UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Ph.D. Dissertation

# Modeling of Advanced Threat Actors: characterization, categorization and detection

**Author**

Antonio Villalón Huerta

**Ph.D. Supervisors**

Dr. José Ismael Ripoll Ripoll
Dr. Héctor Marco Gisbert

February, 2023

Technical University of Valencia

# *Abstract*

School of Computer Engineering
Department of Computer Engineering

Doctor of Philosophy
Computer Engineering

## Modeling of Advanced Threat Actors: characterization, categorization and detection

Antonio Villalón Huerta

Information and its related technologies are a critical asset to protect for people, organizations and even whole countries. Our dependency on information technologies increases every day, so their security is a key issue for our wellness. The benefits that information technologies provide are questionless, but their usage also presents risks that, linked to our growing dependency on technologies, we must mitigate. Advanced threat actors are mainly categorized in criminal gangs, with an economic goal, and countries, whose goal is to gain superiority in strategic affairs such as commercial or military ones. These actors exploit technologies, particularly cyberspace, to achieve their goals.

This PhD Thesis significantly contributes to advanced threat actors' categorization and to the detection of their hostile activities. The analysis of their features is a must not only to know better these actors and their operations, but also to ease the deployment of countermeasures that increase our security. The detection of these operations is a mandatory first step to neutralize them, so to minimize their impact.

Regarding characterization, this work delves into the analysis of advanced threat actors' tactics and techniques. This analysis is always required for an accurate detection of hostile activities in cyberspace, but in the particular case of advances threat actors, from criminal gangs to nation–states, it is mandatory: their activities are stealthy, as their success in most cases relies on not being detected by the target.

Regarding detection, this work identifies and justifies the key requirements to establish an accurate response capability to face advanced threat actors. In addition, this work defines the tactics to be deployed in Security Operations Centers to optimize their detection and response capabilities. It is important to highlight that these tactics, with a kill–chain arrangement, allow not only this optimization,

but particularly a homogeneous and structured approach, common to all defensive centers.

In my opinion, one of the main bases of my work must be the applicability of its results. For this reason, the analysis of threat actors' tactics and techniques is aligned with the main public framework for this analysis, MITRE ATT&CK. The results and proposals from this research can be directly included in this framework, improving the threat actors' characterization, as well as their cyberspace activities' one. In addition, the proposals to improve these activities' detection are directly applicable both in current Security Operations Centers and in common industry technologies. In this way, I consider that this work significantly improves current analysis and detection capabilities, and at the same time it improves hostile operations' neutralization. These capabilities increase global security for all kind of organizations and, definitely, for our whole society.

Universidad Politécnica de Valencia

# *Resumen*

Escuela Técnica Superior de Ingeniería Informática
Departamento de Informática de Sistemas y Computadores

Doctor en Filosofía
Ingeniería Informática

## Modelado de Actores Hostiles Avanzados: caracterización, categorización y detección

Antonio Villalón Huerta

La información y los sistemas que la tratan son un activo a proteger para personas, organizaciones e incluso países enteros. Nuestra dependencia en las tecnologías de la información es cada día mayor, por lo que su seguridad es clave para nuestro bienestar. Los beneficios que estas tecnologías nos proporcionan son incuestionables, pero su uso también introduce riesgos que ligados a nuestra creciente dependencia de las mismas es necesario mitigar. Los actores hostiles avanzados se categorizan principalmente en grupos criminales que buscan un beneficio económico y en países cuyo objetivo es obtener superioridad en ámbitos estratégicos como el comercial o el militar. Estos actores explotan las tecnologías, y en particular el ciberespacio, para lograr sus objetivos.

La presente tesis doctoral realiza aportaciones significativas a la caracterización de los actores hostiles avanzados y a la detección de sus actividades. El análisis de sus características es básico no sólo para conocer a estos actores y sus operaciones, sino para facilitar el despliegue de contramedidas que incrementen nuestra seguridad. La detección de dichas operaciones es el primer paso necesario para neutralizarlas, y por tanto para minimizar su impacto.

En el ámbito de la caracterización, este trabajo profundiza en el análisis de las tácticas y técnicas de los actores. Dicho análisis siempre es necesario para una correcta detección de las actividades hostiles en el ciberespacio, pero en el caso de los actores avanzados, desde grupos criminales hasta estados, es obligatorio: sus actividades son sigilosas, ya que el éxito de las mismas se basa, en la mayor parte de casos, en no ser detectados por la víctima.

En el ámbito de la detección, este trabajo identifica y justifica los requisitos clave para poder establecer una capacidad adecuada frente a los actores hostiles avanzados. Adicionalmente, proporciona las tácticas que deben ser implementadas en los Centros de Operaciones de Seguridad para optimizar sus capacidades de de-

tección y respuesta. Debemos destacar que estas tácticas, estructuradas en forma de *kill–chain*, permiten no sólo dicha optimización, sino también una aproximación homogénea y estructurada común para todos los centros defensivos.

En mi opinión, una de las bases de mi trabajo debe ser la aplicabilidad de los resultados. Por este motivo, el análisis de tácticas y técnicas de los actores de la amenaza está alineado con el principal marco de trabajo público para dicho análisis, MITRE ATT&CK. Los resultados y propuestas de esta investigación pueden ser directamente incluidos en dicho marco, mejorando así la caracterización de los actores hostiles y de sus actividades en el ciberespacio. Adicionalmente, las propuestas para mejorar la detección de dichas actividades son de aplicación directa tanto en los Centros de Operaciones de Seguridad actuales como en las tecnologías de detección más comunes en la industria. De esta forma, este trabajo mejora de forma significativa las capacidades de análisis y detección actuales, y por tanto mejora a su vez la neutralización de operaciones hostiles. Estas capacidades incrementan la seguridad global de todo tipo de organizaciones y, en definitiva, de nuestra sociedad.

Universitat Politècnica de València

# *Resum*

Escola Tècnica Superior d'Enginyeria Informàtica
Departament d'Informàtica de Sistemes y Computadors

Doctor en Filosofía
Enginyeria Informàtica

## Modelat d'Actors Hostils Avançats: caracterització, categorització i detecció

Antonio Villalón Huerta

La informació i els sistemas que la tracten són un actiu a protegir per a persones, organitzacions i fins i tot països sencers. La nostra dependència en les tecnologies de la informació es cada dia major, i per això la nostra seguretat és clau per al nostre benestar. Els beneficis que aquestes tecnologies ens proporcionen són inqüestionables, però el seu ús també introdueix riscos que, lligats a la nostra creixent dependència de les mateixes és necessari mitigar. Els actors hostils avançats es categoritzen principalment en grups criminals que busquen un benefici econòmic i en països el objectiu dels quals és obtindre superioritat en àmbits estratègics, com ara el comercial o el militar. Aquests actors exploten les tecnologies, i en particular el ciberespai, per a aconseguir els seus objectius.

La present tesi doctoral realitza aportacions significatives a la caracterització dels actors hostils avançats i a la detecció de les seves activitats. L'anàlisi de les seves característiques és bàsic no solament per a conéixer a aquests actors i les seves operacions, sinó per a facilitar el desplegament de contramesures que incrementen la nostra seguretat. La detección de aquestes operacions és el primer pas necessari per a netralitzar-les, i per tant, per a minimitzar el seu impacte.

En l'àmbit de la caracterització, aquest treball aprofundeix en l'anàlisi de lestàctiques i tècniques dels actors. Aquesta anàlisi sempre és necessària per a una correcta detecció de les activitats hostils en el ciberespai, però en el cas dels actors avançats, des de grups criminals fins a estats, és obligatòria: les seves activitats són sigiloses, ja que l'éxit de les mateixes es basa, en la major part de casos, en no ser detectats per la víctima.

En l'àmbit de la detecció, aquest treball identifica i justifica els requisits clau per a poder establir una capacitat adequada front als actors hostils avançats. Adicionalment, proporciona les tàctiques que han de ser implementades en els Centres d'Operacions de Seguretat per a optimitzar les seves capacitats de detecció

i resposta. Hem de destacar que aquestes tàctiques, estructurades en forma de kill-chain, permiteixen no només aquesta optimització, sinò també una aproximació homogènia i estructurada comú per a tots els centres defensius.

En la meva opinio, una de les bases del meu treball ha de ser l'aplicabilitat dels resultats. Per això, l'anàlisi de táctiques i tècniques dels actors de l'amenaça està alineada amb el principal marc públic de treball per a aquesta anàlisi, MITRE ATT&CK. Els resultats i propostes d'aquesta investigació poden ser directament inclosos en aquest marc, millorant així la caracterització dels actors hostils i les seves activitats en el ciberespai. Addicionalment, les propostes per a millorar la detecció d'aquestes activitats són d'aplicació directa tant als Centres d'Operacions de Seguretat actuals com en les tecnologies de detecció més comuns de la industria. D'aquesta forma, aquest treball millora de forma significativa les capacitats d'anàlisi i detecció actuals, i per tant millora alhora la neutralització d'operacions hostils. Aquestes capacitats incrementen la seguretat global de tot tipus d'organitzacions i, en definitiva, de la nostra societat.

# *Acknowledgements*

First of all, I would like to thank my advisors, Ismael Ripoll and Héctor Marco, for their continuous patience and expert advice. 2020 was a bad year, and in the worse days they convinced me to start this hard and funny journey. During this time, they have been the best fellow travelers I could imagine.

Of course, in this acknowledgements section I must thank my parents. Without them, this thesis would have not been finished, and not even begun. Not only for obvious biological reasons and for buying me my first computer many years ago, but particularly because they worked tirelessly to help me earn my MSc degree (and they also forced me to study very hard to get it). So dad and mom, this work is also for you (at least, one copy!).

Last but certainly not least, it would have been impossible for me to finish this thesis without the support of my wife. So special thanks to Raquel for her constant encouragement and for allowing me to stay whole nights and weekends in front of a computer. With two little daughters at home it has been particularly difficult, so thanks++.

To Raquel, and also to our two daughters, Teresa and Lucía, is all this work.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Motivation

Since ancient times, information has been a valuable asset to protect and also a main target to gather. Nowadays not only information, but also information and communication technologies are a valuable asset for people, organizations and societies. For our daily basis, we rely on the interconnection of devices and their accessibility, in what is usually called cyberspace [447]: a global domain within the information environment consisting of the interdependent network of information technology infrastructures and resident data, including the Internet, telecommunications networks, computer systems, and embedded processors and controllers.

The high benefits cyberspace –in fact, technology– has contributed to are questionless, but also are the risks it introduces. Threat actors such as foreign countries, terrorists groups or organized crime, are well aware of the exploitation of cyberspace to achieve their goals, from cyber crime to cyber war. For them, the risks of working on the cyberspace are low, and the achieved benefits are high. In fact, as [130] stated years ago, cyberspace is the new front line, and the Pentagon's term for cyberspace is the fifth domain [140], which represents the battlefield Internet has become. After land, sea, air and space, cyberspace is a domain where hostile operations are performed.

This work focuses on advanced threat actors. They have the knowledge, the resources and the intent to achieve their goals through cyberspace. We must face two kinds of advanced threat actors: nation–state and non–state ones. Nation–state actors perform targeted operations, most of them strategically motivated, while non–state actors have operational goals, most of them economical. When dealing with strategic operations, cyberspace exploitation is the most common activity for nation–state actors. Regarding non–state actors, Human Operated Ransomware is the most common threat to all organizations, including critical infrastructures (where their impact can be very high).

