitting from the operator's actions.

e.g., A well-funded customer might use multiple operators, with shared or different goals, all using different capabilities and infrastructures.

Such a granular perspective allows a much more detailed knowledge of adversarial motive and resourcing, ultimately leading to better threat intelligence, attribution, and overall risk mitigation.

The socio-political needs, presented in one of the upcoming sections, are used within this model to generate intel about the motive(s).

Capability

This represents the tools and techniques being used by the attacker during the event.
Both unsophisticated, commodity, and custom highly complex tools are to be considered.

The authors also define, under the capability feature, two concepts.
The adversary arsenal and the capability capacity.
The former is the sum of all adversarial capabilities, the latter is the sum of all the targets exposed to the adversary arsenal.
In other words, how many different weapons the threat actor has and how many different targets those weapons can have.

Infrastructure

This feature represents the structures, digital or analogue, that the adversary employs to carry out its attacks, to maintain C2, and to collect information from the victims (e.g., exfiltrating data).

Like every feature of the DM, the infrastructure can be scoped as broadly or granularly as needed.

Examples of infrastructures are the Internet, or specific domains or protocols such as DNS, or even USB drives purposefully "dropped" in a parking lot next to the office of the target company[79].

The authors of the DM, once again, further break down the concept of infrastructure into three different objects:

- Type 1 infrastructure is one completely owned and operated directly by the attacker.
- Type 2 infrastructure is one under the control of an intermediary that can be wilful and informed or not. (e.g., zombie botnets, compromised service accounts)

Service providers are those entities providing the availability of both the above types. e.g., ISPs

Victim

This feature represents the target of adversarial capabilities.

Just like the previous features, it can be described using a multitude of different parameters, at various levels of granularity, such as target domains, systems, websites, perimeters.

Once again, the authors make an important distinction when analysing the victim feature of an event.

They differentiate between *victim persona* and *victim assets*.
The difference is self-explanatory, but nonetheless extremely important.
The persona is mainly useful when the analytic effort is non-technical, for example when studying cyber-victimology and socio-political correlations, while the assets are the main concern when a technical analysis is being made, for example when analysing the TTPs employed by the enemy.

Figure 11 illustrates how a Diamond Model event and its analytic features can be used to describe a real-life attack scenario.

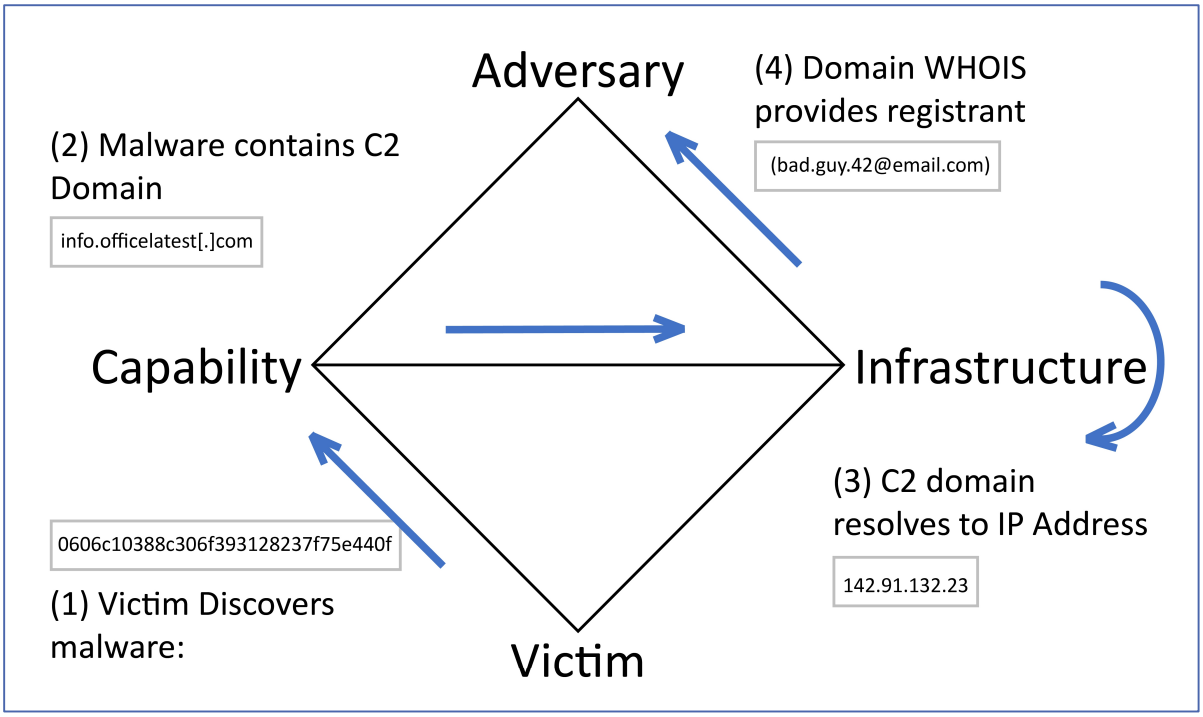

Figure 11 - The Diamond Event in a real-world scenario

3.3.1.1.2 META-FEATURES

The features described above are the core ones that make up an event.
The authors of the DM realised that there could be different - and optional - meta-features to be added to the event object to better analyse an intrusion.

The model can accommodate any number of meta-features.
Describing all the possible ones in detail is beyond the scope of this project, however, a shortlist with a brief explanation of those suggested by the authors can be found below:

- *Timestamp*: a chronological notion of the event, date and time it started, or how long it lasted. Useful to understand periodicity, patterns, or to write a decay function to analyse attackers' changes as time passes.
- *Phase*: Axiom 4 of the DM states that every malicious activity happens in one or more phases, a universally accepted notion[80].
  This meta-feature describes with an ordered n-tuple P the phases of an activity. e.g. $P=(p_1,p_2,p_3,...p_n)$
- *Result*: This describes the possible results of the analysed malicious activity. This result can be expressed in many ways. e.g., A n-tuple stating which part of the CIA triad is compromised.
- *Direction*: This meta-feature describes the directionality of an activity; it can provide useful intelligence to better place detective and mitigating controls. e.g., Victim-to-Infrastructure, Infrastructure-to-Victim, Infrastructure-to-Infrastructure [...]

- *Resources*: This is used by analysts to describe and list any element supporting the malicious activity being studied. e.g., Software, funds, knowledge, access to systems.
- *Methodology*: This meta describes a general class of activity. There can be multiple for each event and once again the DM allows analysts to fill this feature with whatever methodology they want, with the specificity and granularity they need. e.g., spear-phishing, credential dumping.

3.3.2 EXTENDED DIAMOND MODEL

The Diamond Model can be extended by adding any number of meta-features.

In addition to those explained in the previous section, Pendergast, Caltagirone and Bentz present two very peculiar ones.
The social-political and the technology meta-features.

When these two are added to the event object, we have the extended Diamond Model. These meta-features are extremely peculiar because they overlay two core-features defining a precise relationship.
Figure 12, below, illustrates the graphical representation of the extended Diamond Model.

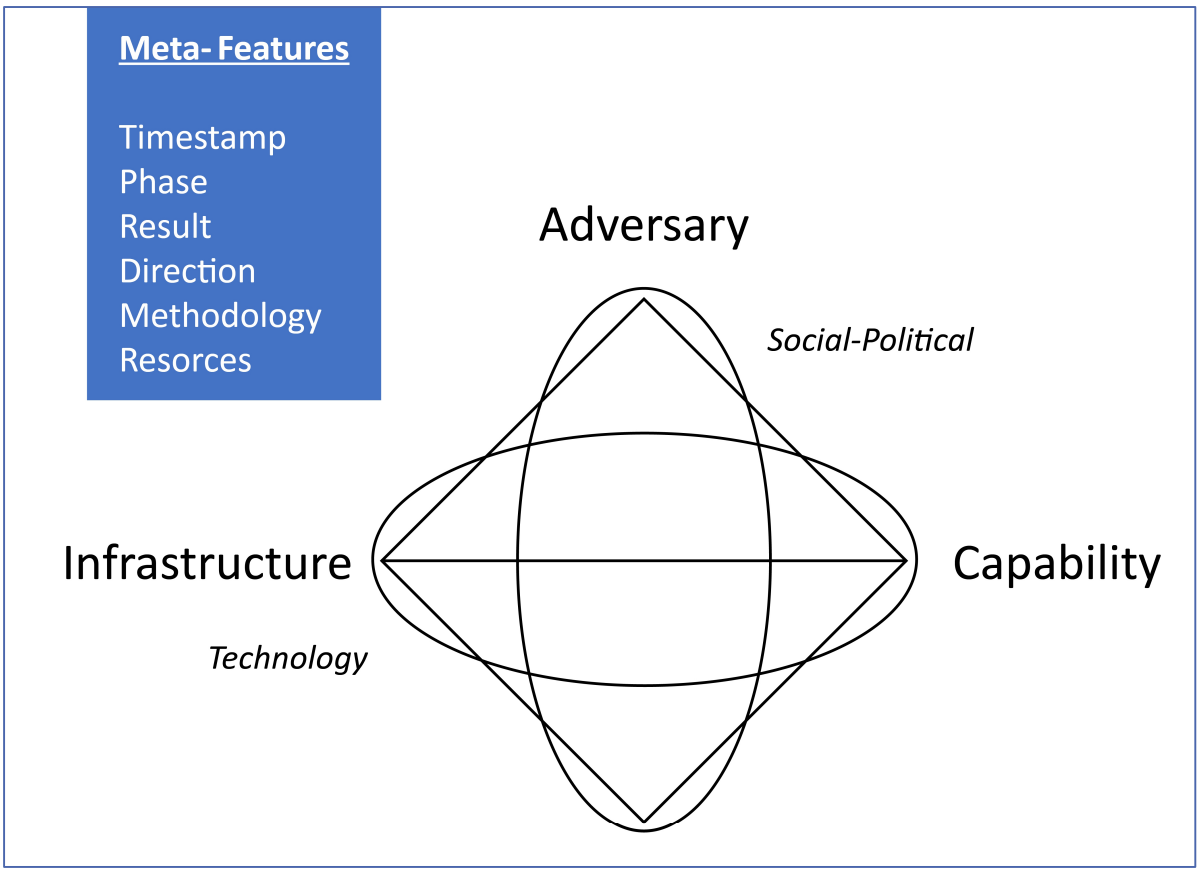

Figure 12 - The extended Diamond Model

3.3.2.1 SOCIAL-POLITICAL META-FEATURE

The social-political meta-feature (figure 12) stems from the assumption (Axiom 6) that there is always a connection, direct or indirect, fleeting or persistent, between adversary and victim.

It describes the relationship as a consumer-provider one, where the adversary is looking for something that can be provided by the unwitting victim.

The social and political needs of the adversary are the foundation of this relationship.

e.g., A nation state (the adversary) might want to reduce the technology gap with its enemies (the social-political need) by conducting espionage (the intent) against another nation state (the victim).

This meta-feature is extremely important as it is the foundation of one of the analytic pivoting techniques explained in section 3.3, the victim-centred approach.
In addition to that, it allows expertise from criminology and victimology fields to be brought bring into the analysis of attacks and intrusions, answering questions such as:

- If we have multiple victims, do they have shared traits?
- Can we create a set of victims out of those traits?
- Can adversarial intent be deduced from the set of victims?
- What adversary might have the needs behind such an intent?

Moreover, the degree of persistence or lack thereof of adversaries, and their tendency to endure or flee after detection or public exposure is also defined within this relationship meta-feature[81].

This last analytical dimension - persistence - is of the utmost importance when dealing with APTs and the authors of the DM illustrate it in a very detailed way, beyond the scope of this research.

3.3.2.2 TECHNOLOGY META-FEATURE

The technology meta-feature (figure 12), in a similar fashion to the social-political one, defines the special relationship between the capability and the infrastructure vertices of the Diamond Event.

It describes the technology that allows the infrastructure and the capability to liaise.

e.g. If a malware resolves domains and uses http verbs to communicate with its C2 server, the technologies used are TCP/IP, DNS, HTTP.

Knowledge of this meta-feature allows analysts to focus on the technology used itself, and not on the underlying infrastructure and capabilities. It serves as an indirection layer that allows a capability-agnostic and infrastructure agnostic approach.

3.3.3 ANALYTIC PIVOTING

It is evident so far that one of the main purposes of the Diamond Model is improving analysts' tradecraft.
In this sense, the model support of *analytic pivoting* is one of its greatest strengths.
Analytic pivoting is a practice where as much data as possible is drawn from a single, specific element, and in conjunction with other data sources, related elements are found.
The very graphical representation of the Diamond is an illustrated example of this analytic method.

A Diamond representing an event E is made up of four vertices with edges connecting them. The edges are the graphical representation of the pivoting opportunities, which can be leveraged by analysts that only have knowledge of one vertex, to shed some light on the other three.
Figure 13 will illustrate this concept.

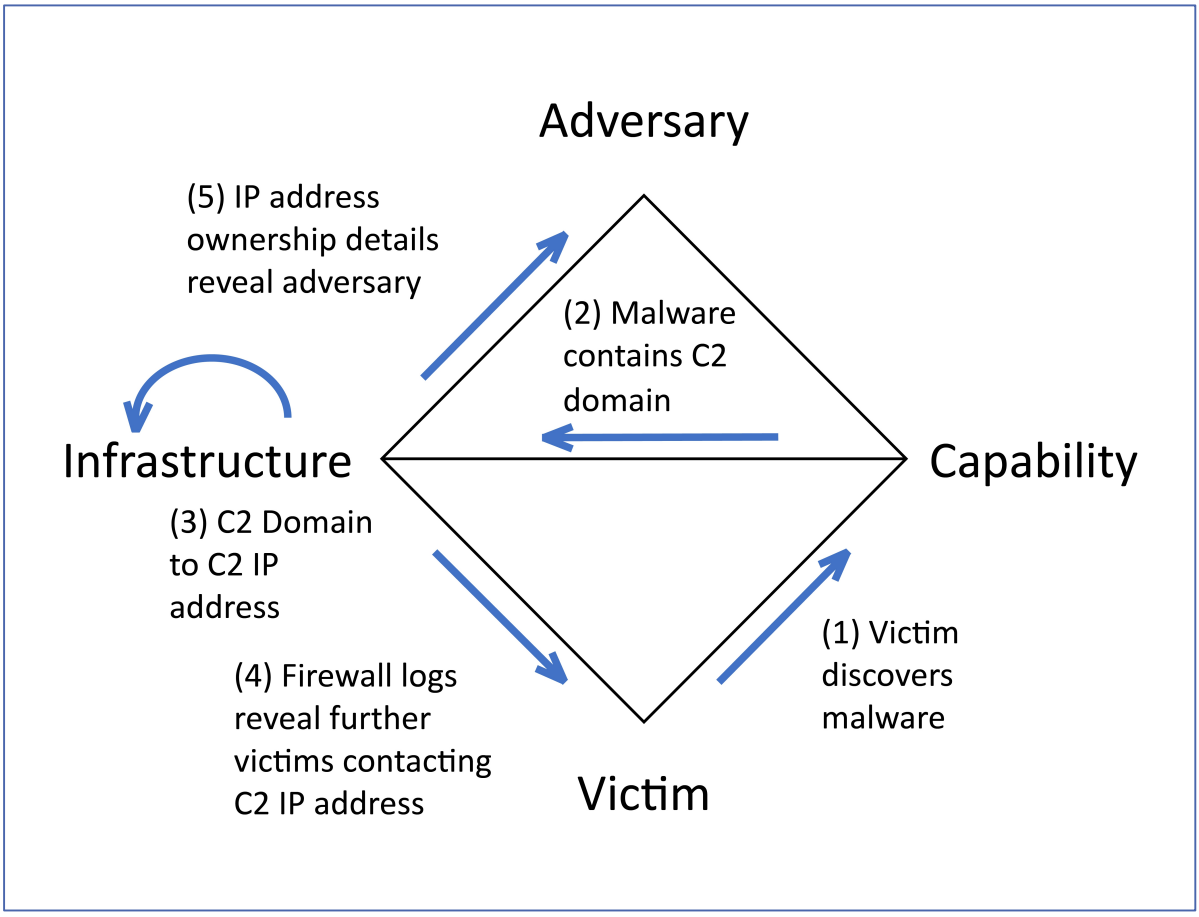

Figure 13 - An example of analytic pivoting of the Diamond Model

This is an extremely broad example of the analytic pivoting supported - and suggested - by the model.
As mentioned above, every edge is a pivoting opportunity, every blue arrow is a real-life example of successful pivoting.
In detail:

1. The Victim finds malware in its systems, the analysts have performed the first pivot, from Victim to Capability (e.g., the malware family).
2. The malware gets examined and reversed, a C2 domain is found. This is the second pivot, leading the analytic process to the Infrastructure vertex.
3. Examining the C2 domain, two pivots can happen by resolving the domain to its address(es):
    a. analysts can get insight into the Adversary via verifying the ownership of the IPs
    b. analysts can check within their network for more connections to those IPs, discovering new Victims

When, during the analytic process, the focus is explicitly set on a specific feature to discover other activities or features, we have a *centred approach*, e.g., capability-centred approach, adversary-centred approach, et cetera.

3.3.4 ACTIVITY THREADS AND GROUPS

3.3.4.1 ACTIVITY THREADS

As per the fourth axiom of the DM, explained in the meta-features section, adversaries always operate in two or more phases, where the success of each phase is necessary to reach the next one.

The model, in its efforts to augment analysts' tradecraft, also provides a way to order events in phases and link them by their causal interrelation.
The result of this cataloguing in phases and causally linking is called an *activity thread*.
A formal, mathematical definition of activity threads is given by the authors of the Diamond Model. The project will not analyse it in detail, as it is beyond the research scope, but it can be found in Appendix C at the bottom of the paper.
An example of how activity threads are shaped and work, however, will be given below.

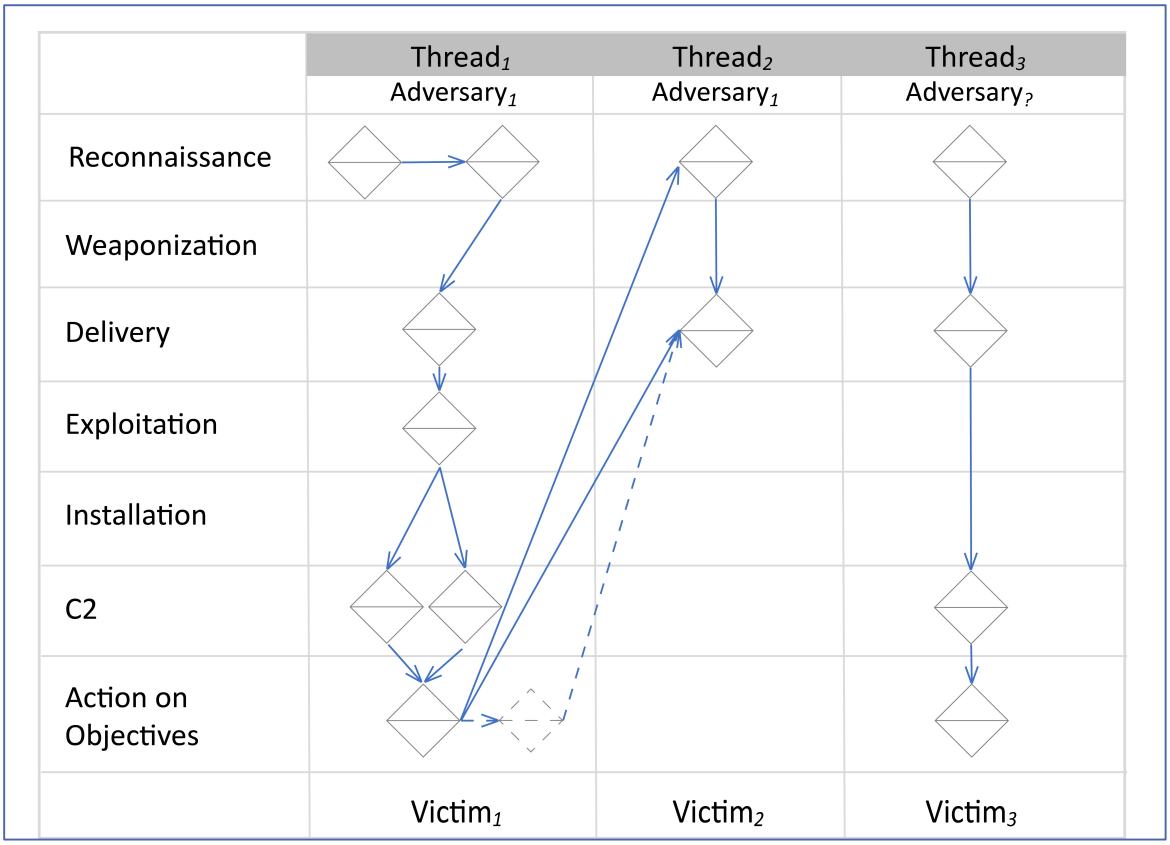

Figure 14 - Diamond Model activity threads

Figure 14 represents the activity of one known adversary ($Adversary_1$) against two different targets ($Victim_1$ and $Victim_2$) and the actions of an unidentified adversary ($Adversary_?$) against a third target ($Victim_3$).

The arcs denote a causal relationship between the events, if dotted they represent a *possible* causal relation that still needs to be corroborated.
Every arc can be associated with a degree of confidence.
The columns of the graph represent the vertical correlation within one adversary-victim relationship.

When an activity spans across multiple columns, we have horizontal correlation.

Horizontal correlation can be extremely useful to identify defensive gaps and synthesise information on adversarial action across multiple activity-threads as well as to organise adversaries in threat groups.

3.3.4.2 ACTIVITY THREADS FOR INTELLIGENCE ANALYSIS

Activity threads being phase-separated, formal, and highly visualisable makes them the perfect support for the full hypothesis lifecycle (generation, documentation, testing).

They can be easily integrated with modern intelligence analysis methods to generate actionable intelligence.

The project will now illustrate this process, its importance is highlighted by the very authors of the model.

Let us assume a malicious activity has been identified by some detective controls.

As per axiom #4 analysts know that malicious activity is necessarily made up of consecutive phases, each containing at least an event, by representing the malicious activity into an activity-graph, they will be able to:

- identify knowledge gaps. e.g., analysts might know everything about a specific event and its core-features but might miss knowledge about the resources meta-feature.
- generate and document hypotheses. e.g., in relation to the previous step, analysts might start generating hypotheses to answer the question “What resources are needed to launch a similar attack?”.

These hypotheses, once generated and documented, can be polished and put to test.

Modern intelligence analysis offers several methods to achieve that, ranging from Occam’s Razor to conservatism to Delphi techniques to ACH (Analysis of Competing Hypotheses). Several of these methods are staples of modern intelligence analysis, developed and formalised by Dr Richards Heuer, a researcher often referenced by the author of the Diamond Model[82]–[84].