The modeling of advanced threat actors is a key component of counterintelligence activities, so it is a must for the prevention, detection and neutralization of hostile operations. These activities are the core ones for Security Operation Centers (SOC) all around the world. To successfully perform them, it is mandatory to define models to globally describe advanced threat actors, from their sponsors to their lower level indicators. The goal of such models is to help analysts to identify threat actors, their activities, their targets and their tactics, techniques and procedures. Although many efforts in this direction have been done, the problem is not fully addressed and different current approaches must be improved. For example, in the case of tactics and techniques, the MITRE ATT&CK framework is the main reference nowadays, but it presents relevant problems that can hinder the analysts' work, such as a plain structure for all techniques inside each tactic or, more important, the incorrect usage of basic concepts when dealing with tactics such as reconnaissance.

Facing advanced threat actors is not an easy task. A deep knowledge about their goals, tactics and techniques and even about their tools is required. This work delves into the characterization of advanced threat actors by analyzing their tactics and techniques, this is, their modus operandi. Through a better characterization of advanced threat actors and their techniques, the detection of hostile operations is improved. The modeling of techniques provides an in–depth understanding of cyberspace operations and a valuable intelligence which is hard to modify by advanced threat actors. In this way, this dissertation improves the knowledge about advanced threat actors and their activities.

In addition to the modeling of advanced threat actors, to detect and neutralize their activities an accurate monitoring and response approach is also mandatory. From an intelligence perspective, data from multiple sources must be acquired, processed and analyzed in order to establish a monitoring capability. This work not only provides an intelligence–driven approach to the detection of advanced hostile activities, but also a detection and response arrangement of tactics to improve defensive capabilities. These contributions ease the early detection of hostile activities in all kind of organizations, thus improving their global security.

## 1.2 Goals and contributions

The goal of this dissertation is to provide accurate approaches for the analysis of advanced threat actors and their operations. This work must improve the knowledge about elements such as sponsors, goals and objectives, strategies or tools, but particularly about the tactics, techniques and procedures (TTP) executed by these actors. TTP are the key element for the planning and execution of prevention, detection and response activities for two main reasons:

- TTP are the higher level observable to extract from a targeted infrastructure.

Other high level features of a hostile operation, such as strategy or goals, can not be directly observed, so they must be inferred with a given probability.

- TTP are the observable which is harder to modify for hostile actors. Low level indicators, such as IP addresses or network domains, or even tools, can be easily modified by advanced threat actors, so their usefulness for detection is limited.

Being tactics and techniques the most relevant element for cyberspace defense, this work delves into the characterization of advanced threat actors by focusing on their TTP; particularly, three key tactics and their associated techniques in hostile cyberspace operations are dissected: delivery, persistence and attack (this is, operations linked to interruption or manipulation). In the modeling of all of them, this research follows a logical structure that can be easily expanded and adapted to main current frameworks, such as MITRE ATT&CK. In this way, the main results for this work can be directly used in commonly accepted industry standards, providing practical real world results to the community.

Characterization of advanced threat actors is a must for their analysis but, particularly, for their detection. It is important to highlight that detection and response are the final goals of defensive mechanisms, from people to technology. Regarding detection, the arrangement of tactics to enable an accurate detection capability, in the form of kill–chain, is presented: SOC Critical Path. This is a especially relevant contribution, as this novel model provides defensive center not only a set of mandatory tactics, but also a continuous improvement approach to increase global security. In addition, the key requirements for an accurate detection based on tactics and techniques are identified. As stated before, TTP are the most relevant observable to detect hostile activities, and thus to respond to them.

The main contributions of this work are as follows:

- To analyze and dissect key tactics for hostile operations performed by advanced threat actors, providing taxonomies for them. This contribution significantly improves not only the knowledge about advanced threat actors, but also their identification, thus providing accurate detection capabilities.

- To improve the main framework for the analysis of advanced threat actors' tactics and techniques, MITRE ATT&CK, easing not only the detection of these tactics and techniques but also the identification of hostile actors' capabilities.

- To identify the key requirements for an accurate detection of hostile activities regarding advanced threat actors; the stealthiness of these activities makes them harder to detect, so different approaches for an improved detection scheme following common intelligence models are proposed.

- To define the mandatory tactics to detect and respond to security incidents, providing a novel arrangement for them in the form of kill–chain model. This common set of tactics allows defensive centers the planning and continuous improvement of their work.

## 1.3 Thesis outline

In this section the dissertation's outline is presented. This introduction chapter is complemented with background and state of the art sections, in order to clarify concepts about the definition, characterization and detection of advanced threat actors. In further chapters, I delve into the characterization and detection of advanced threat actors, as stated before: analyzing some of their key tactics and techniques and detecting the key requirements for their detection, as well as a novel detection kill–chain model.

In chapter 2 a classification approach for techniques linked to delivery is proposed. Delivery is a key stage for offensive cyber operations because it defines the initial access to a target, this is, the actions that a threat actor performs to gain an initial foothold into a targeted infrastructure. It is the first moment in which a hostile actor and its target establish a real contact. The proposed analysis and classification approaches significantly reduce the amount of effort needed to identify, analyze, and neutralize hostile activities from advanced threat actors, in this case in their initial access stage.

In chapter 3 a taxonomy for Persistence techniques is presented; as with Delivery, Persistence is a key tactic in offensive cyberspace operations: keeping its presence in the targeted infrastructure over time is a main goal for the threat actor. This research allows the detection of novel techniques and the identification of appropriate countermeasures to face them. A novel concept, persistence point, is presented as the basic element for the proposed structure of persistence techniques. The identification of these persistence points is a useful tool for the detection of persistence in novel platforms and technologies. As with the previous tactic, the main goal here is to ease the work of identifying, analyzing and neutralizing hostile activities.

To end with the characterization approaches, in chapter 4 this work goes one step further common complex operations and delves intro destructive and manipulation ones (Cyberspace Attack or Computer Network Attack, CNA), proposing a taxonomy for tactics and techniques for these operations. Cyberspace Attack actions are today a major threat, especially for cyber physical systems. They are usually linked to state-sponsored actors, and they are much less analyzed than Computer Network Exploitation activities (CNE), those regarding intelligence gathering. While in CNE operations the main tactics and techniques are defined and well structured, in CNA there is a lack of such consensuated approaches. This situation hinders the modeling of threat actors, which prevents an accurate definition of counter measures to identify and to neutralize malicious activities. The novel proposal presented here significantly reduces the amount of effort need to identify, analyze, and neutralize advanced threat actors.

After the contributions to the modeling of advanced threat actors, and dealing with the detection of hostile operations, I define a defensive kill chain approach where tactics for teams in charge of cyber defense activities are structured and arranged. This novel model, SOC Critical Path (SCP), is detailed in chapter 5 as a technology–independent approach that provides an arrangement of mandatory steps, in the form of tactics, to be executed by Computer Network Defense teams to detect hostile cyber operations. By adopting this novel model, defensive teams increase the performance and the effectiveness of their capabilities through a common framework that formalizes the steps to follow for the detection and neutralization of threats. This proposal can be used not only to identify detection and response gaps, but also to implement a continuous improvement cycle over time.

To improve the detection capabilities of defensive teams, in chapter 6 the requirements for an effective detection of advanced threats through indicators of compromise are analyzed. Given that current Cyber Threat Intelligence sharing focuses on atomic and computed indicators that are expressed in machine–readable formats and that are easily loaded into security devices, I analyze the benefits and disadvantages of this tactical intelligence sharing. These indicators are easily avoided by advanced threat actors, so their usefulness is very limited, especially in terms of time of life. I analyze why higher level indicators, those linked to operational intelligence and expressed as tactics, techniques and procedures, are not widely used, identifying key requirements for a successful detection, specification and sharing of these indicators. As these operational indicators provide high value intelligence, their detection, specification and sharing improves detection maturity for defensive teams, thus improving the protection of organizations.

After all the main contributions of this work have been presented, in chapter 7 I discuss the different solutions and their role in the characterization, categorization and detection of advanced threat actors. I analyze how this work improves defensive capabilities and helps organizations to prevent, detect and respond to advanced threat actors, thus improving global security for organizations.

Finally, in chapter 8 the main conclusions of this work are presented, particularly those regarding the characterization of advanced threat actors and the detection of their hostile activities. In this final chapter the main conclusions of my contributions are exposed, highlighting that advanced threat actors, both state–sponsored and non–state, are a real risk for most organizations, as cyberspace provides enormous benefits for these actors.

## 1.4 Background and state of the art

### 1.4.1 Information Operations

Information Operations (IO) are defined [447] as the integration of specified capabilities involving information and information systems; more specific definitions refer to actions taken to affect adversary information and information systems while defending ones own information and information systems. IO capabilities are categorized in either core (the ones which are essential to the conduct of IO by providing critical operational effects or preventing the adversary from doing so), supporting (those which provides additional, though less critical, operational effects) or related (those which contribute to the accomplishment of the IO mission).

IO core capabilities are the following ones [425]:

**Electronic Warfare (EW).** Any military action involving the use of electromagnetic and directed energy to dominate the electromagnetic spectrum or to attack the enemy; EW comprises three major areas of activity [238]:

> **Electronic Attack (EA).** Activities involving the use of electromagnetic energy, directed energy, or anti–radiation weapons to attack personnel, facilities, or equipment with the intent of degrading, neutralizing, or destroying enemy combat capability.
>
> **Electronic Protection (EP).** Activities involving passive and active means taken to protect personnel, facilities, and equipment from any effects of friendly or enemy employment of electronic warfare that degrade, neutralize, or destroy friendly combat capabilities.
>
> **Electronic Warfare Support (ES).** Activities involving actions to search for, intercept, identify, and locate or localize sources of intentional and unintentional radiated electromagnetic energy for the purpose of immediate threat recognition, targeting, planning and conduct of future operations.

**Military Deception (MILDEC).** Actions executed to deliberately mislead adversary military decision makers as to friendly military capabilities, intentions, and operations, thereby causing the adversary to take specific actions (or inactions) that will contribute to the accomplishment of the friendly mission.

**Computer Network Operations (CNO).** Capabilities to attack adversary computer networks, defend our own and exploit enemy computers to enable intelligence gathering. CNO comprises three main activities:

> **Computer Network Attack (CNA).** Actions taken through the use of computer networks to disrupt, deny, degrade, or destroy information resident in computers and computer networks, or the computers and networks themselves; from its own definition, CNA operations are those regarding more or less destructive actions against a target.
>
> **Computer Network Defense (CND).** Actions taken through the use of computer networks to protect, monitor, analyze, detect and respond to unauthorized activity in own information systems and computer networks[1].
>
> **Computer Network Exploitation (CNE).** Actions taken to enable operations and intelligence collection capabilities conducted through the

[1] Please note the term "respond" in this context; it could range from a passive information assurance tool, such as a firewall, to an active response against an intrusion.

use of computer networks to gather data from target or adversary automated information systems or networks. As [414] states, CNE is about computer espionage: about gaining access to a computer systems and retrieving information from them.