The model’s attempt to be a tool to enhance analytic tradecraft is even more evident with activity-threads.

3.3.4.3 ACTIVITY-ATTACK GRAPHS FOR MITIGATION

The model leverages activity threads and competing hypotheses to also produce actionable intelligence on defensive gaps and requirements.

It does so by integrating activity threads with *attack-graphs*.

Attack graphs are graphical representations of all the practicable attack paths a malicious actor can use to strike a target asset[80].

Activity-attack graphs can be used to answer the following question:

“We have a known adversary with a documented attack pattern, what can we expect from variations of this pattern? How can we be more resilient to changes in adversarial behaviour?”

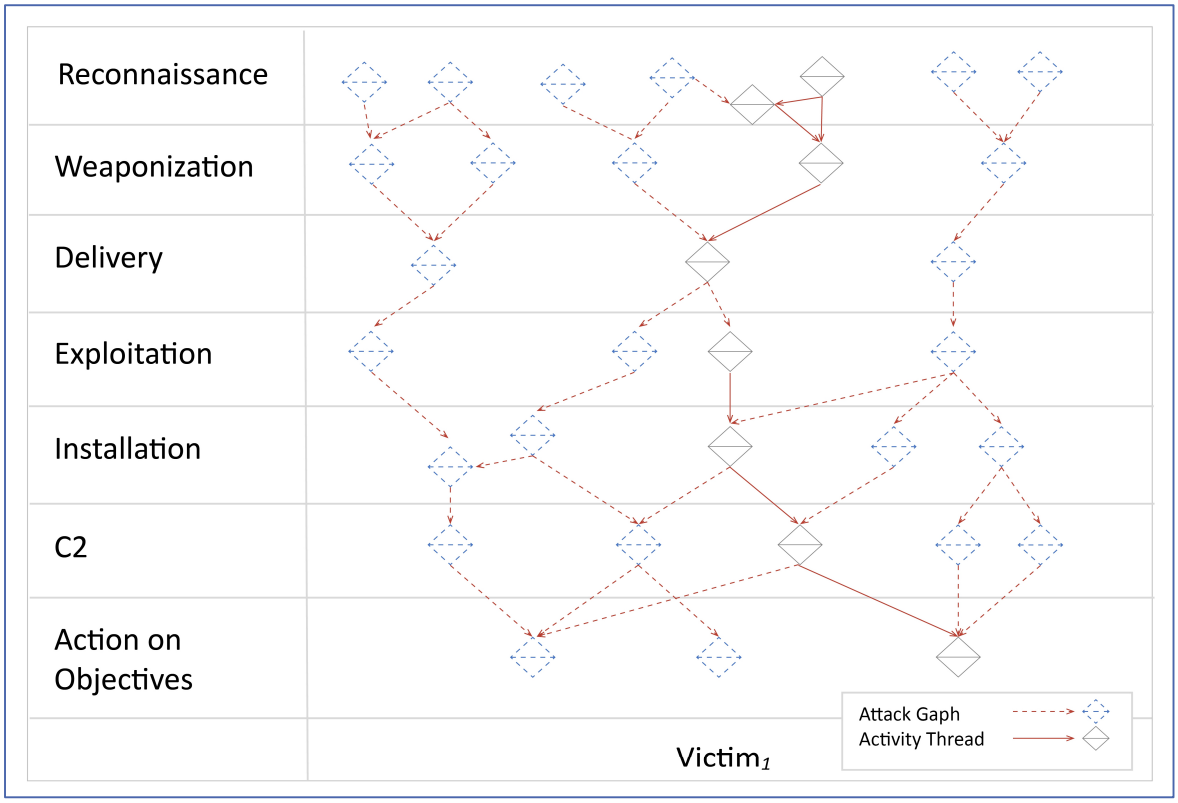

Figure 15 - Diamond Model activity-attack graphs

The activity-attack graph, as represented in figure 15 illustrates exactly how analysts can visualise both *what has already happened*, with an activity graph represented by a solid line, and *what might happen*, with all possible attack paths represented by the dotted attack graphs.

The knowledge represented in that graph can be leveraged by defenders to better understand their defensive gaps and requirements.

Once again it can be appreciated how this model is designed to help analysts produce superior intelligence by offering a simple and intuitive way to generate and document hypotheses.

3.3.4.4 ACTIVITY GROUPS

Diamond Model events and activity-threads can be clustered together based on similarities they might have.

The model provides a formal, structured, and thorough process to create *activity groups* to answer higher level analytic problems (e.g., establishing life cycle timelines, identifying common capabilities developers, defining common adversaries behind attacks and campaigns).

The process is made up of the following steps:

1. The problem the grouping is aiming to solve must be defined.
2. The features on which to establish similarity have to be specified.
3. An activity group is created and populated by grouping known events with similar features.
4. When new events are discovered, they are classified in an existing group, if showing the relevant similarities
5. The group is analysed to try and give an answer to the problem defined in #1.
6. Periodic consistency checking and redefinition of groups is required.

There is an extremely formal mathematical definition of activity-groups and related analytic activities, despite it being outside the scope of the project, it can be found - unabridged - at the end of the paper as Appendix D.

With activity-groups, the authors of the DM are once again providing analysts with a powerful tool to enhance their tradecraft, especially in the attribution field.

3.3.5 KEY POINTS AND ADDITIONAL CONSIDERATIONS

The Diamond Model is intentionally generic while still containing all the essential and atomic dimensions of intrusion analysis.

Its nature is fundamental, mathematical, and inherit flexible, as explained in section 3.3.1. **It was intended to be an underlying conceptual framework for better analytic tradecraft**, **for it leverages graphs** (see section 3.3.4.1-3) **and clustering theories** (see section 3.3.4.4) **and analytic pivoting** (see section 3.3.5) **to augment analysts' efficiency, precision, and correlation and information synthesis capabilities.**

It was intended as a tool to integrate with, and to broaden the perspective of, other intrusion models and attack-graphs (see section 3.3.4)**.**
This is also corroborated by the constant references the DM authors make to the CKC, as can be seen in figure 14 and in figure 15, where activity threads and activity-attack graph elements are always mapped to the Lockheed Martin's kill chain phases.

Additionally, due to its flexibility, the Diamond Model can easily be integrated with most decision making and planning frameworks[75].

It can form the basis for cyber ontologies, but it is not one. It doesn't aim to be a protocol, taxonomy, or knowledge management system[85].

These concepts will be further expanded in the next chapter, where the project will compare the Diamond Model with Lockheed Martin's Cyber Kill Chain and with MITRE's ATT&CK framework.

4 MODEL COMPARISON

All the models presented thus far are well-respected industry standards for intrusion and attack analysis, however, they are vastly different from one another.

The project will now compare them.

4.1 AN ADDITIONAL NOTE ON COMPARISON METHODOLOGY

There is no academic research explicitly dedicated to the comparison of intrusion analysis models.

One of the challenges of this project, aside from thoroughly understanding and illustrating the defining characteristics of the three models, has been finding a valid logic to compare them.

The comparison needed to be detailed and comprehensive enough, to appreciate even the slightest differences, yet not overzealously punctilious, as the time and resource constraints of the dissertation would not allow that.

The following method, divided into two phases, has been conceived by the author to achieve an exhaustive comparison of the models within the aforementioned scope and constraints.

The **first phase** will focus on the specific differences between the models.
The differences examined will be, in order of importance:

- The **purpose** of the models
- The **abstraction** level of the models
- The **design** of the models

This is an ordered list, as the entire method revolves around the following axiom.

From the purpose of a model stem, to address its architectural and operative requirements, its abstraction level and its design.

The **second phase** will reinforce those dissimilarities by providing examples of complementary and integrated uses of the models.
By illustrating how an entity complements another one, their differences become even more evident.

4.2 DIFFERENCES

As per the methodology note above, the project will now examine the different purposes, abstraction levels, and design choices of the models.

4.2.1 DIFFERENT PURPOSES

The key to fully grasp their differences relies on understanding how their purposes diverge.

Lockheed Martin's Cyber Kill Chain, the Diamond Model, and the MITRE ATT&CK framework were created by different researchers, to solve different problems, and as such, they all serve different purposes.

The CKC model was created as the foundation of an intelligence-driven approach to protect the cyber perimeter of organisations.

The purpose of the model is to help defenders enhance their assets' resilience by understanding and controlling the attackers' Cyber Kill Chain, the ordered and precise phase-based sequence of adversarial intrusion activities. It gives analysts a high-level guideline to recognise and represent attack patterns.

The MITRE ATT&CK framework was created in an attempt to establish a globally shared and curated knowledge base of documented attacks, tactics, techniques, procedures, groups, mitigations and the relations between them.

It achieves this by having an extremely specific taxonomy, a solid object structure, strict rules ordering the relationships among different objects, and very formal procedures to be expanded by incorporating new and updated information.

The Diamond Model was created to improve defenders' analytic capabilities, especially on the grouping and attribution of intrusion events.

It provides analysts with the most essential, formal, and foundational conceptual framework to structure their workflow and augment their hypotheses generation, documentation, and testing capabilities.

To summarise:

- **The CKC is a model used to produce actionable intelligence by analysing adversarial kill chains.**
- **The MITRE ATT&CK framework is a, globally shared and curated, knowledge base of intrusion TTPs, groups, and mitigations and the relations between them.**
- **The Diamond Model is a tool to improve defenders' analytic processes and intelligence analysis tradecraft, with a focus on clustering and attribution.**

The DM is by far the most different of the three.

While LM's and MITRE's models' main focuses are adversarial activity and behaviour, the Diamond Model is primarily concerned with analysts' workflow, cognitive processes, and decision making.

Such a remarkable eschatological difference, by itself, requires the project to temporarily set aside the Diamond Model, to focus on the other, more comparable, two.

4.2.2 DIFFERENT ABSTRACTION LEVELS: CKC VS MITRE ATT&CK

Unlike the Diamond Model, both Lockheed Martin's Cyber Kill Chain and MITRE ATT&CK framework do precisely focus on intrusions and attacks, but they do so in two very distinct ways, at different *abstraction levels*.

The Cyber Kill Chain has a higher abstraction level, useful to understand higher level adversarial processes, goals, patterns, intents.

However, it lacks the notions required by the defenders to describe the tactical goals and the implementations of specific malicious actions, as well as the related controls and modi operandi of the threat actors that leverage them.

Those entities are better described with a **mid-level abstraction model such as the MITRE ATT&CK framework** with its strict taxonomy and relational structure, filled with the intelligence gathered from several millions of real-world, documented and catalogued, intrusions.

A scenario to better understand the above statement:

- A privilege escalation is being reported by the SOC of an organisation.
- Analysts using the CKC can reconstruct the phases preceding the privilege escalation and can try to synthesise information about the next step of the attacker.
- However, the kill chain approach has no notion of the exact techniques and procedures used.
- Analysts will need MITRE's framework to describe the detailed adversarial behaviour and to answer questions like "What technique did the attacker use to escalate privileges? How was that technique implemented?".
- In this scenario, for example, the malicious actor might have escalated privileges by *Abuse of Elevation Control Mechanisms* (MITRE technique)[86] and more specifically by exploiting *Setuid and Setgid* (MITRE sub-technique)[87].

In this abstraction-based representation, an even lower tier would be represented by vulnerability[88] (or malware[89]) databases, where specific software and code examples are found, but completely devoid of any intrusion analysis context dimension, such as TTPs, adversarial intent, et cetera.

Figure 16 helps illustrate these three levels of abstraction.

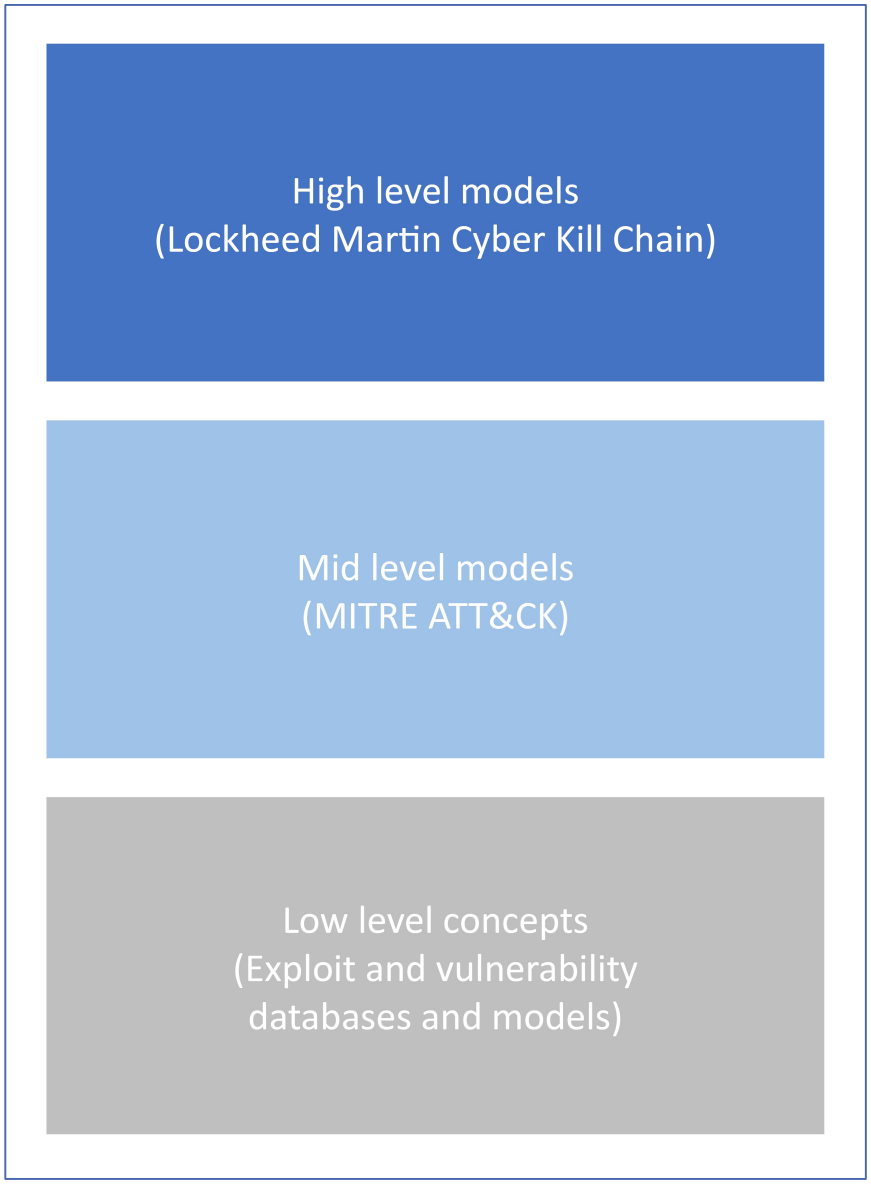

Figure 16 - Abstraction levels

This CKC and ATT&CK abstraction difference is further corroborated by the simple comparison between the phases of the former and the tactics of the latter.
Those are the topmost entities of each model, but Lockheed Martin's kill chain has seven phases, while MITRE's framework has fourteen tactics (see table 6).
Furthermore, the latter expands those 14 tactics into hundreds of known and documented techniques, sub-techniques, procedures, and related threat actors and mitigations.

Cyber Kill Chain phases	MITRE ATT&CK tactics
Reconnaissance	Reconnaissance
Weaponisation	Resource Development
Delivery	Initial Access
Exploitation	Execution
Installation	Persistence
C2	Privilege Escalation
Action on objectives	Defence Evasion
Ø	Credential Access
Ø	Discovery
Ø	Lateral Movement
Ø	Collection
Ø	C2
Ø	Exfiltration
Ø	Impact

Table 6 - CKC/ATT&CK topmost entity comparison

These are the main abstraction differences between the CKC and MITRE ATT&CK framework.

The project has purposefully not applied these distinctions to the Diamond Model, for its focus is not on cyber attacks themselves, but rather on the cognitive processes involved in their analysis.

The DM does not belong to a different abstraction level, but to an entirely different domain.

The comparison will now continue, including once again Pendergast, Caltagirone and Bentz's work, and will focus on the design differences.

4.2.3 DIFFERENCES IN MODEL DESIGN

As illustrated in the previous sections, the three models serve different purposes and work at different abstraction levels.

This also influences their design.

The dissertation will focus on three aspects of the design of the models:

- *Rigorousness*: how strictly and formally are the core elements defined?
- *Upgradability*: how easy is it to incorporate new ideas and elements?
- *Versatility*: what degree of freedom do have analysts using existing elements?

The core elements taken into consideration, for each model, will be:

- Diamond Model - Event, features, meta-features
- Mitre ATT&CK - Anything defined in its whitepaper as an object. (TTPs, groups, mitigations, software)
- LM's Cyber Kill Chain - The kill chain phases

These are considered, within the design documents of each model, their core elements[42][52][78].

4.2.3.1 RIGOUR IN THE DEFINITION OF CORE ELEMENTS

One of the parameters chosen to compare the design of the models is how formal, strict, and logically-sound their core elements are defined.

MITRE's framework and the Diamond Model provide extremely formal, logical, and mathematical definitions of their core elements.

The project has shown that the former has strict rules structuring both its core objects (TTPs, groups, mitigations) and the relations between them (see sections 3.2.2 and 3.3.1). The latter has extremely detailed mathematical descriptions of events, confidence, groups, and clustering rules (see section 3.3.1 and appendices C and D).

The Cyber Kill Chain model does not have such a formal and strict definition of its core elements.

This has practical consequences.
For example, the CKC is the easiest of the three models to adopt and implement, as its core elements are easy to understand and even a basic understanding of them is sufficient to bootstrap it.

However, MITRE ATT&CK and the Diamond Model are not as simple to understand, due to the complex definition of their respective core elements.
The former, in its ontology-like rigour (see table 5), requires an in-depth understanding, especially when automating it, integrating it, or using its schema to create custom knowledge bases.
The latter, while being easier to understand at a basic level, requires a very thorough understanding of the mathematical properties of its elements to be used with full consciousness and in the most advantageous way. These mathematical properties, as mentioned above, are extremely strict and rigorous (see appendices C and D).

The versatility and expandability of a model are in no way related to the rigorousness applied to the definition of its core terms.
The next section will illustrate how the more formal models are the most ductile and expandable.

4.2.3.2 UPGRADABILITY

New technologies and behaviours require the models to expand themselves, incorporating new concepts to better understand the updated landscape.

Both the Diamond Model and the ATT&CK framework are designed to incorporate new ideas and grow to suit analysts' renewed needs.

The former can include any number of additional meta-features that might be needed to fulfil new analytic requirements.

The latter, being designed as a globally shared and curated knowledge base, expects to grow as the information security field advances and its philosophy whitepaper provides awfully specific rules and processes to incorporate new ideas or to update existing ones.

The CKC, instead, lacks this kind of expandability, and this is one of its greatest shortcomings.

This flaw in its approach has some serious consequences when paradigm shifts in technology, such as the massive shift towards cloud services that has drastically changed the information security landscape, happen (see section 3.1.5).

An example of this is the following.
A perimeter-based approach such as that of the Cyber Kill Chain can hardly adapt to the off-premises, multi-tenant-based reality of IaaS, PaaS, and SaaS where perimeter boundaries become more and more blurred and counterintuitive. This is one of the reasons why the CKC struggles, when used on its own, to describe cloud-related malicious activities.
MITRE's framework, instead, adapted to this shift quite seamlessly by introducing an alternative to the Enterprise ATT&CK Matrix called the Cloud ATT&CK Matrix (see Appendix B).
This new Matrix contains all the TTPs, groups, and mitigations relevant to cloud environments and was created using the very rules of the framework that specify object structures and their relationships.

The Diamond Model, in this case, does not even need to be expanded, for its event atomic element, the existing core and meta-features, and the analytic processes are so essential and versatile that they can be directly adapted to cloud environments.

4.2.3.3 VERSATILITY

The very reasons that make the Diamond Model easily expandable also make it versatile.
It was designed to be a cognitive and analytic, as flexible as possible, tool.
Its focus is nurturing a healthy and well-structured hypothesis life cycle and that can only be achieved by giving analysts freedom, and that very freedom is granted by the atomic nature of its core elements. **The DM is likely the most versatile of the three.**

The Cyber Kill Chain, instead, is the least ductile.

To illustrate this concept the project will focus on the phases of the CKC and the tactics of MITRE ATT&CK, their respective topmost entities (see table 6).

The CKC approach requires analysts to work with fixed, sequential and ordered phases. There must be seven phases, they cannot be changed, they must happen in a precise order and direction.
e.g. all seven of them are required and they always go from reconnaissance to actions on objectives.

On the other hand, the ATT&CK framework does not require its topmost objects, the tactics, to be used in a strict order.
MITRE's tactics are not required to happen in a precise sequence for they are just descriptions of documented adversarial behaviour, not necessarily links of a chain. The different assumption here is that attackers can pursue any tactic at any time.
A real-world scenario to better illustrate this concept:

1. We run an organisation that heavily relies on AWS services living inside a VPC[90] with multiple subnets[91].
2. An adversary gains a foothold into a host of a specific subnet of our AWS VPC.
3. He maps the subnet to find other hosts (Discovery tactic)[92].
4. A new host is found, the attacker moves to it (Lateral Movement tactic)[93].
5. On one of these new hosts, the system and network can be mapped once more (Discovery tactic again). A second NIC is found, leading into a new subnet.
6. The hostile actor can now pursue the Lateral Movement tactic again, moving to the newfound subnet.

In steps three to six, the adversary is switching back and forth between tactics, not following a strict sequence.

The main differences between the three models have just been presented, the project will now further highlight them, by illustrating how the models can be integrated, complementing each other's gaps and weaknesses.

4.3 COMPLEMENTARY USE

This section will exemplify, using a real-world scenario, how the different models can be used in an integrated way, strengthening each other.

The scenario will mimic a practical situation, using realistic code snippets, email headers, IPs, services, and entities.
All the above, while realistic, are completely fictitious, purposefully crafted for the sake of this dissertation.

4.3.1 REAL WORLD SCENARIO

'BigPharmaAI' is a large enterprise based in the United Kingdom, it is specialised in using artificial intelligence to conduct pharmaceutical research.
It is a huge multinational corporation, with its own SOC and a large team of information security analysts.