**Psychological Operations (PSYOP).** Planned operations to convey selected information and indicators to foreign audiences to influence their emotions, motives, objective reasoning, and ultimately the behavior of foreign governments, organizations, groups, and individuals [238]. In modern doctrines, PSYOP is referred as Military Information Support Operations (MISO).

**Operations Security (OPSEC).** The process of identifying essential elements of friendly information and taking measures to mask them from disclosure to adversaries, to determine if friendly essential intelligence can be associated with adversaries operations.

In addition to these core capabilities, IO Supporting Capabilities provide additional operational effects: Information Assurance (IA), Physical Security, Physical Attack, Counterintelligence (CI), and Combat Camera (COMCAM). Related Capabilities of Public Affairs (PA), Civil-Military Operations (CMO), and Defense Support to Public Diplomacy (DSPD) contribute to the accomplishment of the IO mission. With the exception of counterintelligence, an issue that is mandatory when dealing with threat actors and that will be covered in this chapter, the focus of this work are not these additional IO capabilities, either supporting or related, but the core ones.

### 1.4.2 Cyberspace Operations

Cyberspace Operations (CO) are defined as [625] the employment of cyberspace capabilities where the primary purpose is to achieve objectives in or through cyberspace. CO is a broader term than Computer Network Operations, as it comprises not only pure CNO but all the capabilities inside IO in or through cyberspace, such as Electronic Warfare or even Psychological Operations. [450] identifies and defines the following Cyberspace actions:

- **Offensive Cyberspace Operations.**

  **Cyberspace Attack.** Cyberspace attack actions create noticeable denial effects (degradation, disruption or destruction) in cyberspace or manipulation that leads to denial effects in the physical domains. Cyberspace Attack includes actions to deny or manipulate the services of IT infrastructure or the information they handle.

  **Cyberspace Exploitation.** Cyberspace exploitation actions include military intelligence activities, maneuver, information collection, and other enabling actions required to prepare for future military operations. Cyberspace exploitation includes activities to gain intelligence and support operational preparation of the environment for current and future operations through actions such as gaining and maintaining access to networks, systems, and nodes of military value; maneuvering to positions of advantage; and positioning cyberspace capabilities to facilitate follow–on actions.

- **Defensive Cyberspace Operations.**

  **Cyberspace Security.** Cyberspace security actions are taken within protected cyberspace to prevent unauthorized access to, exploitation of, or damage to computers, electronic communications systems, and other IT, including Platform IT, as well as the information contained therein, to ensure its availability, integrity, authentication, confidentiality, and non repudiation. Cyberspace security actions occur in advance of a specific security compromise and protect from threats within cyberspace by reducing or eliminating vulnerabilities that may be exploited by an adversary and/or implementing measures to detect malicious cyberspace activities.

  **Cyberspace Defense.** Cyberspace defense actions are taken within protected cyberspace to defeat specific threats that have breached or are threatening to breach the cyberspace security measures and include actions to detect, characterize, counter, and mitigate threats, including malware or the unauthorized activities of users, and to restore the system to a secure configuration.

Nowadays, most advanced threat actors sponsored by nation–states are able to launch Offensive Cyberspace Operations, which include not only offensive CNO (CNA or CNE), but also other offensive IO operations through cyberspace, such as Electronic Warfare [393] or even PSYOP [159] [155] [317], what literature defines as cyber enabled PSYOP [367].

### 1.4.3 Intelligence

NATO [454] defines intelligence as the product resulting from the directed collection and processing of information regarding the environment and the capabilities and intentions of actors, in order to identify threats and offer opportunities for exploitation by decision–makers; the same work also defines the intelligence cycle as the sequence of activities whereby information is obtained, assembled, converted into intelligence and made available to users. Although there are different versions of this cycle, its approach can be summarized by considering the following five steps:

1. **Direction**. Determination of intelligence requirements, planning the collection effort, issuance of orders and requests to collection agencies and maintenance of a continuous check on the productivity of such agencies.

2. **Collection**. The exploitation of sources by collection agencies and the delivery of the information obtained to the appropriate processing unit for use

in the production of intelligence.

3. **Processing**. The conversion of information into usable data suitable for analysis.

4. **Analysis**. Integration, evaluation, interpretation, etc. of information to turn it into intelligence: a contextualized, coherent whole.

5. **Dissemination**. The timely conveyance of intelligence, in an appropriate form and by any suitable means, to those who need it.

In figure 1.1 the intelligence cycle as exposed in [449] is shown; in this work, the authors define a sixth step: evaluation and feedback –at the centre of the figure–: the disseminated intelligence feeds a new iteration of the cycle.

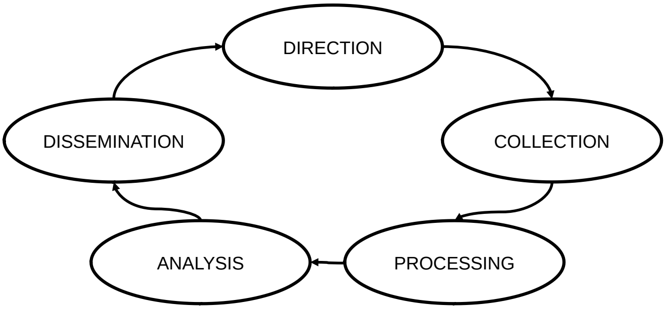

Figure 1.1: Intelligence Cycle

Intelligence (as a product) is regarded as the final result of a set of actions, which are sequentially launched; this set of actions, the intelligence cycle, is a simple explanation of a complex process, which is also defined as intelligence (as a process). It starts when someone (an authority, a government, etc.) has particular information needs in order to make the best decision about a particular subject. At this point the cycle starts first by identifying the requirements and planning the acquisition of the information to be processed and analyzed in order to generate intelligence.

Once planned, the next stage is to acquire information, and this acquisition can be performed through different intelligence collection disciplines [80], commonly referred as “the INTs”; the essential elements of these INTs are not formally defined [139], and neither they are consensuated between authors, but they define the families of sources the information can be gathered from: a simple public website, a satellite, an intercepted artifact, a mole, etc. These intelligence collection disciplines are discussed in this section.

Once the relevant information has been gathered, processing and exploitation turn it into a form suitable for the production of finished intelligence [520]; this stage includes tasks such as decryption, translation or data conversion and, as a part of the cycle, it is a requirement for the next one: analysis, in which the intelligence, the final product, is generated. This analysis must include the information gathered and processed no matter which collection discipline it comes from. In this sense, it is possible to refer to all–source intelligence, defined [38] as the intelligence products, organizations, and activities that incorporate all sources of information and intelligence, including open source information, in the production of intelligence.

Finally, once the intelligence as a product has been generated, it is delivered to and used by the customer, the entity which had the information needs stated before, in a suitable form for its use and by a variety of means. This product will be used to help the decision making process and, possibly, to start a new iteration of the intelligence cycle.

### Intelligence gathering disciplines

As stated before, intelligence collection disciplines are not consensuated between authors, a fact that motivates different discussions. There are five commonly accepted disciplines by the US Intelligence Community [376] [377] [482] [138]: geospatial (formerly imagery) intelligence (GEOINT), signals intelligence (SIGINT), measurement and signatures intelligence (MASINT) –which includes technical intelligence or TECHINT–, human intelligence (HUMINT) and open source intelligence (OSINT).

Imagery intelligence, IMINT, is defined as the technical, geographic, and intelligence information derived through the interpretation or analysis of imagery and collateral material [109], and it is considered inside GEOINT in many works [509] [139], although it is also considered as an independent discipline in many others [235] [110]. Most references consider GEOINT as the integration of IMINT and geospatial information [448] [109]. For this reason, in this work I will deal with GEOINT as a global discipline comprising IMINT. It is important to note that, strictly, no collection system collects GEOINT [137]: geospatial information is collected via IMINT, OSINT, SIGINT, HUMINT, MASINT, etc.

The role of TECHINT, intelligence gathered from the collection, processing, analysis and exploitation of data and information pertaining to foreign equipment and materiel [58], is much more discussed topic. It is considered inside MASINT by the references which identify only five main disciplines and by some US military works. However, it is considered a discipline by itself in many references [110] [305], and other works identify TECHINT as all intelligence gathered from technical sources (vs. human sources) [250] [568] [158] [307]. Other authors, such as [271], differentiate between main sources and smaller sources for intelligence gathering disciplines. A discussion about those disciplines and their consideration can be found at [552]. In addition to these differences, there have been also some efforts to add new intelligence collection disciplines to the list, such as those proposals in [608], [196] or [41], generating even more confusion into the community.

This work will not enter into the discussion about which disciplines have to be considered: I will simply deal with the five well–accepted disciplines. TECHINT will be included inside the MASINT discipline and, in the same way, IMINT will be included inside GEOINT, although highlighting that imagery intelligence plays a key role in the cyber battle space (much more than GEOINT, as cyber is a domain of conflict not directly related to GEO in many cases). In summary, the following five disciplines are considered, without detailing subcategories for the purpose of this work:

**Human Intelligence (HUMINT).** Intelligence collected and provided from human sources [449].

**Geospatial Intelligence (GEOINT).** Intelligence gathered from geospatial data through the application of geospatial techniques and by skilled interpretation, in which the location and movement of activities, events, features and people play a key role [154]. As stated, in this work IMINT is considered as a part of GEOINT.

**Measurement and Signature Intelligence (MASINT).** Technically derived intelligence that enables detection, location, tracking, identification and description of unique characteristics of fixed and dynamic target sources [377]. As stated, it includes TECHINT, intelligence gathered from the collection, processing, analysis and exploitation of data and information pertaining to foreign equipment and materiel [58].

**Signals Intelligence (SIGINT).** Intelligence produced by exploiting foreign communications systems and non communications emitters [449], which comprises three subcategories: communications intelligence (COMINT), electronic intelligence (ELINT), and foreign instrumentation signals intelligence (FISINT).

**Open Source Intelligence (OSINT).** Intelligence gathered from publicly available information that is collected, exploited, and disseminated in a timely manner to an appropriate audience for the purpose of addressing a specific intelligence requirement [666].

### Counterintelligence

Previously, I have referred to counterintelligence (CI) as an IO Supporting Capability, those that provide additional operational effects. [238] defines CI as Information gathered and activities conducted to identify, deceive, exploit, disrupt or protect against espionage, other intelligence activities, sabotage, or assassinations conducted for or on behalf of foreign powers, organizations or persons or their agents, or international terrorist organizations or activities; in [190] John Ehrman presents some competing definitions of the term. When dealing with CI, Spanish Law [239] refers to the actions to prevent, detect and neutralize the threat, being this threat in the form of foreign services, groups or persons.

What seems clear from the different approaches to CI is that it is defined by a set of actions (information gathering, tactical activities, etc.) conducted to counter (prevention, detection, deception, neutralization, etc.) hostile intelligence activities from an adversarial threat such as foreign services, organizations or persons. CI is directly linked to intelligence (in fact, some works [187] differentiate between positive intelligence and counterintelligence under the "intelligence" umbrella term). It is also linked to information assurance [573], although it is mandatory to distinguish between counterintelligence and security measures: security measures are defensive devices applied as protection against the things which counterintelligence seeks knowledge of, and they relate directly to the item to be secured, denying or inhibiting access to particular information, material or areas [633].