One day their SIEM lights up like a Christmas tree, apparently, there is an ongoing spearphishing campaign targeting the company.
Fraudulent emails containing unsafe links are reaching its employees working on the R&D of a vaccine for a new flu strain.
These emails all share similar headers, such as the one presented below, parts of which are purposefully highlighted for later referencing.

```
Received: from [76.211.4.202] ([76.211.4.202:52123]
helo=sender.freemail.com) by mail44.sender (envelope-from
<mail-salesforce@freemail.com>) (ecelerity 2.2.2.45 r(34222M))
with ESMTPS (cipher=AES256-SHA) id19/15-58796-2E6D1A23; Mon,
15 Mar 2021 08:50:45 +0000

Received: from mallory.dk ([58.124.222.101]) by
mail.freemail.com (mrgmxdk156) with ESMTPA (Exim 4.63) id
0Lao5G-1VDXgv3QRT-00kOPw for <
john.doe@researchers.bigpharmaai.com >; Mon, 15 Mar 2021
08:50:45 +0000

From: Salesforce Accounting Team <hostmaster@mail-
salesforce.com>

To: john.doe@researchers.bigpharmaai.com

Subject: [IMPORTANT]Invoice confirmation

Thread-Topic: [IMPORTANT]Invoice confirmation

Thread-Index: Acqvbp3C811djHLbQ9eTGDmyBL925w==

Sender: "mail-salesforce@freemail.com" <mail-
salesforce@freemail.com>

Date: Mon, 15 Mar 2021 08:50:45 +0000

Message-ID: <fbb225f033ba3043569af9f230ddc4db@mallory.dk>

Reply-To: Salesforce Accounting Team <accounting@mail-
salesforce.com>

Content-Language: en-GB

X-MS-Exchange-Organization-AuthAs: Anonymous

X-MS-Has-Attach: No

X-MS-TNEF-Correlator: 74726f6c6c

X-Mailer: mlx 2.2.9

Content-Type: multipart/alternative;
boundary="_000_8a648fa1a6b71b2694fa19f39ceb8f5bmallorydk_"
MIME-Version: 1.0
```

The mail, claiming to come from Salesforce, a reputable CRM solutions vendor, contains a short text “The invoice can be found here, please confirm its contents” and a link to http://mail-salesforce.ru/mail/auth.php.
Both the senders’ identity claims and the destination link domains are clearly malicious.

The dissertation will now illustrate how the three models can be leveraged concurrently to produce sharper cyber threat intelligence and an overall better protection of the organisation. The focus will be on the analytic process itself, not on the defensive response that stems from it.

At a top level, ‘BigPharmaAI’ analysts using the Cyber Kill Chain will identify this activity as an attempt at the Delivery phase.

Additionally, based on their expertise and prior knowledge of similar attacks, they might also be able to backtrack and understand what the Reconnaissance and Weaponisation phases might have been like and what subsequent phases might look like. (See section 3.1.4.2)

The MITRE ATT&CK framework helps analysts better describe adversarial behaviour.
They can catalogue these emails as an adversary trying to achieve the tactical goal of *Initial access*[94], by using the *Phishing* technique[95], and in particular the *Spearphishing Link* sub-technique[96].
By consulting the framework, in an automated or manual way, they will also be able to:

- Identify procedures usually linked to these (sub-)techniques. (See section 3.2.2.5)
- Identify which known groups make use of such procedures. (See section 3.2.2.6)
- Run an immediate defensive gap assessment by identifying the most common mitigations available to prevent these attacks from succeeding, such as User Training[64] and Restricting Web-Based Content[97]. (See section 2.2.2.3)

The Diamond Model can be used to guide the analysts' thought process.
Its event atomic element can already be populated, both with the data from the attack itself and with the intelligence generated by the other two models:

- The event is the phishing email attack.
- The elements highlighted in grey in the email headers are potential capability indicators, while the blue ones are infrastructure-related indicators.
- The Kill Chain phase and the MITRE tactic can be mapped to the Phase meta-feature of the Diamond event.
- The MITRE techniques and sub-techniques can be added to the methodology meta-feature.
- The defenders also know the victim, as well, their organisation: BigPharmaAI and its vaccine R&D department.

Figure 17 represents this Diamond Model event.

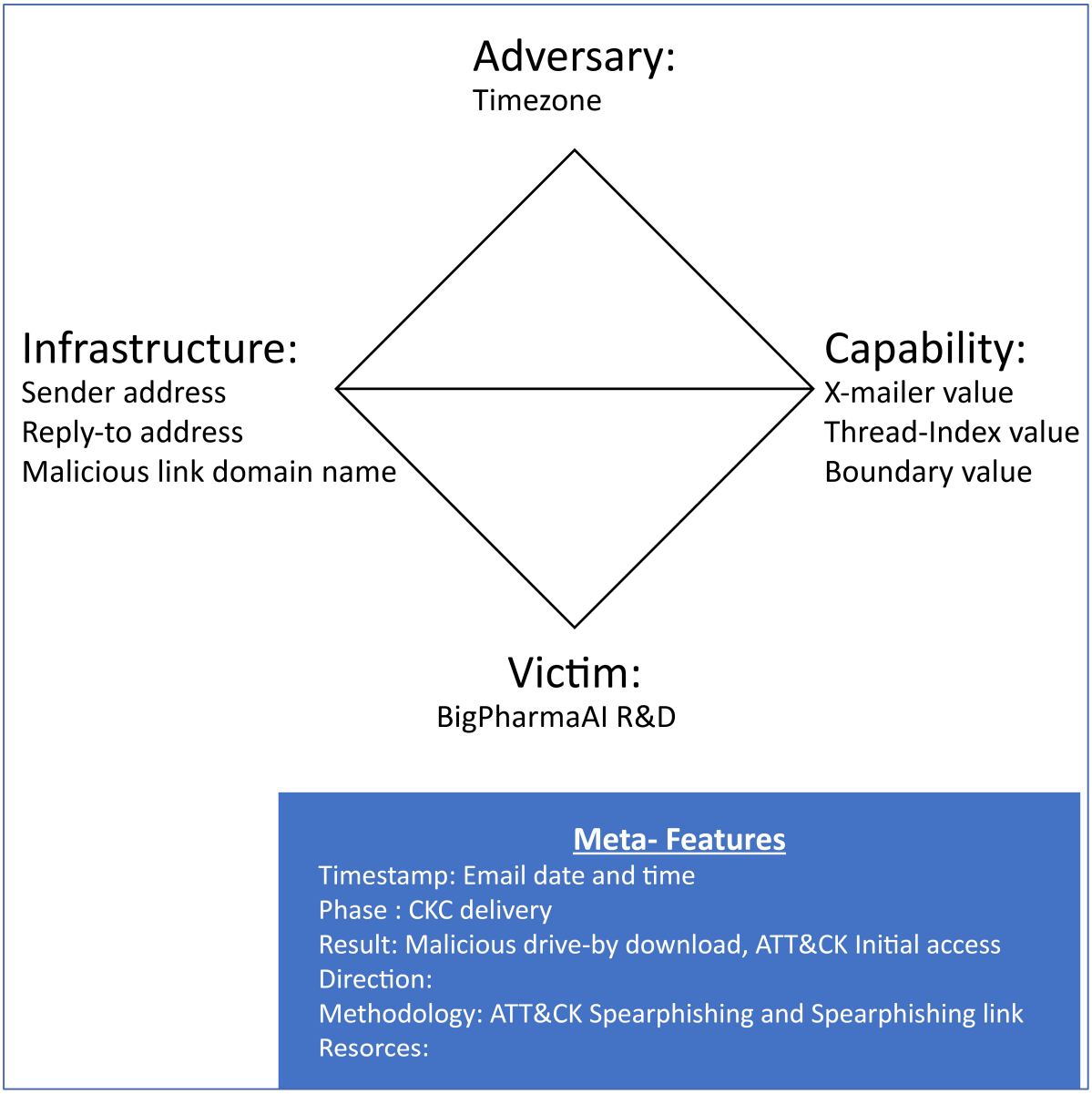

Figure 17 - Diamond Model with ATT&CK and CKC Meta-features

Starting from this diamond, analysts can:

- attempt to fill in, or generate hypotheses about, other details of the core-features or meta-features of the event (e.g. resources and/or expertise required to perform this malicious activity).
- cluster this event with others in pre-existing groups based on their similarities.
- create new clustering rules and groups for future reference if no similarities are found.
- try to understand the identity of the adversary - the threat actor or threat customer – using analytic pivoting and/or thanks to clustering.
- create activity threads and activity-attack graphs.

The analysts of BigPharmaAi, in the scenario, recognise several similarities with previously recorded activities, allowing them to group this event in a pre-existing cluster due to the correspondences of *IP ranges*, *domains*, *time zone*, *mailing client*, *victimology* (see section 3.3.4.4).
To the analysts' knowledge, events in this cluster are, with probable[39] confidence, attributable to a specific actor, APT29[98].

These analytic results generated by applying the Diamond Model, with their associated confidence level, can be fed back to the other two models to produce additional intelligence.

For example, analysts are now able to query the MITRE framework with the supplementary knowledge of the supposed threat actor behind the attack.
This allows them to check the typical techniques, procedures, and software employed by APT29 to update their defences accordingly.

Now that the threat actor and its TTPs are known, analysts can once again resort to the Cyber Kill Chain and, in the light of this newly acquired intelligence, revise their original reconstruction of the kill chain and synthesis of future attack steps.

This shows how each model can produce a different level and type of intelligence on adversarial behaviour that can be used as input for the other models in a virtuous circle. Every time new intelligence is generated within this circle, a new re-evaluation of the defences of the organisation can be performed, allowing a better assessment of defensive gaps and refinement of the security requirements and related spending plans.

It is worth mentioning that this re-evaluation '*can*' be performed; its actual frequency depends on several different factors such as the organization's risk appetite, cost/benefit analyses, time constraints, regulatory/normative requirements, and ISMS policies[10].

5 CONCLUSIONS

To conclude, this thesis will refer to its objectives, stated in section 1.1.

The **first objective** was to explain why analysing intrusions from advanced actors is a fundamental task in the Information Security field.
This has been done in chapter 2, where a common conceptual ground has been established and the complexities of malicious activities from APTs and their high-profile targets have been illustrated.

The **second objective** of the thesis was to thoroughly analyse the three models one by one, explaining how they work and what makes them unique.
This was done in chapter 3, where every model was reviewed and described in detail within its own sub-chapter. A *"key-points and additional considerations"* section exists for each model, providing a summary of all its qualities and features.

The **third objective** of this dissertation was to propose a method to compare the models, it was provided in section 4.1.

It was first applied in section 4.2 to show how different the CKC, DM, and ATT&CK approaches are, both in their purposes, abstraction levels, and design.
It was also applied in section 4.3, where a real-world scenario was used to illustrate how the three models could be used concurrently, in an extremely beneficial and seamless integration.
These two applications of the comparative method fulfilled the **fourth objective**, the comparison of the models.

From those objectives stems the very conclusion of this thesis:
The three models can be harmoniously and advantageously integrated despite being conceived by different authors to serve different purposes.

Lockheed Martin's Cyber Kill Chain will give analysts a top-level method to describe adversarial attack patterns.
MITRE's ATT&CK framework will grant them access to a globally shared knowledge base of documented tactics, techniques, procedures, threat groups, and mitigations and their relationships; all in the form of a structured ontology, accessible manually or via automated tools.
The Diamond Model will enhance defenders' analytic tradecraft, structuring and streamlining their cognitive workflow, by applying analytic pivoting techniques and clustering, game, and graph theories.

When properly integrated, the three models synergise, and the analytic prowess of the whole is greater than the sum of its parts.

6 REFERENCES

[1] N. Machiavelli, *Il principe*. Milan, Italy: Mondadori, 2019.