CI is a field much less structured than intelligence, maybe because of the fact that it has been considered the most complex and least understood of all intelligence disciplines [233]; although there have been some efforts to establish a CI process, it is not as structured as the intelligence cycle. In what authors agree is in the fact that CI can be offensive and defensive, and of course both approaches have to be coordinated and integrated to achieve global goals. Defensive CI includes all actions taken to identify and counter intelligence activities from an hostile threat; it is directly related to security measures [498] (in fact, there is some confusion between CI and security [497]), to knowledge gathered about threats, to risk assessment, etc. in summary, to the prevention and minimization of the threat. On the other hand, offensive CI is defined [640] by interactions with the adversary to directly collect information about their intelligence collection operations or to deceive them.

CI comprises four principles [497] [498]: to deter, detect, deceive and neutralise the opposition's efforts to collect information, regardless of why this data is collected; defensive CI is linked to deterrence and detection, while offensive CI is linked to detection[2], deception and neutralization.

### Words of Estimative Probability

When dealing with intelligence concepts, as well as with security ones, in most cases it is very difficult (almost impossible) to have a complete certainty about a particular fact. For this reason, analysts must deal with probabilities and confidence levels: as an example, it is mandatory to attribute an operation to a certain actor with a high/medium/low confidence level. This degree of confidence is defined by multiple parameters, such as source integrity, external facts analysis, etc., and in many cases analysts tend to use terms such as "very probable", "low probability" or "high confidence". These vague terms have different meanings, depending on the analyst that is using them, and they can be also understood in multiple ways, depending on the receiver of the intelligence as a product.

To address this situation, in 1964 Sherman Kent [326] stated the concept of Words of Estimative Probability (WEP) or Language of Uncertainty (LoU). The

[2] Please note that detection can be both defensive and offensive; the offensive side would be, for example, the hunting for intruders: the so called threat hunting.

goal of WEP is to set forth the community's findings in such a way as to make clear to the reader what is certain knowledge and what is reasoned judgment, and within this large realm of judgment what varying degrees of certitude lie behind each key judgment. Sherman Kent, one of the fathers of modern intelligence analysis, was trying to fill the gap between common expressions in natural language and their associated statistical probabilities in a manner that, no matter who reads an intelligence report, the interpretation of the words is the same. In table 1.4.3 the relationship between probability ranges and the right proposed words to express those probabilities in a non–numeric manner is presented, as in Kent's original work.

Table 1.1: Sherman Kent's Words of Estimative Probability.

| 100% **Certainty** | | | |
|---|---|---|---|
| The General Area of Possibility | 93% | give or take about 6% | **Almost certain** |
| | 75% | give or take about 12% | **Probable** |
| | 50% | give or take about 10% | **Chances about event** |
| | 30% | give or take about 10% | **Probably not** |
| | 7% | give or take about 5% | **Almost certainly not** |
| 0% **Impossibility** | | | |

If we simplify a lot, this approach can be seen only as the clarification of the meaning of some words. However, specifying this correct meaning and associating it with probabilistic terms (this is, with numerical terms) is a must if we want to deal with real intelligence; in fact, a vaguely defined terminology which leads to confusion has been identified as a main problem when referring to cyber threat intelligence [387].

In addition to WEP, when dealing with intelligence information it is mandatory to consider the intelligence source and information reliability rating systems used in intelligence analysis. Usually, those systems rate both the reliability of the source and the information itself. For example, [451] defines the methodology used by the United States Army to rate both of them. In its appendix B we can find the matrices to rate source reliability, ranging from *Reliable (A)* o *Unreliable (E)*, and information content, from its highest degree of confidence (1) to its lower one (5), including a level indicating that no determination can be done about the information. Both matrices are shown in tables 1.4.3 and 1.4.3.

### 1.4.4 Cyber intelligence

In this context of intelligence concepts, it is a must to refer to cyber intelligence, a vague term used in the cyber arena. Cyber Intelligence, CYINT or CYBINT, is somewhat related to cyberspace, a concept that as stated before has no single definition. While HUMINT is considered as intelligence from human sources, CYBINT concept can not be the equivalent, intelligence from cyberspace. The

Table 1.2: Evaluation of Source Reliability.

| | | |
|---|---|---|
| **A** | **Reliable** | **No doubt** of authenticity, trustworthiness, or competency; has a history of complete reliability |
| **B** | **Usually Reliable** | **Minor doubt** of authenticity, trustworthiness, or competency; has a history of valid information most of the time |
| **C** | **Fairly Reliable** | **Doubt** of authenticity, trustworthiness, or competency but has provided valid information in the past |
| **D** | **Not Usually Reliable** | **Significant doubt** about authenticity, trustworthiness, or competency; history of invalid information |
| **E** | **Unreliable** | **Lacking** in authenticity, trustworthiness, or competency; history of invalid information |
| **F** | **Cannot be Judged** | **No basis** exists for evaluating the reliability of the source |

term is generally used to convey the idea of widely scoped and better qualified knowledge of actual or potential events regarding cyberspace that may endanger an organization [77]. CYBINT can not be considered a collection discipline, but an analytic one: this is, with its focus on the analysis stage of the intelligence cycle, relying on data collected from the information gathering disciplines stated before [27] [557] (SIGINT, HUMINT, MASINT, OSINT and GEOINT).

In 2011 Intelligence and National Security Alliance (INSA) published [201] the first formal and high level approach to the "emerging discipline" of CYBINT, providing a framework to approach to the development of intelligence in the cyber domain and stating it as a new discipline in the Intelligence Community, but without providing an accurate definition of the term. The same year some authors stated the earliest definitions of cyber intelligence. In [481] Pythagoras Petratos defined it as the process of obtaining specific types of valuable information and knowledge trough the Internet. Pythagoras Petratos also defined [609] cyber intelligence, as the process of collecting, relating, analysing, and reporting information about a topic, an organization or a person, from sources available on the internet and other open sources. These initial definitions make a clear reference to intelligence gathered *from* Internet, and have been superseded during the decade with more accurate terms that fit better the concept that today we have about cyber intelligence.

With the early concept of intelligence *from* Internet, in 2012 Matthew M Hurley discussed [282] what cyber intelligence is, differentiating "from" and "for" cyber, depending on the scope of the information gathering activities, the means employed to carry them out and the final purpose they serve. Intelligence "from"

Table 1.3: Evaluation of Information Content.

| | | |
|---|---|---|
| **1** | **Confirmed** | **Confirmed** by other independent sources; **logical** in itself; **consistent** with other information on the subject |
| **2** | **Probably True** | **Not confirmed**; **logical** in itself; **consistent** with other information on the subject |
| **3** | **Possibly True** | **Not confirmed**; **reasonably logical** in itself; **consistent** with other information on the subject |
| **4** | **Doubtfully True** | **Not confirmed**; possible but **not logical**; **no other information** the subject |
| **5** | **Improbable** | **Not confirmed**; **not logical** in itself; **contradicted** by other information on the subject |

is knowledge produced through the analysis of any valuable information collected within or through cyberspace [77], while intelligence "for" refers to capabilities to enable cyberspace operations regardless of the source, method or medium: this is, different collection disciplines providing valuable intelligence to these operations.

In 2013 [54] stated that CYBINT should not be limited to an understanding of network operations and activities, but should include the collection and analysis of information that produces timely reporting, with context and relevance to a supported decision maker. Although yet undefined, what was clear is that the term refers to a multifaceted approach to framing, thinking about, and reacting to cyber adversarial activity, not only regarding intelligence *from* cyber space.

Although still in 2020 there is no consensus about a formal CYBINT definition (some discussion can be found at [316], [557] and [77]), a useful and simple approach was proposed in [618], which defines CYBINT as the acquisition and analysis of information to identify, track, and predict cyber capabilities, intentions, and activities that offer courses of action to enhance decision making. This definition fits well in what is usually understood as CYBINT, as the product derived from the analytic discipline, focusing in cyber intelligence *for* cyberspace but also including intelligence gathered *from* cyberspace, as long as it is useful for cyber activities. Really, when we refer to intelligence gathered from cyberspace to satisfy information needs outside this battlefield, we could simply refer to classical collection disciplines. For the purpose of this work, this definition will be used.

In addition to CYBINT, a term that is usually used among the information security community is Cyber Threat Intelligence, or CTI, first defined [397] as evidence–based knowledge, including context, mechanisms, indicators, implications and actionable advice, about an existing or emerging menace or hazard to assets that can be used to inform decisions regarding the subject's response to that menace or hazard. CTI focuses [142] on all source intelligence about threats: programs, intentions, capabilities, research and development, tactics, targets, operational activities and indicators, potential impacts, infrastructure and data, characterization, structures, etc. The term is also used without the cyber prefix –this is, Threat Intelligence or TI–, and its goal is [148] to help organizations in recognizing the indicators of cyber attacks, extracting information about the attack methods, and consequently responding to the attack accurately and in a timely manner.

CTI can be considered as a subset of CYBINT: CYBINT includes CTI, but CTI does not represent all the elements of CYBINT [195]. While CTI focuses on the single analysis of threats, cyber intelligence includes this analysis, but also the analysis of areas such as geopolitics, military or diplomacy; CTI, from its definition to its goal or its components, focuses on threats, not in their external context.

Both in CYBINT and in CTI it is possible to identify different levels to deal with; in fact, it is possible to identify these levels in all intelligence–related activities. Each of these levels refers to intelligence (CYBINT o CTI) with a specific goal, time of life, type of product, etc., and they are defined as follows [54] [308] [9][3]:

**Strategic.** Level at which an actor determines global strategic security objectives and guidance, and develops and uses resources to achieve these objectives. In the cyber domain, strategic intelligence provides knowledge to understand threats and risks at a senior management level: main actors and their motivations, victims and their relations, links to geopolitical events, etc. The final product is usually in the form of written reports with a long lifetime and a non–technical approach, about *who* and *why*.

**Operational.** Level at which campaigns and major operations are planned, conducted, and sustained to achieve strategic objectives within theaters or other operational areas. In the cyber domain, operational intelligence provides knowledge about the context and trends of past incidents [400]: tactics, techniques, patterns, actor profiles, etc. The final product is in the form of short written reports with a medium lifetime, about *how* and *where*.

**Tactical.** Level at which battles and engagements are planned and executed to achieve military objectives assigned to tactical units or task forces. This is the most basic form of intelligence, and in the cyber domain it provides knowledge about the identification of threats targeting the infrastructure in the form of hashes, IP addresses, domains or detection rules. The final product is in the form of low–level indicators in a machine–readable format, such as Yara rules, IDS signatures, OpenIOC indicators or blacklists. All of them are suitable to load in different security devices; tactical intelligence has a short lifetime and tries to answer *what* is happening or what is to happen.

Finally, a term not widely used among professionals but commonly accepted in mass media is **cyber espionage**, a concept also related to CYBINT that, as its name implies, refers to espionage activities performed through the cyberspace: the

[3] Other works [548] [426] [360] change the definitions and layers of tactical and operational levels of intelligence, while others [442] include a fourth level, called technical, at the lowest part of the heap.

covert gathering of information through computers and networks, a very similar concept to what in military doctrine is called Computer Network Exploitation (CNE).

### Cyber counterintelligence

The same way today we refer to cyber intelligence, in our “cyber” approach many works refer to cyber counterintelligence (CCI), being this a concept a that has been developed in conference papers, master thesis and post graduate studies especially since 2010 [184]. A literature companion can be found on [183].