[2] "Learning the Associations of MITRE ATT & CK Adversarial Techniques." [Online]. Available: https://ieeexplore.ieee.org/document/9162207. [Accessed: 03-Mar-2021]

[3] T. Cruz and P. Simoes, *ECCWS 2019 - Proceedings of the 18th European Conference on Cyber Warfare and Security*. ACPIL, 2019 [Online]. Available: https://play.google.com/store/books/details?id=o655xgEACAAJ

[4] "Script kiddie," *Kaspersky*. [Online]. Available: https://encyclopedia.kaspersky.com/glossary/script-kiddie/. [Accessed: 12-Feb-2021]

[5] "Hacktivism 101: A Brief History and Timeline of Notable Incidents," *Trendmicro*. [Online]. Available: https://www.trendmicro.com/vinfo/pl/security/news/cyber-attacks/hacktivism-101-a-brief-history-of-notable-incidents. [Accessed: 12-Feb-2021]

[6] I. Periodicals, "IEEE REFERENCE GUIDE." [Online]. Available: https://ieeeauthorcenter.ieee.org/wp-content/uploads/IEEE-Reference-Guide.pdf. [Accessed: 15-Mar-2021]

[7] "National Information Assurance Glossary," *DNI*, 26-Apr-2010. [Online]. Available: https://www.dni.gov/files/NCSC/documents/nittf/CNSSI-4009_National_Information_Assurance.pdf. [Accessed: 15-Jan-2021]

[8] "Operational technologies," *NCSC*. [Online]. Available: https://www.ncsc.gov.uk/guidance/operational-technologies. [Accessed: 22-Feb-2021]

[9] "The Human Factor in IT Security: How Employees are Making Businesses Vulnerable from Within." [Online]. Available: https://www.kaspersky.com/blog/the-human-factor-in-it-security/. [Accessed: 22-Feb-2021]

[10] "ISO/IEC 27001," *ISO*, 2020. [Online]. Available: https://www.iso.org/isoiec-27001-information-security.html. [Accessed: 12-Feb-2021]

[11] "Information Security Resources," *SANS*. [Online]. Available: https://www.sans.org/information-security. [Accessed: 12-Feb-2021]

[12] J. Saltzer and M. Schroeder, "The Protection of Information in Computer Systems" [Online]. Available: http://web.mit.edu/Saltzer/www/publications/protection/. [Accessed: 12-Feb-2021]

[13] B. Binde and R. McRee, "Assessing Outbound Traffic to Uncover Advanced Persistent Threat," *SANS*, Mat 5, 2011. [Online]. Available: https://www.sans.edu/student-files/projects/JWP-Binde-McRee-OConnor.pdf. [Accessed: 12-Feb-2021]

[14] "Regin Whitepaper," *Kaspersky*, 24-Nov-2014. [Online]. Available: https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/03/08070305/Kaspersky_Lab_whitepaper_Regin_platform_eng.pdf. [Accessed: 12-Feb-2021]

[15] "Regin: Top-tier espionage tool enables stealthy surveillance," *Broadcom/Symantec*, 27-Aug-2015. [Online]. Available: https://docs.broadcom.com/doc/regin-top-tier-espionage-tool-15-en. [Accessed: 12-Feb-2021]

[16] "Website," *FireEye*. [Online]. Available: https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf. [Accessed: 12-Feb-2021]

[17] "APT1 Indicators," *FireEye*. [Online]. Available: https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report-appendix.zip. [Accessed: 12-Feb-2021]

[18] "What Is an Advanced Persistent Threat (APT)?" [Online]. Available: https://www.cisco.com/c/en/us/products/security/advanced-persistent-threat.html. [Accessed: 12-Feb-2021]

[19] "Advanced Persistent Threats: A Symantec Perspective," *Symantec/Broadcom*. [Online]. Available: http://index-of.es/Varios/b-advanced_persistent_threats_WP_21215957.en-us.pdf. [Accessed: 12-Feb-2021]

[20] "FireEye Threat Intelligence Follow the Money: Dissecting the Operations of FIN6," *FireEye*. [Online]. Available: https://www.fireeye.com/products/financial-services/rpt-fin6.html. [Accessed: 12-Feb-2021]

[21] "[Report] M-Trends 2020," *FireEye*. [Online]. Available: https://content.fireeye.com/m-trends/rpt-m-trends-2020. [Accessed: 12-Feb-2021]

[22] "Treasury Sanctions North Korean State-Sponsored Malicious Cyber Groups," *U.S. DEPARTMENT OF THE TREASURY*. [Online]. Available: https://home.treasury.gov/news/press-releases/sm774. [Accessed: 12-Feb-2021]

[23] "U.S. Charges Russian GRU Officers with International Hacking and Related Influence and Disinformation Operations," *U.S. DEPARTMENT OF JUSTICE*, 04-Oct-2018. [Online]. Available: https://www.justice.gov/opa/pr/us-charges-russian-gru-officers-international-hacking-and-related-influence-and. [Accessed: 12-Feb-2021]

[24] "Advisory: APT29 targets COVID-19 vaccine development," *U.K. NCSC*. [Online]. Available: https://www.ncsc.gov.uk/news/advisory-apt29-targets-covid-19-vaccine-development. [Accessed: 12-Feb-2021]

[25] "Coronavirus," *U.S. CISA*. [Online]. Available: https://www.cisa.gov/coronavirus. [Accessed: 12-Feb-2021]

[26] A. Conway, "New data from Microsoft shows how the pandemic is accelerating the digital transformation of cyber-security - Microsoft Security," *Microsoft Security*, 19-Aug-2020. [Online]. Available: https://www.microsoft.com/security/blog/2020/08/19/microsoft-shows-pandemic-accelerating-transformation-cyber-security/. [Accessed: 12-Feb-2021]

[27] "Advanced persistent threats: minimising the damage," *Network Security*, vol. 2014, no. 4, pp. 5–9, Apr. 2014, doi: 10.1016/S1353-4858(14)70040-6. [Online]. Available: http://dx.doi.org/10.1016/S1353-4858(14)70040-6. [Accessed: 13-Feb-2021]

[28] "[No title]." [Online]. Available: https://research.nccgroup.com/wp-content/uploads/2020/07/ncc-group-whitepaper-soc-maturity-and-capability.pdf. [Accessed: 01-Mar-2021]

[29] N. Tsagourias and M. Farrell, "Cyber Attribution: Technical and Legal Approaches and Challenges," *European Journal of International Law*, Aug. 2020, doi: 10.1093/ejil/chaa057. [Online]. Available: https://academic.oup.com/ejil/article-abstract/31/3/941/5897247?redirectedFrom=fulltext. [Accessed: 13-Feb-2021]

[30] G. B. Moskowitz, *Social Cognition: Understanding Self and Others*. Guilford Publications, 2013 [Online]. Available: https://play.google.com/store/books/details?id=6ABPAgAAQBAJ

[31] S. M. Kassin, S. Fein, and H. R. Markus, *Social Psychology*. Wadsworth, 2011 [Online]. Available: https://play.google.com/store/books/details?id=M51sRAAACAAJ

[32] J. Rollins, "U.S.-China Cyber Agreement [October 16, 2015]," Library of Congress. Congressional Research Service, CRS Insight, IN10376, Oct. 2015 [Online]. Available: https://www.hsdl.org/?abstract&did=. [Accessed: 27-Feb-2021]

[33] "U.S. Charges Five Chinese Military Hackers For Cyber Espionage Against U.S. Corporations And A Labor Organization For Commercial Advantage," 14-Jul-2015. [Online]. Available: https://www.justice.gov/usao-wdpa/pr/us-charges-five-chinese-military-hackers-cyber-espionage-against-us-corporations-and. [Accessed: 27-Feb-2021]

[34] C. N. N. Tal Kopan, "White House readies cyber sanctions against China ahead of state visit," 31-Aug-2015. [Online]. Available: https://www.cnn.com/2015/08/31/politics/china-sanctions-cybersecurity-president-obama/index.html. [Accessed: 27-Feb-2021]

[35] T. Steffens, *Attribution of Advanced Persistent Threats: How to Identify the Actors Behind Cyber-Espionage*. Springer Nature, 2020 [Online]. Available:

https://books.google.com/books/about/Attribution_of_Advanced_Persistent_Threa.html?hl=&id=6FryDwAAQBAJ

[36] M. Berninger, “Going ATOMIC: Clustering and Associating Attacker Activity at Scale.” [Online]. Available: https://www.fireeye.com/blog/threat-research/2019/03/clustering-and-associating-attacker-activity-at-scale.html. [Accessed: 25-Feb-2021]

[37] LII, “Beyond a reasonable doubt.” [Online]. Available: https://www.law.cornell.edu/wex/beyond_a_reasonable_doubt. [Accessed: 16-Mar-2021]

[38] LII, “Burden of proof.” [Online]. Available: https://www.law.cornell.edu/wex/burden_of_proof. [Accessed: 16-Mar-2021]

[39] S. Kent, *Sherman Kent and the Board of National Estimates: Collected Essays*. 1994 [Online]. Available: https://play.google.com/store/books/details?id=ypiz6KHttqEC

[40] The Economist, “War in the fifth domain,” 01-Jul-2010. [Online]. Available: https://www.economist.com/briefing/2010/07/01/war-in-the-fifth-domain. [Accessed: 13-Feb-2021]

[41] “On Incident Handling and Response: A state-of-the-art approach,” *Comput. Secur.*, vol. 25, no. 5, pp. 351–370, Jul. 2006, doi: 10.1016/j.cose.2005.09.006. [Online]. Available: http://dx.doi.org/10.1016/j.cose.2005.09.006. [Accessed: 13-Feb-2021]

[42] E. Hutchins, M. Cloppert, and R. Amin, “Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains.” [Online]. Available: https://www.lockheedmartin.com/content/dam/lockheed-martin/rms/documents/cyber/LM-White-Paper-Intel-Driven-Defense.pdf. [Accessed: 13-Feb-2021]

[43] “DYNAMIC TARGETING AND THE TASKING PROCESS,” *U.S. Air Force Doctrine*. [Online]. Available: https://www.doctrine.af.mil/Portals/61/documents/Annex_3-60/3-60-D17-Target-Dynamic-Task.pdf. [Accessed: 13-Feb-2021]

[44] “Dictionary of Military and Associated Terms.” [Online]. Available: https://fas.org/irp/doddir/dod/jp1_02.pdf. [Accessed: 13-Feb-2021]

[45] Interference Security, “ICMP Reverse Shell.” [Online]. Available: https://resources.infosecinstitute.com/topic/icmp-reverse-shell/. [Accessed: 13-Feb-2021]

[46] “U.S. Department of Defense Joint publication - Information Operations.” [Online]. Available: https://fas.org/irp/doddir/dod/jp3_13.pdf. [Accessed: 13-Feb-2021]

[47] G. Engel, “Deconstructing The Cyber Kill Chain,” *Dark Reading*, 18-Nov-2014. [Online]. Available: https://www.darkreading.com/attacks-breaches/deconstructing-the-cyber-kill-chain/a/d-id/1317542. [Accessed: 13-Feb-2021]

[48] P. Reidy, “Combating the insider threat at the FBI,” in *BlackHat 2013* [Online]. Available: https://media.blackhat.com/us-13/US-13-Reidy-Combating-the-Insider-Threat-At-The-FBI-Slides.pdf. [Accessed: 13-Feb-2021]

[49] “Cyber Threat Intelligence Reports,” *FireEye*. [Online]. Available: https://www.fireeye.com/current-threats/threat-intelligence-reports.html. [Accessed: 13-Feb-2021]