CCI is defined [184] as a subset of multidisciplinary CI aimed at deterring, preventing, degrading, exploiting, and neutralising adversarial attempts to collect, alter or in any other way to breach the confidentiality, integrity or availability of valued information assets through cyber means. Although we could discuss the goals, for example, by including identification or penetration of adversarial threats, this definition focus on the obvious key concept for CCI: the cyber means. CCI is a part of CI as a multidisciplinary concept or approach, in this case focusing on the “cyber” field.

As in classical CI, CCI can be both defensive and offensive, with the same principles than in CI: while defensive CCI is linked to deterrence and detection, offensive CCI is linked to detection, deception and neutralization. For example, Threat Intelligence is a defensive CCI type effort [185], while Threat Hunting is an offensive one. In [187] Petrus Duvenage et al. propose a CCI approach that combines both offensive and defensive views into a single process. A maturity model for CCI is presented in [302] and [303].

## 1.4.5 Threat modeling

A **threat** is defined [603] as any circumstance or event with the potential to adversely impact organizational operations (including mission, functions, image, or reputation), organizational assets, individuals, other organizations, or the Nation through an information system via unauthorized access, destruction, disclosure, or modification of information, and/or denial of service; in this reference the authors identify four types of threat sources:

**Adversarial.** Individuals, groups, organizations, or states that seek to exploit the organization dependence on cyber resources.

**Accidental**. Erroneous actions taken by individuals in the course of executing their everyday responsibilities.

**Structural.** Failures of equipment, environmental controls, or software due to aging, resource depletion, or other circumstances which exceed expected operating parameters.

**Environmental.** Natural disasters and failures of critical infrastructures on which the organization depends, but which are outside the control of the organization.

As they represent the only deliberate ones[4], for the purpose of this work I will focus on adversarial threats, this is, on threat actors; more specifically, I will focus on advanced threat actors, those defined in the previous section, which are able to perform a full range of offensive cyber operations, from CNE to PSYOP, and which seriously impact in the target information, infrastructure or reputation.

A model is defined [76] as an abstract representation of some domain of human experience, used to structure knowledge, to provide a common language for discussing that knowledge, and to perform analyses in that domain. Models are mandatory in order to better understand and discuss abstract entities, representing and structuring common knowledge and experiences, and allowing analysts to profile attackers, from their goals to their TTP. They are used in many fields of investigation, from health to telecommunications or human behaviour.

Once threat and model have been defined, we can consider threat modeling in its simplest definition [567] as using models to find security problems. It is difficult to establish a single global definition for threat modeling, as it refers to threat agents, to risks or to techniques regarding requirements specification or design analysis [566]. In any case, it is considered a process whose goal is to mitigate security risks. [458] defines threat modeling as the formal process of identifying, documenting and mitigating security threats to a software system.

Threat modeling has multiple benefits for analysts and defenders. The main one is that, by using abstract models, it allows those analysts to be independent of specific technologies or systems, thus facilitating a global definition of security requirements and the implementation of defense mechanisms [429]. In fact, threat modeling is about risk evaluation, a task that facilitates the prioritization of security initiatives and its associated budget in each case.

Three well defined structured approaches for threat modeling can be identified [567] [76]:

**System–Centric.** Models the system, data and boundaries in the environment and then determines what threats are relevant. This is, by far, the most used approach while doing threat modeling.

**Asset–Centric.** Identifies the organizational assets that could be affected by threats, characters the threats that could affect those assets and situates the assets in terms of systems. While asset centric modeling seems a logical approach, it presents a complex problem: the identification of the relationships between assets and threats [567], so asset–centric is not a common (neither recommended) approach.

[4] Please note that structural threats, while referring to depletion, are not considered inside deliberate actions but inside infrastructure failures.

**Threat–Centric.** Models the threat, generally or specifically, and then applies the model to a relevant environment. This approach is not considered effective because of the gap between attackers and what they really want to do, thus not leading to reproducible results.

Many threat modeling methodologies have been developed over time; from STRIDE (Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privilege), defined in 1999 and adopted by Microsoft some years later, to PASTA (Process for Attack Simulation and Threat Analysis) or ATASM (Architecture, Threats, Attack Surfaces, and Mitigations). Most of them have been primary used to find security flaws in the development of new products or systems [615] [673] [5] [538]: this is, a system–centric approach, the most common one used to threat modeling. In fact, [567], a key reference in the topic, recommends not to use attackers or assets as the center of the threat modeling process. A summary and analysis of some of threat modeling methodologies can be found in [399], [565] and of course in [567].

This work will delve into the modeling of advanced threat actors (this is, adversarial sources –specifically, as stated before, the advanced ones–), following a threat–centric approach (this is, focusing on the threat). The goal is not to protect specific infrastructures or organizations, but to better know this kind of attackers, which will help us to enable a better protection against them.

### 1.4.6 Threat actors' categorization

Threats can be categorized [591] [370] [592] by three elements with foundational dependencies between them, defining an ontology, in what is defined as the COI model:

**Capability.** An actor's means to perform an act.

**Opportunity.** Relationships between the actor and the situation elements to be acted upon.

**Intent.** The actor's intended actions and its hoped for effects.

This threat ontology considers these three elements to be the principal factors in predicting intentional actions [591]; these elements are, from a counterintelligence point of view, what analysts consider when developing counter measures [249]. In the context of this work, capability and intent are related to the threat actor, while opportunity is related to its target, so I will focus on the first two elements for the modeling of threat actors.

These threat actors can be defined as persons or entities that are responsible for incidents that impact, or have the potential to impact, into the safety or security of another person or entity. They can range from script kiddies to the most sophisticated threat groups, including terrorists or governments; many works have established typologies for threat actors (a summary can be found in [168]). This work will focus on the last ones, on advanced threat actors, such as intelligence services, military units or organized criminal groups.

In [539] Mirko Sailio et al. present the threat actors that different expert organizations identified, as well as their classification. However, although the authors identify common definitions for threat actors, they do not provide a global approach for their classification. In [389] Vasileios Mavroeidis et al. define an ontology to infer the types of threat actors, in which the authors identify attributes and types for these actors. Attributes include motivation, limits or resources, among others. Such an ontology is a useful tool for the categorization of threat actors; however, the clear identification of all of these attributes is not only a difficult work, but in many cases it is simply not possible. In [384] Louis Marinos identifies different types of threat actors with high capabilities, advanced threat actors, those who perform activities such as cyber espionage, hacktivism, terrorism, crime and war. However, not all authors agree on the activities performed by those advanced threat actors. Most western intelligence services have released reports linking advanced threat actors to nation–state actors, this is, to threat actors sponsored by, or directly part of, a government. In the case of the Spanish Centro Nacional de Inteligencia, through its CERT, its yearly report "Cyber Threats and Trends" [431] has been stating since 2008 that advanced threat actors perform cyber espionage and cyber attack campaigns (inside Computer Network Operations), as well as Psychological Operations. If we take a look at the core capabilities inside Information Operations, hostile actions are related to EW (Electronic Attack), CNO (Computer Network Attack and Computer Network Exploitation), MILDEC and PSYOP; this is, all capabilities –in its offensive role– but OPSEC, which is a purely defensive discipline. As stated before, in [450] the US Military refer to Cyberspace Operations, defined [625] as the employment of cyberspace capabilities where the primary purpose is to achieve objectives in or through cyberspace. As stated before, this is a broader term that comprises not only CNO, directly linked to the most technical aspects of operations, but all offensive capabilities when used in cyberspace.

When dealing with advanced threat actors, it is mandatory to refer to the definition of Advanced Persistent Threats, APT; this term is apparently originated from US Air Force in 2006, including Col. Gregory Rattray as the individual who coined this term [333], probably as a polite way to refer to Chinese cyber attacks exfiltrating gigabytes of information from US targets.

Some years after the term was coined, Richard Bejtlich [62] gave one of the earliest formal definitions of an APT; analyzing each of the letters from the acronym, Bejtlich stated that APT can be defined as follows:

**Advanced** means the adversary can operate in the full spectrum of computer intrusion. They can use the most pedestrian publicly available exploit against a well–known vulnerability, or they can elevate their game to research new vulnerabilities and develop custom exploits, depending on the target posture.

**Persistent** means the adversary is formally tasked to accomplish a mission. They are not opportunistic intruders. As an intelligence unit, they receive directives and work to satisfy their masters. Persistent does not necessarily mean they need to constantly execute malicious code on victim computers. Rather, they maintain the level of interaction needed to execute their objectives.

**Threat** means the adversary is not a piece of mindless code. This point is crucial. Some people throw around the term “threat” with reference to malware. If malware had no human attached to it (someone to control the victim, read the stolen data, etc.), then most malware would be of little worry (as long as it did not degrade or deny data). Rather, the adversary here is a threat because it is organized and funded and motivated. Some people speak of multiple “groups” consisting of dedicated “crews” with various missions.

As we can realize from this definition, APT refers to groups of people formally accomplished with missions, with appropriate budget –this is “advanced”– and capabilities to operate against a full range of targets. From this definition we can understand APT as a form to refer to an advanced threat actor, in most cases an intelligence service or military unit or department, performing CNA or CNE operations. This last point is a key one, as APT are usually considered with a focus on stealth computer and network intrusions, both to gather information (CNE) or to destroy or manipulate their target (CNA).

Some months after Bejtlich’s definition, in its Special Publication 800–39 [530], NIST defined an Advanced Persistent Threat as an adversary that possesses sophisticated levels of expertise and significant resources which allow it to create opportunities to achieve its objectives by using multiple attack vectors. These objectives typically include establishing and extending footholds within the information technology infrastructure of the targeted organizations for purposes of exfiltrating information, undermining or impeding critical aspects of a mission, program, or organization; or positioning itself to carry out these objectives in the future. The advanced persistent threat: (i) pursues its objectives repeatedly over an extended period of time; (ii) adapts to defenders’ efforts to resist it; and (iii) is determined to maintain the level of interaction needed to execute its objectives.

NIST definition explicitly refers to advanced hostile actors establishing a foothold within its target infrastructure, focusing on the exfiltration of information -present or future-: this is, in CNE. Dissecting these two definitions, we can extract some interesting conclusions about what is and what is not an APT. In addition to these –and much more– definitions, many analysts and companies have tried to re–define the concept “APT”; in fact, during last years the term APT has become a buzzword, with many confusing explanations about what they are or what they are not, mixing concepts and finally presenting APT as advanced malware, campaigns or specific actions.

As stated before, APT definition and groups have been mainly associated to CNE, to cyber espionage, although many of the best known groups have performed destructive actions (CNA), especially those linked to the military, such as APT28 or Sandworm. In fact, nowadays we know that APT groups also perform activities such as PSYOP or EW. Definitely, all of the offensive IO core capabilities. The first part of the definition from [530] really does cover this concept, although it focuses on information exfiltration: an adversary that possesses sophisticated levels of expertise and significant resources, that allow it to create opportunities to achieve its objectives by using multiple attack vectors. In summary, APT not only performs CNA/CNE, but any kind of offensive Information Operations through cyberspace. The issue here is the own name: if we associate the term APT to CNE or CNA, another concept must be used to identify those actors operating on the full range of offensive IO, so the term Advanced Threat Actors will be adopted, no matter if I am referring to an APT or to another kind of offensive actor.

An advanced threat actor will be tasked to accomplish a mission; this mission can be related to goals such as intelligence gathering, damage or control of an infrastructure or influence. This mission is tasked by a superior authority that can be defined as the **sponsor**: that is, the individual, group, organization, state, etc. that gives an order. Please note that superiority can be stated by elements such as a military range, a more powerful economic position or a legal status, or simply by threatening, blackmailing or kidnapping.

### 1.4.7 Threat actors' characterization

In this section relevant concepts for the characterization of threat actors are presented. By characterization, I refer to the extraction and analysis of threat actors' features in order to identify their interests, goals or operations, among others. In this section, threat actors' tactics, techniques and procedures will be defined, as they are the key element for their characterization and for their later detection. In addition, the MITRE ATT&CK framework will be presented, as the *de facto* standard nowadays for the identification of advanced threat actors' tactics and techniques. This section will be finished by introducing Threat Agent Risk Assessment, a methodology to identify threat agents and their operations.

#### Tactics, techniques and procedures

Every threat actor, those that perform Information Operations and those that perform other kind of activities, develops Tactics, Techniques and Procedures (TTP) to achieve its goals; in [308] we can find the following definitions for these terms:

**Tactics** The employment and ordered arrangement of forces in relation to each other.

**Techniques** Non–prescriptive ways or methods used to perform missions, functions, or tasks.

**Procedures** Standard, detailed steps that prescribe how to perform specific tasks.

Tactics specify what a threat actor is doing, at the highest level of description, to accomplish a certain mission, while techniques specify how a tactic is implemented and procedures describe a particular implementation of a technique. Those Tactics, Techniques and Procedures represent the behavior of the actor, very similar to what we usually call its *modus operandi*, from the highest–level description (tactic) to the lowest–level one (procedure) [304].

Although TTP play a central role in cyber threat intelligence [56] and it is a widely used term, the different concepts they represent are not well expressed among different works: many books, scientific papers or industrial reports use the concept "TTP" without proper definitions of each term, or even mixing concepts between tactics, techniques and procedures. The previous definitions, from the military, are not clear between information security professionals [391].

Tactics represent what a threat actor is doing, as stated before; this term is a high level one that does not get into how it is accomplished. It is possible to refer to MITRE ATT&CK to identify tactics such as *Lateral movement* or *Exfiltration*: they represent what a threat actor wants, regardless of the particular techniques or procedures performed in each case. For example, to achieve *Exfiltration* –the tactic–, a threat actor can apply a variety of techniques, ranging from the execution through module load to the use of Powershell.

As stated in its definition, techniques are non–prescriptive, meaning there is no procedural sequencing with techniques [595]; they apply to an individual threat actor or even persons, which is a key point for intelligence analysts handling with multiple incidents (for example, for attribution hypotheses) and his/her own way to operate

Finally, as its own definition states, procedures specify how to perform specific tasks in the form of sequences of actions: this is, they are a particular implementation of a technique. These sequences are well–defined and standardized, so no matter who is executing a procedure, it will be always the same (it is their main differences with techniques, which depend on individual threat actors). Obviously, procedures are not so largely discussed as techniques, because they are supposed to be internal to groups –specifically, in our case, to threat actors–. Looking at the previous example regarding tactics and techniques, a procedure would specify how the execution through module loading could be performed, in a manner than an operator could execute it without thinking, just as a cooking recipe.

### MITRE ATT&CK

MITRE ATT&CK (Adversarial Tactics, Techniques, and Common Knowledge) is a globally accessible knowledge base of adversary tactics and techniques based on real–world observations. This knowledge, contributed by analysts all around the world, can be used as a base for the development of specific threat models and methodologies. Started in 2013 and published in 2015, ATT&CK develops a process for modeling an adversary's post–compromise behavior at a granular level, allowing the intelligence sharing among the community [597]; an relevant description of the framework and the work done can be found at [598].

As of November, 2022, ATT&CK has defined fourteen enterprise tactics (this is, those regarding the activities of an attacker into its victim) and 193 enterprise techniques associated with those tactics, with 401 sub techniques. In addition to that, ATT&CK defines fourteen mobile tactics, related to the compromise of mobile devices, and 66 mobile techniques with 41 sub techniques[5].

The enterprise tactics defined by ATT&CK represent those executed by a threat actor during an operation; although they are described in a particular, logical order, from the initial access to the final actions in the targeted infrastructure, they must be seen as linked to the global stages followed by an advanced threat actor, so they can be executed in parallel or simply not executed if the threat actor does not need them. The fourteen enterprise tactics that MITRE ATT&CK defines are the following ones:

1. **Reconnaissance.** The adversary is trying to gather information they can use to plan future operations.
2. **Resource development.** The adversary is trying to establish resources they can use to support operations.
3. **Initial access.** The adversary is trying to get into your network.
4. **Execution.** The adversary is trying to run malicious code.
5. **Persistence.** The adversary is trying to maintain their foothold.
6. **Privilege escalation.** The adversary is trying to gain higher–level permissions.
7. **Defense evasion.** The adversary is trying to avoid being detected.
8. **Credential access.** The adversary is trying to steal account names and passwords.
9. **Discovery.** The adversary is trying to figure out your environment.
10. **Lateral movement.** The adversary is trying to move through your environment.
11. **Collection.** The adversary is trying to gather data of interest to their goal.
12. **Command and Control.** The adversary is trying to communicate with compromised systems to control them.
13. **Exfiltration.** The adversary is trying to steal data.
14. **Impact.** The adversary is trying to manipulate, interrupt, or destroy your systems and data.

Each of the tactics MITRE ATT&CK has defined represents the highest level of abstraction within the model, specifying what the threat actor tries to perform during an operation; under them, the next level of abstraction is defined by

[5] Please note that MITRE ATT&CK is an ongoing work, so at the moment of reading the number of tactics and techniques, or even the structure of the framework, may have changed.

techniques, linked to at least one tactic (they can be linked to more than one), that specify how a tactic is performed. Beside tactics and techniques, ATT&CK identifies software, a generic term for elements such as tools, artifacts or malware, that can be used to implement one or more of the techniques stated before.

MITRE also links APT groups to tactics, techniques and software; with 135 identified groups at the time of this writing, everyone of them is named, aliased, described and linked to specific techniques (including pre–attack and mobile) and software. In this way, an analyst can establish relations between those entities to model an adversary and its activities against a target and, most important, to establish defense mechanisms to prevent, detect and respond to a threat.

ATT&CK represents an enormous effort to provide the community with an unified framework to identify the activities of advanced threat actors, from their TTP to the software they use, correlate information among those entities and improve not only the knowledge about APT, but also the defense mechanisms required to counter them. However, as it will be detailed in this work, the framework must be improved over time with contributions from analysts.

### Threat Agent Risk Assessment

Threat Agent Risk Assessment (TARA) is a methodology developed by Intel IT [6] [529] to identify threat agents that are pursuing objectives which are reasonably attainable an could cause unsatisfactory losses. TARA does not attempt to identify every single weak point of a system, product or organization. However, it identifies which threat agents pose the greatest risk, their motivation and the likely methods they will employ [499]

TARA relies on three groups (libraries) of collected data:

**Threat Agent Library (TAL).** Relevant threat agents and their corresponding attributes. This library was first introduced in [113].

**Methods and Objectives Library (MOL).** Lists known threat agent objectives and the most likely methods they will employ to reach them.

**Common Exposure Library (CEL).** Enumerates known information security vulnerabilities and exposures, mapping vulnerabilities against existing controls to show which exposures are residual.

From the above list, for the purpose of this work it is interesting to focus on TAL, the Threat Agent Library, and MOL, the Methods and Objectives Library. The first one, TAL, develops a common set of attributes for threat agents, both for hostile and for non hostile ones; this common set is defined as follows [113]:

**Access.** The extent of the threat agent's access to the target's assets: internal or external.

[6] MITRE has also developed a methodology called TARA. The only common thing with Intel's one is its acronym.

**Outcome.** The agent's primary goal: theft/acquisition, business advantage, damage, embarrassment or technical advantage.

**Limits.** Legal and ethical limits that may constraint the agent: code of conduct, legal or extra–legal.

**Resources.** The organizational level at which a threat agent typically works: individual, club, contest, team, organization or government

**Skills.** The special training or expertise a threat agent typically possesses: none, minimal, operational or adept.

**Objective.** The action that the threat agent intends to take in order to achieve a desired outcome: copy, destroy, injure, take or do not care.

**Visibility.** The extent to which the threat agent intends to conceal or reveal its identity: overt, covert, clandestine or do not care.

In [114] Timothy Casey introduced the *Motivation* attribute to the previous list as a combination of different types of motivations, including the *Defining Motivation*, the single most prevalent and descriptive motivation of the agent archetype, and *Personal Motivation*, that of the individual threat agent.

On the other side, Methods and Objectives Library, MOL, provides information about threat agent objectives and the most likely methods they might use to accomplish these objectives [529]. Objectives represent what the threat agent wants to accomplish, while methods represent the likely vias through which an attack might occur. In [529] Matthew Rosenquist provides examples both for objectives (theft/exposure, data loss, sabotage, etc.) and for methods (copy/expose, destroy/delete/render unavailable, etc.).

TARA couples the MOL with the TAL to emerge a picture of the types of likely possible attacks based upon many factors derived from both libraries; when they are overlaid with the CEL, not described in this work, the areas of high exposure are identified, as they are those with sufficient controls deployed to reduce risk.

### 1.4.8 Threat actors' detection

In this section relevant concepts for the detection of threat actors are presented. Please note that when I refer to detection of threat actors, I am actually referring to the detection of their operations. For this reason, in this section I will focus on models for the description of offensive Computer Network Operations. The Cyber Kill Chain® (CKC) framework is introduced as a relevant model for the identification of hostile operations. Another model in use is The Diamond Model for Intrusion Analysis, also detailed in this section, which allows analysts to track the operations of adversary groups over time. Finally, a model to assess the capabilities of an organization at detecting hostile activities is the Detection Maturity Level (DML), which can be also seen as a clear way to characterize threat actors.

### Cyber Kill Chain®

The Cyber Kill Chain® framework, developed by Lockheed Martin, is part of the Intelligence Driven Defense® model for identification and prevention of cyber intrusions activity, identifying what a threat actor must complete in order to achieve their objective. It was first described in [284] as a seven steps process suitable for CNA or CNE operations, as shown in figure 1.2.

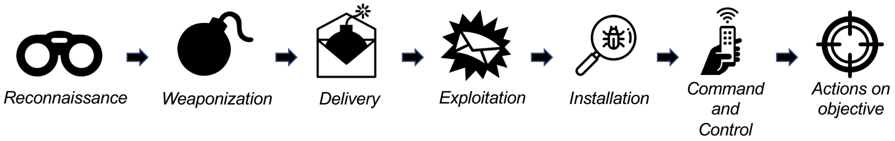

Figure 1.2: Cyber Kill Chain

These seven steps are defined as follows [284] [219]:

1. **Reconnaissance.** Research, identification and selection of targets.
2. **Weaponization.** Before really attacking a target, the threat actor has to couple a remote access trojan with an exploit into a deliverable payload.
3. **Delivery.** Transmission of the weapon to the targeted environment to launch a particular operation.
4. **Exploitation.** After the weapon is delivered, exploitation triggers intruders code.
5. **Installation.** Installation of an implant, such as a remote access trojan or a backdoor, on the victim system allows the adversary to maintain persistence inside the environment.
6. **Command and Control.** Compromised hosts must beacon outbound to an Internet controller server to establish a C2 channel, thus allowing the threat actor to control its target remotely.
7. **Actions on Objectives.** After progressing through the first six phases, intruders can take actions to achieve their original objectives, such as information theft, denial or hop to a third party infrastructure.