[50] M. Laliberte, “A Twist On The Cyber Kill Chain: Defending Against A JavaScript Malware Attack,” *Dark Reading*, 21-Sep-2016. [Online]. Available: https://www.darkreading.com/attacks-breaches/a-twist-on-the-cyber-kill-chain-defending-against-a-javascript-malware-attack/a/d-id/1326952. [Accessed: 13-Feb-2021]

[51] B. E. Strom *et al.*, “Finding Cyber Threats with ATT&CK-Based Analytics,” *MITRE Technical papers*, Jul. 2017 [Online]. Available: https://www.mitre.org/sites/default/files/publications/16-3713-finding-cyber-threats%20with%20att%26ck-based-analytics.pdf. [Accessed: 13-Feb-2021]

[52] B. E. Strom, A. Applebaum, D. P. Miller, K. C. Nickels, A. G. Pennington, and C. B. Thomas, “ATTACK design and philosophy,” *MITRE*. [Online]. Available: https://attack.mitre.org/docs/ATTACK_Design_and_Philosophy_March_2020.pdf. [Accessed: 13-Feb-2021]

[53] “Remote Services,” *MITRE*. [Online]. Available: https://attack.mitre.org/techniques/T1021/. [Accessed: 13-Feb-2021]

[54] “Cobalt Strike,” *MITRE*. [Online]. Available: https://attack.mitre.org/software/S0154. [Accessed: 13-Feb-2021]

[55] “Cobalt strike - Advanced threat tactics for penetration testers,” *Cobalt strike*. [Online]. Available: https://cobaltstrike.com/downloads/csmanual38.pdf. [Accessed: 13-Feb-2021]

[56] “Rocke,” *MITRE*. [Online]. Available: https://attack.mitre.org/groups/G0106. [Accessed: 13-Feb-2021]

[57] W. Largent, “Rocke: The Champion of Monero Miners.” [Online]. Available: http://blog.talosintelligence.com/2018/08/rocke-champion-of-monero-miners.html. [Accessed: 13-Feb-2021]

[58] “Magic Hound.” [Online]. Available: https://attack.mitre.org/groups/G0059/. [Accessed: 13-Feb-2021]

[59] B. Lee and R. Falcone, “Magic Hound Campaign Attacks Saudi Targets,” 15-Feb-2017. [Online]. Available: https://unit42.paloaltonetworks.com/unit42-magic-hound-campaign-attacks-saudi-targets/. [Accessed: 13-Feb-2021]

[60] "APT18," *MITRE*. [Online]. Available: https://attack.mitre.org/groups/G0026/. [Accessed: 13-Feb-2021]

[61] J. Grunzweig, M. Scott, and B. Lee, "New Wekby Attacks Use DNS Requests As Command and Control Mechanism," 24-May-2016. [Online]. Available: https://unit42.paloaltonetworks.com/unit42-new-wekby-attacks-use-dns-requests-as-command-and-control-mechanism/. [Accessed: 13-Feb-2021]

[62] "Application Developer Guidance," *MITRE*. [Online]. Available: https://attack.mitre.org/mitigations/M1013/. [Accessed: 13-Feb-2021]

[63] "Network Segmentation," *MITRE*. [Online]. Available: https://attack.mitre.org/mitigations/M1030/. [Accessed: 13-Feb-2021]

[64] "User Training," *MITRE*. [Online]. Available: https://attack.mitre.org/mitigations/M1017/. [Accessed: 13-Feb-2021]

[65] Editorial Team, "Our Work with the DNC: Setting the record straight," 05-Jun-2020. [Online]. Available: https://www.crowdstrike.com/blog/bears-midst-intrusion-democratic-national-committee/. [Accessed: 13-Feb-2021]

[66] "Contribute." [Online]. Available: https://attack.mitre.org/resources/contribute/. [Accessed: 23-Mar-2021]

[67] "Getting started with ATT&CK," *MITRE*. [Online]. Available: https://www.mitre.org/sites/default/files/publications/mitre-getting-started-with-attack-october-2019.pdf. [Accessed: 13-Feb-2021]

[68] D. Lu, "ATT&CK Structure Part II: From Taxonomy to Ontology," 09-Jul-2019. [Online]. Available: https://www.tripwire.com/state-of-security/mitre-framework/attck-structure-ontology/. [Accessed: 27-Feb-2021]

[69] "Introduction to STIX." [Online]. Available: https://oasis-open.github.io/cti-documentation/stix/intro.html. [Accessed: 27-Feb-2021]

[70] "ATT&CK™ content available in STIX™ 2.0 via public TAXII™ 2.0 server," May 2018 [Online]. Available: https://www.mitre.org/capabilities/cybersecurity/overview/cybersecurity-blog/attck%E2%84%A2-content-available-in-stix%E2%84%A2-20-via. [Accessed: 03-Mar-2021]

[71] "OWL - Semantic Web Standards." [Online]. Available: https://www.w3.org/OWL/. [Accessed: 27-Feb-2021]

[72] "MITRE ATT&CK Dashboard." [Online]. Available: https://www.fireeye.com/mandiant/security-validation/mitre-attack-dashboard.html. [Accessed: 13-Feb-2021]

[73] M. Sentonas, "CrowdStrike Falcon Dominance Evident in MITRE ATT&CK Evaluation," 24-Apr-2020. [Online]. Available: https://www.crowdstrike.com/blog/crowdstrike-falcon-mitre-attack-evaluation-results-second-iteration/. [Accessed: 13-Feb-2021]

[74] "ATT&CK in MDR services," *Kaspersky*. [Online]. Available: https://www.kaspersky.com/enterprise-security/mitre/mdr. [Accessed: 13-Feb-2021]

[75] S. Caltagirone, "A summary of the Diamond Model of intrusion analysis," *Active response*. [Online]. Available: http://www.activeresponse.org/wp-content/uploads/2013/07/diamond_summary.pdf. [Accessed: 13-Feb-2021]

[76] D. Ross, "Game Theory," *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, 2019 [Online]. Available: https://plato.stanford.edu/archives/win2019/entries/game-theory/

[77] Contributors to Wikimedia projects, "Cluster analysis," 21-May-2004. [Online]. Available: https://en.wikipedia.org/wiki/Cluster_analysis. [Accessed: 28-Feb-2021]

[78] S. Caltagirone, A. Pendergast, and C. Betz, "The Diamond Model of Intrusion Analysis." [Online]. Available: https://www.activeresponse.org/wp-content/uploads/2013/07/diamond.pdf. [Accessed: 13-Feb-2021]

[79] A. Ahmad, J. Webb, K. C. Desouza, and J. Boorman, "Strategically-Motivated Advanced Persistent Threat: Definition, Process, Tactics and a Disinformation Model of Counterattack," *Comput. Secur.*, vol. 86, Jul. 2019, doi: 10.1016/j.cose.2019.07.001. [Online]. Available: https://www.researchgate.net/publication/334274476_Strategically-Motivated_Advanced_Persistent_Threat_Definition_Process_Tactics_and_a_Disinformation_Model_of_Counterattack. [Accessed: 13-Feb-2021]

[80] B. Schneier, "Attack Trees," *Schneier on Security*. [Online]. Available: https://www.schneier.com/academic/archives/1999/12/attack_trees.html. [Accessed: 13-Feb-2021]

[81] C. Stoll, "Stalking the wily hacker." [Online]. Available: https://dl.acm.org/doi/10.1145/42411.42412. [Accessed: 13-Feb-2021]

[82] R. J. Heuer, *Psychology of Intelligence Analysis*. Pickle Partners Publishing, 2020 [Online]. Available: https://play.google.com/store/books/details?id=_oLUDwAAQBAJ

[83] R. H. Pherson and R. J. Heuer Jr., *Structured Analytic Techniques for Intelligence Analysis*. CQ Press, 2020 [Online]. Available: https://play.google.com/store/books/details?id=R6qiDwAAQBAJ

[84] R. J. Heuer, "The Evolution of Structured Analytic Techniques" [Online]. Available: https://www.e-education.psu.edu/geog885/sites/www.e-education.psu.edu.geog885/files/file/Evolution_SAT_Heuer.pdf. [Accessed: 19-Feb-2021]

[85] L. Obrst, P. Chase, and R. Markeloff, "Developing an Ontology of the Cyber Security Domain." [Online]. Available: http://ceur-ws.org/Vol-966/STIDS2012_T06_ObrstEtAl_CyberOntology.pdf. [Accessed: 13-Feb-2021]

[86] "Abuse Elevation Control Mechanism," *Mitre*. [Online]. Available: https://attack.mitre.org/techniques/T1548/. [Accessed: 20-Feb-2021]

[87] "Abuse elevation control mechanism: Setuid and setgid," *MITRE*. [Online]. Available: https://attack.mitre.org/techniques/T1548/001/. [Accessed: 20-Feb-2021]

[88] "CVE - common vulnerabilities and Exposures (CVE)," *MITRE*. [Online]. Available: https://cve.mitre.org/. [Accessed: 18-Feb-2021]

[89] "Kaspersky Threats — Threats," *Kaspersky*. [Online]. Available: https://threats.kaspersky.com/en/threat/. [Accessed: 18-Feb-2021]

[90] "What is Amazon VPC?," *AWS*. [Online]. Available: https://docs.aws.amazon.com/vpc/latest/userguide/what-is-amazon-vpc.html. [Accessed: 19-Feb-2021]

[91] "VPCs and subnets." [Online]. Available: https://docs.aws.amazon.com/vpc/latest/userguide/VPC_Subnets.html. [Accessed: 20-Feb-2021]

[92] "Discovery," *MITRE*. [Online]. Available: https://attack.mitre.org/tactics/TA0007/. [Accessed: 19-Feb-2021]

[93] "Lateral Movement," *MITRE*. [Online]. Available: https://attack.mitre.org/tactics/TA0008/. [Accessed: 19-Feb-2021]

[94] "Initial Access," *MITRE*. [Online]. Available: https://attack.mitre.org/tactics/TA0001/. [Accessed: 02-Mar-2021]

[95] "Phishing." [Online]. Available: https://attack.mitre.org/techniques/T1566/. [Accessed: 02-Mar-2021]

[96] "Phishing: Spearphishing Link," *MITRE*. [Online]. Available: https://attack.mitre.org/techniques/T1566/002/. [Accessed: 02-Mar-2021]

[97] "Restrict Web-Based Content," *MITRE*. [Online]. Available: https://attack.mitre.org/mitigations/M1021/. [Accessed: 02-Mar-2021]

[98] “APT29.” [Online]. Available: https://attack.mitre.org/groups/G0016/. [Accessed: 03-Mar-2021]

APPENDIX A - MITRE ATT&CK ENTERPRISE MATRIX

The latest version of the MITRE ATT&CK Enterprise Matrix can always be found here: https://attack.mitre.org/matrices/enterprise/.
The most updated version as of 2021-03-05 is reproduced below, split in five parts for better accessibility and readability.

Reconnaissance	Resource Development	Initial Access
Active Scanning	Acquire Infrastructure	Drive-by Compromise
Gather Victim Host Information	Compromise Accounts	Exploit Public-Facing Application
Gather Victim Identity Information	Compromise Infrastructure	External Remote Services
Gather Victim Network Information	Develop Capabilities	Hardware Additions
Gather Victim Org Information	Establish Accounts	Phishing
Phishing for Information	Obtain Capabilities	Replication Through Removable Media
Search Closed Sources		Supply Chain Compromise
Search Open Technical Databases		Trusted Relationship
Search Open Websites/Domains		Valid Accounts
Search Victim-Owned Websites		

Execution	Persistence	Privilege Escalation
Command and Scripting Interpreter	Account Manipulation	Abuse Elevation Control Mechanism
Exploitation for Client Execution	BITS Jobs	Access Token Manipulation
Inter-Process Communication	Boot or Logon Autostart Execution	Boot or Logon Autostart Execution
Native API	Boot or Logon Initialisation Scripts	Boot or Logon Initialisation Scripts
Scheduled Task/Job	Browser Extensions	Create or Modify System Process
Shared Modules	Compromise Client Software Binary	Domain Policy Modification
Software Deployment Tools	Create Account	Event-Triggered Execution
System Services	Create or Modify System Process	Exploitation for Privilege Escalation
User Execution	Event-Triggered Execution	Hijack Execution Flow
Windows Management Instrumentation	External Remote Services	Process Injection
	Hijack Execution Flow	Scheduled Task/Job
	Implant Container Image	Valid Accounts
	Office Application Startup	
	Pre-OS Boot	
	Scheduled Task/Job	
	Server Software Component	
	Traffic Signalling	
	Valid Accounts	

Defense Evasion	Credential Access	Discovery
Abuse Elevation Control Mechanism	Brute Force	Account Discovery
Access Token Manipulation	Credentials from Password Stores	Application Window Discovery
BITS Jobs	Exploitation for Credential Access	Browser Bookmark Discovery
Deobfuscate/Decode Files or Information	Forced Authentication	Cloud Infrastructure Discovery
Direct Volume Access	Forge Web Credentials	Cloud Service Dashboard
Domain Policy Modification	Input Capture	Cloud Service Discovery
Execution Guardrails	Man-in-the-Middle	Domain Trust Discovery
Exploitation for Defence Evasion	Modify Authentication Process	File and Directory Discovery
File and Directory Permissions Modification	Network Sniffing	Network Service Scanning
Hide Artifacts	OS Credential Dumping	Network Share Discovery
Hijack Execution Flow	Steal Application Access Token	Network Sniffing
Impair Defences	Steal or Forge Kerberos Tickets	Password Policy Discovery
Indicator Removal on Host	Steal Web Session Cookie	Peripheral Device Discovery
Indirect Command Execution	Two-Factor Authentication Interception	Permission Groups Discovery
Masquerading	Unsecured Credentials	Process Discovery
Modify Authentication Process		Query Registry
Modify Cloud Compute Infrastructure		Remote System Discovery
Modify Registry		Software Discovery
Modify System Image		System Information Discovery
Network Boundary Bridging		System Network Configuration Discovery
Obfuscated Files or Information		System Network Connections Discovery
Pre-OS Boot		System Owner/User Discovery
Process Injection		System Service Discovery
Rogue Domain Controller		System Time Discovery
Rootkit		Virtualisation/Sandbox Evasion
Signed Binary Proxy Execution		
Signed Script Proxy Execution		
Subvert Trust Controls		
Template Injection		
Traffic Signalling		
Trusted Developer Utilities Proxy Execution		
Unused/Unsupported Cloud Regions		
Use Alternate Authentication Material		
Valid Accounts		
Virtualisation/Sandbox Evasion		

Weaken Encryption		
XSL Script Processing		

Lateral Movement	Collection	Command and Control
Exploitation of Remote Services	Archive Collected Data	Application Layer Protocol
Internal Spearphishing	Audio Capture	Communication Through Removable Media
Lateral Tool Transfer	Automated Collection	Data Encoding
Remote Service Session Hijacking	Clipboard Data	Data Obfuscation
Remote Services	Data from Cloud Storage Object	Dynamic Resolution
Replication Through Removable Media	Data from Configuration Repository	Encrypted Channel
Software Deployment Tools	Data from Information Repositories	Fallback Channels
Taint Shared Content	Data from Local System	Ingress Tool Transfer
Use Alternate Authentication Material	Data from Network Shared Drive	Multi-Stage Channels
	Data from Removable Media	Non-Application Layer Protocol
	Data Staged	Non-Standard Port
	Email Collection	Protocol Tunnelling
	Input Capture	Proxy
	Man in the Browser	Remote Access Software
	Man-in-the-Middle	Traffic Signalling
	Screen Capture	Web Service
	Video Capture	

Exfiltration	Impact
Automated Exfiltration	Account Access Removal
Data Transfer Size Limits	Data Destruction
Exfiltration Over Alternative Protocol	Data Encrypted for Impact
Exfiltration Over C2 Channel	Data Manipulation
Exfiltration Over Other Network Medium	Defacement
Exfiltration Over Physical Medium	Disk Wipe
Exfiltration Over Web Service	Endpoint Denial of Service
Scheduled Transfer	Firmware Corruption
Transfer Data to Cloud Account	Inhibit System Recovery
	Network Denial of Service
	Resource Hijacking
	Service Stop
	System Shutdown/Reboot

APPENDIX B - MITRE ATT&CK CLOUD MATRIX

The latest version of the MITRE ATT&CK Cloud Matrix can always be found here: https://attack.mitre.org/matrices/enterprise/cloud/.
The most updated version as of 2021-03-05 is reproduced below, split in three parts for better accessibility and readability.

Initial Access	Persistence	Privilege Escalation
Drive-by Compromise	Account Manipulation	Domain Policy Modification
Exploit Public-Facing Application	Create Account	Valid Accounts
Phishing	Implant Container Image	
Trusted Relationship	Office Application Startup	
Valid Accounts	Valid Accounts	

Defense Evasion	Credential Access	Discovery
Domain Policy Modification	Brute Force	Account Discovery
Impair Defenses	Forge Web Credentials	Cloud Infrastructure Discovery
Modify Cloud Compute Infrastructure	Steal Application Access Token	Cloud Service Dashboard
Unused/Unsupported Cloud Regions	Steal Web Session Cookie	Cloud Service Discovery
Use Alternate Authentication Material	Unsecured Credentials	Network Service Scanning
Valid Accounts		Permission Groups Discovery
		Software Discovery
		System Information Discovery
		System Network Connections Discovery

Lateral Movement	Collection	Exfiltration	Impact
Internal Spearphishing	Data from Cloud Storage Object	Transfer Data to Cloud Account	Defacement
Use Alternate Authentication Material	Data from Information Repositories		Endpoint Denial of Service
	Data Staged		Network Denial of Service
	Email Collection		Resource Hijacking

APPENDIX C - DIAMOND MODEL ACTIVITY THREAD MATHEMATICAL DESCRIPTION

This appendix illustrates, verbatim, the mathematical definitions and properties of Activity threads as presented the Diamond Model whitepaper[78]:

"Formally, we can define the activity thread as a directed graph, ***AT***, where AT = (V, A) is an ordered pair such that:

- |V | ≥ 1, there exists at least one event in the thread
- *AT* is a finite graph
- *V* is the set of all events partitioned into sub-sets such that all events in a sub-set share the same adversary and victim and are further partitioned into *p* labeled tuples where *p* is the number of defined phases and each event is placed into the tuple which matches its phase
- *A* is the set of ordered pairs of arcs such that *arc(x, y)* is defined if and only if the adversary successfully executed event y because of event *x* and event *x* directly preceded event *y*
- There can exist more than one arc to any one event. For example, given three events *x*, *y*, and *z* there can exist a path from *x* to *y arc(x, y)* as well as a path from *z* to *y arc(z, y)*.
- There can exist more than one arc from any one event. For example, given three events x, y, and z there can exist a path from *x* to *y arc(x, y)* as well as a path from *x* to *z arc(x, z)*.
- There can exist only one path from one node to another (i.e., each arc ordered pair is unique within the graph). For example, given two events *x* and *y* there can only exist one path from *x* to *y arc(x, y)*.
- Arcs are labeled with a 4-tuple ⟨Confidence, And/Or, Hypothetical/Actual, Provides⟩ where:
  - Confidence: defining the analytic confidence in the existence of a causal relationship between *x* and *y*
  - And/Or: defines whether the path from *x to y* is necessary and required for *y* to be successful (AND) or whether the path is an alternative and optional route to achieve *y* from *x* (OR)
  - – Hypothetical/Actual: distinguishes a hypothesized arc from an actual arc supported by evidence

- – Provides: defining the resources *x* provides to *y* to be successful matching with the requirements listed in the resources event meta-feature⟧

APPENDIX D - DIAMOND MODEL GROUP-RELATED MATHEMATICAL DEFINITIONS

This appendix illustrates, verbatim, the mathematical definitions and properties of Groups (and related elements) as presented the Diamond Model whitepaper[78]:

Analytic Problem	Which events and threads are likely to be conducted by the same adversary who utilizes a certain process ($Process_1$)? (e.g., attribution)
Feature Space	$Infrastructure_{IP}$, $Infrastructure_{Domain}$, $Capability_{MD5}$, $Victim_{IP}$, $Victim_{Organization}$, Methodology, $Process_1$, $Process_2$, $Process_3$
Feature Vector	$\langle \text{Infrastructure}_{IP}, Capability_{MD5}, Process_1 \rangle$
Outcome	All events and threads will be grouped by similarities in infrastructure IP, capability MD5 hash, and a defined adverarsy process $Process_1$

⟦**Activity Group definition**

Formally, we define an activity group, ***AG*** as a set of events and activity threads which share one or more similarities in features or adversary processes:

$$AG = \{et_1, et_2, \ldots, et_n\}$$

Where:

- $n \geq 1$, there must be at least one element in an activity group
- et_n is either: A singular event or an activity thread
- All events or processes in AG share one or more similarities satisfying the activity group creation function used to partition the events and threads⟧

⟦Analytic problem definition

We define an analytic problem, ***PR***, as an intrusion analysis problem statement which requires clustering and classification (i.e., grouping) to address in part or full. ⟧

⟦Feature Space definition

We define the feature space, ***FS***, as the set of all core-, meta-, and sub-features which define events as well as any and all adversary processes.
Further, we define the feature vector to address an analytic problem, FV_{PR}, as:

$$FV_{PR} = \langle\langle f_1,\ w_{f1}\rangle,\langle f_2,\ w_{f2}\rangle,\ldots,\langle f_n,\ w_{fn}\rangle\rangle$$

Where:

- $n \geq 1$, there must be at least one element in the feature vector
- $f_n \in FS$, every feature in the feature vector must exist in the feature space
- $FV \subset FS$, the feature vector is a sub-set of the feature space
- f_n is a necessary element to group events and threads to address the analytic problem *PR*
- $wf_n \in \mathbf{R}$ and $0 < w_{fn} \geq 1$, the weight is a real number which describes the relative importance of f_n to all other f, such that $w=1$ is a feature with the greatest importance⟧

⟦Activity Group creation function definition

Formally, we define an activity group creation function, ***AGC*** as:

$$AGC(P\ R,\ FV_{PR},\ ET) \rightarrow AGS$$

$$AGS = \{AG_1,\ AG_2,\ .\ .\ .\ ,\ AGn\}$$

Where:

- *PR* is a defined analytic problem to be satisfied by the function

- FV_{PR} is the feature vector which satisfies the analytic problem *PR*
- *ET* is the set of all events and threads to be grouped
- *AGC* partitions all elements of the event/thread set *ET* into a set of *n* Activity Groups, *AGS*, based on the feature vector FV_{PR}
- The function comprising *AGC* can operate across all elements within the set *ET* using the features and processes defined in FV*PR*
- *AGS* is the set of activity groups such that each activity group, *AGn*, satisfies the definition of an activity group
- It is possible that the creation function establishes no groups because no similarities exist, and therefore $n \geq 0$