The Cyber Kill Chain represents an industry–accepted methodology for understanding how an attacker will conduct the activities necessary to cause harm to the organizations and has been largely discussed in different works [430]. In [684] Wen Zeng and Vasileios Germanos identify some of the discussions regarding the application of the Cyber Kill Chain. Some authors [350] [92] have proposed the addition and removal of different stages in order to improve or adjust the original model, and the topic has been also discussed in technical conferences. In addition, some efforts have been made in order to unify models and variants, such as [489].

In spite of that, the original proposal has been widely used and applied to specific problem regarding advanced threat actors, for example those regarding the modeling of critical structures attack stages [253] [112] [686].

Critics to the CKC are related to its approach as a linear progression not accurately representing the actions of an actor; Mandiant (FireEye) presented in [398] the Mandiant Attack Life Cycle, a model in which the weaponization stage is removed and that introduces a loop to represent the continuous activities of internal recon, lateral movement and persistence performed by an hostile actor, as shown in figure 1.3.

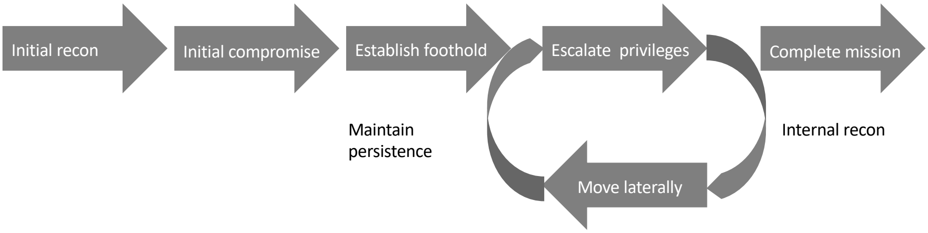

Figure 1.3: Mandiant's Attack Life cycle Model

Mandiant Attack Life cycle is focused on the detection of the activities of an hostile actor against an infrastructure; because of that, the weaponization stage is removed from the model, as stated before. This approach, although useful for defensive teams (it focuses on what can be detected) is not accurate for a global modeling of threat actors.

### The Diamond Model of Intrusion Analysis

The Diamond Model of Intrusion Analysis (DMIA) [105] applies scientific principles to intrusion analysis, providing a simple, formal, and comprehensive method of activity documentation, synthesis, and correlation. The model represents an *adversary* deploying a *capability* over some *infrastructure* to target a *victim*; these activities, called events, are the atomic features of the model, and they define one step in a series that a threat actor must execute to achieve its goal. Each event has these core features (adversary, capability, infrastructure and victim) as well as meta–features (timestamp, phase result, direction, methodology and resources), both of them with an associated confidence value; the last ones are used to arrange events within an activity thread. This model, summarized in figure 1.4, is used to track adversary groups over time rather than the progress of individual attacks [616].

In [105] Sergio Caltagirone et al. improve the original approach defining the extended Diamond Model, in which two additional fundamental meta–features are defined: the Social-Political, determining the Adversary-Victim relationship (a

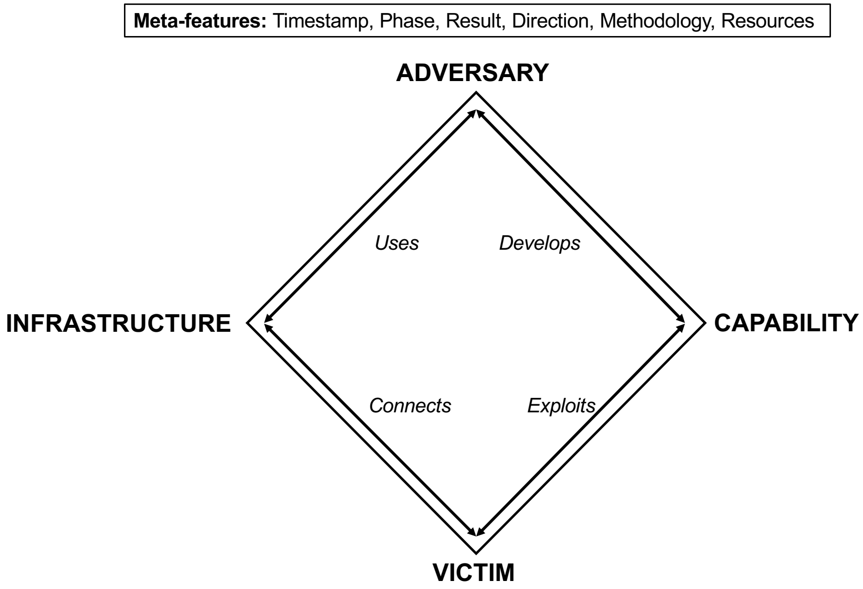
Meta-features: Timestamp, Phase, Result, Direction, Methodology, Resources
ADVERSARY
INFRASTRUCTURE
CAPABILITY
VICTIM
Uses
Develops
Connects
Exploits

Figure 1.4: The Diamond Model of Intrusion Analysis

relationship that always exists), and the Technology one, connecting and enabling the infrastructure and the capability to operate and communicate.

### Detection Maturity Level

In 2014, Ryan Stillions wrote a blog post [594] in which he presented the Detection Maturity Level (DML), a model to assess the maturity of an organization to detect cyber attacks in terms of its capabilities to consume and act upon given threat information. The DML model is composed of nine maturity levels, from the most technical ones (really, level 0 represents no information about the threat) to the highest level ones, such as goals and strategy of the threat actor, as shown in figure 1.5.

This layered model not only represents the detection maturity of a target organization, but it is also an approach towards how an advanced threat actor works. Its nine levels, from down to top, are defined as follows:

**DML-8. Goals.** In DML, the highest level of detection maturity for an organization, which is almost impossible to achieve. Like strategy, goals can not be detected by technical means, representing together what the threat actor wants.

**DML-7. Strategy.** Strategy marks the high level approach to what is to be done,

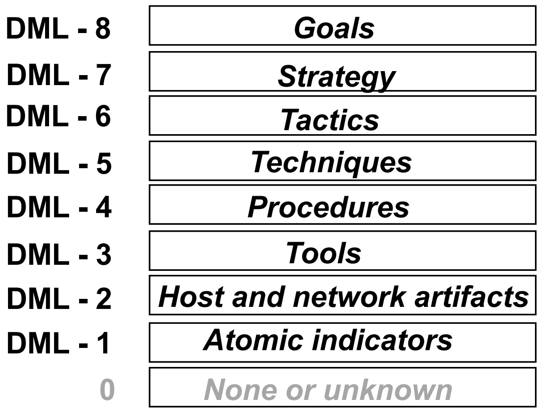

Figure 1.5: DML levels

and cannot be identified by technical means. It is a subjective concept that can not be easily structured, so it is not detected by pure technical means and has to be determined by intelligence capabilities.

**DML-6. Tactics.** Tactics represent what a threat actor is doing regardless which particular techniques or procedures are used. Their detection is a complex task based on the observation and aggregation of multiple activities, including not only the technical ones, so this detection is performed by analysts rather than by correlation or automated systems.

**DML-5. Techniques.** Techniques specify how a threat actor is operating and they are linked to a particular actor, so detection at this level implies that the organization has capabilities to identify individual actors that target it.

**DML-4. Procedures.** The organization is able to detect procedures, this is, sequences of steps that a threat actor follows during an operation. Procedures are specific, detailed, implementations of techniques, and represent the flow of an operation, so being able to detect them is a key capability to respond to an incident in its earliest stages.

**DML-3. Tools.** Detection at this level means capability to identify tools that the threat actor is using, regardless of minor functionality changes to them or the artifacts or atomic indicators they may leave behind.

**DML-2. Artifacts.** Host and network artifacts observed during or after an operation by the threat actor are available as traces for the analysts.

**DML-1. Atomic Indicators.** Only atomic indicators such as IP addresses, filenames or domains used by the threat actor are available, usually in the form of feeds from third parties.

**DML-0. None.** No information about the threat actor is available.

Once an actor is tasked with its mission, this is, it has a defined goal, it will work in its strategy: how it will accomplish its mission, how it will get its goal. From this point, the threat actor develops its tactics, what is to be done, its techniques, how it is to be done, and its procedures, the specific steps to implement the selected techniques for the operation.

An operation is executed through more or less complex tools, ranging from complete exploitation frameworks to simple artifacts; the execution leaves in the targeted infrastructure particular indicators, which in many cases are the initial point to the detection of an operation. This heap, seen from the perspective of the detection and knowledge of the threat, defines the actor's structure in DML.

The DML model has been improved [88] [387] by adding a tenth level of abstraction, DML-9, regarding identity of the threat actor, a useful information to connect multiple attacks to the same actor in order to predict strategy, tactics, techniques and procedures expected to be used in an operation. This hierarchical model can give an approach not only to evaluate the detection capabilities of an organization, but also to the semantic modeling of an advanced threat actor: from a group to specific indicators left after an operation, thus helping the analysts to model the interests and behaviour of the actor and its modus operandi in specific operations.

### STIX/TAXII

In this section STIX/TAXII is presented as the main reference developed to specify and share cyber threat intelligence. Structured Threat Information Expression (STIX) is a language and serialization format used to exchange cyber threat intelligence (CTI) whose goal is to identify and to represent all the elements of cyber threats in a flexible, automatable and human readable way. It is the main reference developed to specify and share cyber threat intelligence. Upon a standardized language, in XML format, STIX provides a common mechanism for addressing structured cyber threat information improving consistency, efficiency, interoperability, and overall situational awareness into a unified architecture that structures and links all those elements of a threat, from its lower level, where observables or indicators are, to the higher one (campaigns and actors) [56]. STIX and TAXII were defined in 2012, sponsored by US Department of Homeland Security, and in 2015 all the intellectual property and trademarks associated with STIX were licensed to OASIS, a nonprofit organization focused on the development and integration of open technological standards. At the time of this writing, the project repository can be found at `https://oasis-open.github.io/`.

STIX defines two kinds of objects to represent CTI, being the first ones the STIX Domain Objects (SDO); an SDO is an object that describes unique concepts represented in CTI, in one way or another: the actor performing a campaign, the campaign itself, the software used and vulnerabilities exploited or the observables in a compromised system, for example. In its version 2, the current one at the time of this writing, STIX defines twelve SDO:[7]: Attack Pattern, Campaign, Course of Action, Identity, Indicator, Intrusion Set, Malware, Observed Data, Report, Threat Actor, Tool and Vulnerability.

The second type of objects STIX defines are STIX Relationship Objects (SRO); as their name implies, they represent types of relationships established between SDO, used to describe CTI. We can find two SRO: Relationship and Sighting.

As stated, the STIX standard defines the Threat Actor SDO as individuals, groups, or organizations believed to be operating with malicious intent. In addition to the common properties and the obvious specific ones such as type, labels or name, this SDO defines properties for a threat actor, such as Roles, Goal, Sophistication or Motivation, among others.

STIX provides a complete and comprehensive model for threat actors' characterization regardless their type: it can be used to model from an unskilled individual to an advanced hostile actor. Beside STIX, Trusted Automated Exchange of Intelligence Information (TAXII) is an application layer protocol for the communication of cyber threat information in a simple and scalable manner, specifically designed to support the exchange of CTI (over HTTPS) represented in STIX[8]. Although TAXII was specifically designed to support the exchange of CTI represented in STIX format, and support for exchanging STIX 2.0 content is mandatory to implement, it can also be used to share data in other formats. STIX and TAXII are independent standards: the structures and serializations of STIX do not rely on any specific transport mechanism, and TAXII can be used to transport non-STIX data.

STIX and TAXII can be considered a way to store and exchange indicators, but neither of them provide methods to integrate or exploit data; STIX is a presentation language and TAXII is a transport protocol, so for some real world cases, just as IOC sharing, management and exploiting, there are many solutions easier to deploy that do not consider these standards.

[7] A detailed description of all of them can be found at OASIS website.

[8] For more information, please visit `https://oasis-open.github.io/cti-documentation/taxii/intro`

# Chapter 2

# A Taxonomy for Threat Actors' Delivery Techniques

*This chapter analyzes the Delivery tactic, a particularly relevant one as it is the first moment in a cyberspace operation in which a threat actor interacts with its victim. A taxonomy for Delivery associated techniques is provided, dissecting the key elements for this tactic in hostile cyberspace operations. By improving the modeling of advanced threat actors and their operations, this work increases the detection capabilities of all kind of organizations. It is aligned with MITRE ATT&CK, the main framework for the identification of threat actors' tactics and techniques.*

This chapter was published as: Villalón-Huerta, A.; Ripoll-Ripoll, I.; Marco-Gisbert, H. A Taxonomy for Threat Actors' Delivery Techniques. Applied Sciences. 2022, 12, 3929. https://doi.org/10.3390/app12083929

## Contents

The main contribution of this paper is to provide an accurate taxonomy for delivery techniques, which allows the detection of novel techniques and the identification of appropriate countermeasures. Delivery is a key stage for offensive cyber operations. During delivery, a threat actor tries to gain an initial foothold into the targeted infrastructure. It is the first step of an offensive cyber operation, where the threat actor interacts with its victim in a hostile way; thus, its success is mandatory for the global achievement of the operation. However, delivery techniques are not well structured among the literature, being in many cases a simple list of techniques with which, if one of them is slightly modified by the threat actor, its detection becomes very difficult. This situation hinders the modeling of hostile actors, a fact that makes it difficult to identify countermeasures to detect and neutralize their malicious activities. In this work, we analyze the current delivery techniques' classification approaches and the problems linked to them. From this analysis, we propose a novel taxonomy that allows the accurate classification of techniques, overcoming the identified problems and allowing both the discovery of new techniques and the detection of gaps in deployed countermeasures. Our proposal significantly reduces the amount of effort needed to identify, analyze, and neutralize hostile activities from advanced threat actors, in particular their initial access stage. It follows a logical structure that can be easy to expand and adapt, and it can be directly used in the industry's commonly accepted standards, such as MITRE ATT&CK.

## 2.1 Introduction

Computer Network Operations (CNO) are defined as the actions taken through the use of computers and networks to gain information superiority or to deny the adversary this enabling capability. CNO is an umbrella term that comprises three main activities [415]: Computer Network Attack (CNA), Computer Network Defense (CND), and Computer Network Exploitation (CNE). CND is about computer and network protection, whereas CNE is focused on information gathering, that is, in cyber espionage, and CNA is related to degradation, disruption, destruction, or manipulation actions. The offensive CNO are those related to CNA and CNE, and they are both defined by a series of mandatory steps for the operation to be successful. In each of these steps, a set of tactics, which define what an hostile actor is doing, are performed by specific techniques, which define how the hostile actor accomplishes a tactic.

In these offensive operations, the initial access, or delivery, is the first mandatory step in which the hostile actor approaches its target in an offensive way and has a direct contact with it. Although some reconnaissance techniques, a previous step, can be also executed with a hostile approach to the target, they are not always mandatory in an operation, as reconnaissance can be achieved by information-gathering techniques, such as passive or semi-passive, which are not considered hostile [587].

The delivery being the first mandatory hostile approach to a target, it is, as well, the first moment where an operation can be detected and neutralized; in fact, delivery is considered a high-risk task for hostile actors, as it leaves traces in the target [675]. For this reason, it is a must for defenders to understand how delivery is performed by threat actors. Without a clear understanding of the tactic and their associated techniques, defense is harder and the success rate for the threat actor increases.

Our paper provides a suitable taxonomy for techniques exploited to achieve the delivery tactic; we have dissected the tactic and identified its key elements, defined them, and designed a taxonomy for them. These key elements are the malicious objects delivered to a target (delivery object) in a specific way (delivery vector) and that break the security perimeter of the target in a specific way (delivery path). With these three elements, we can classify all delivery techniques, and we can identify as well different approaches not commonly exploited but which would allow a hostile actor to achieve persistence.

The contributions of this paper are summarized as follows:

- To identify and define the key elements that compose the delivery tactic.
- To structure the delivery tactic approaches used in offensive Computer Network Operations (CNE or CNA).
- To provide an accurate taxonomy for techniques into the delivery tactic for these operations, thus allowing defenders to detect novel or uncommon techniques, identify specific countermeasures, and improve global security.

The rest of the paper is organized as follows. The background, Section 2.2, provides a brief introduction to the Cyber Kill Chain and to the MITRE ATT&CK framework, as two of the main frameworks for the modeling of offensive operations, both in their steps and in their tactics and techniques. In Section 2.3, we assess the problem of the lack of a unified taxonomy for the delivery tactics and its importance for the modeling of threat actors and operations. Section 2.4 analyzes the prior work in this field, stating that little research has been conducted in this sense. In Section 2.5, we propose a novel taxonomy for the techniques inside the delivery tactic, identifying the key aspects to classify particular techniques. In Section 2.6, we discuss the results of our work, comparing them with other approaches and identifying improvements, as well as future research lines. Finally, Section 2.7 summarizes the outcome of the overall work.

## 2.2 Background

### 2.2.1 Mitre ATT&CK

MITRE ATT&CK (Adversarial Tactics, Techniques, and Common Knowledge) is a globally accessible knowledge base of adversary tactics and techniques based on real-world observations. This knowledge, contributed by analysts all around the world, can be used as the base for the development of specific threat models and methodologies. Started in 2013 and published in 2015, ATT&CK develops a process for modeling an adversary's post-compromise behavior at a fine level. A description of the framework and the work performed can be found at [598].

Tactics specify what a threat actor is doing, at the highest level of description, to accomplish a certain mission. Techniques specify how tactics are implemented, and procedures describe a particular implementation of a technique. These tactics, techniques, and procedures represent the behavior of a threat actor from the highest level description (tactic) to the lowest level one (procedure). MITRE ATT&CK framework is today's de facto standard to structure tactics and techniques of advanced threat actors. As of March 2022, MITRE ATT&CK had defined 14 enterprise tactics—those related to the activities of an attacker onto its victim—and 188 enterprise techniques associated with those tactics and 379 subtechniques. Apart from that, MITRE ATT&CK defines 14 mobile tactics, related to the compromise of mobile devices, and 92 mobile techniques. Beside tactics and techniques, ATT&CK identifies software (a generic term for tools, artifacts, malware, etc.) that can be used to implement one or more of the techniques, and which is out of the scope of this work.

In the ATT&CK Matrix for Enterprise, the framework represents tactics as the adversary's tactical goals for acting [672]. Although ATT&CK does not provide a kill-chain approach to specify the arrangement of tactics, most of them are presented in the logical order that a threat actor follows in hostile operations. All of them can be achieved through different techniques, and a single technique that can be associated with one or more tactics. There is no formal structure in MITRE ATT&CK for techniques in each tactic, all of them being represented in a plain view. For example, for the Command and Control tactic, representing the goal of enabling the remote control of the compromised infrastructure, the framework identifies techniques such as Data Encoding, Data Obfuscation, Protocol Tunneling, or Remote Access Software. The structure of tactics and techniques in MITRE ATT&CK allows analysts to organize which adversarial actions belong to specific techniques and tactics, thus helping defensive teams to understand what a threat actor may be trying to achieve, how this actor is trying to achieve it and how to better defend against the threat [19].

MITRE ATT&CK also links Advanced Persistent Threat groups, APTs, to tactics, techniques, and software. With 110 identified groups at the time of this writing, all of them are named, aliased, described, and linked to specific techniques (including pre-attack and mobile) and software. In this way, an analyst can establish relationships between those entities to model an adversary and its activities against a target and, most importantly, to establish the defense mechanisms to prevent, detect, and respond to a threat.

MITRE ATT&CK represents an enormous effort to provide to the community a unified framework to identify the activities of advanced threat actors, from their TTP to the software they use, correlate information among those entities, and improve, not only the knowledge about APT, but also the defense mechanisms required for their detection and response. It constitutes a framework that, as usual, has to be improved with continuous work and contributions; in this sense, we miss in the MITRE ATT&CK a more defined structure for techniques inside each tactic. The standard specifies all tactics for a cyber kill-chain model but, for each tactic, all related techniques have a plain structure, broken only by the specification of sub-techniques and, particularly, implementations of a specific technique.

### 2.2.2 Cyber Kill Chain ®

The Cyber Kill Chain® framework [284], developed by Lockheed Martin, is part of the Intelligence-Driven Defense® model for the identification and prevention of cyber intrusion activity, identifying what a threat actor must complete in order to achieve its ive. It was first described in [284] as a seven-step process suitable for CNA or CNE operations, as shown in Figure 1.2.

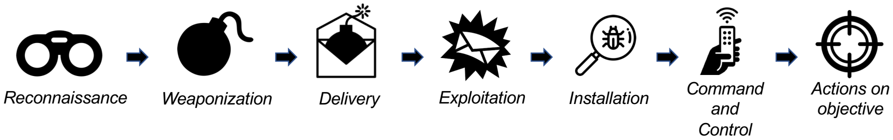

Figure 2.1: Cyber Kill Chain® as defined in [284].

These seven steps are defined as follows [219, 284]:

1. Reconnaissance. Research, identification, and selection of targets.
2. Weaponization. Before attacking a target, the threat actor has to couple a remote access Trojan with an exploit into a deliverable payload.
3. Delivery. The transmission of weapons to the targeted environment to launch a particular operation.
4. Exploitation. After the weapon is delivered, the exploitation triggers an intruders' code.
5. Installation. The installation of an implant, just as a remote access Trojan, a backdoor, or any kind of malicious software, on the victim system, allowing the adversary to maintain persistence inside the environment.
6. Command and Control. The compromised hosts must beacon outbound to an Internet controller server to establish a C2 channel, thus allowing the threat actor to remotely control its target.
7. Actions on objectives. After progressing through the first six phases, the intruders can take actions to achieve their original goals, such as information theft, denial, or hop to a third-party infrastructure.

The cyber kill chain represents an industry-accepted methodology for understanding how an attacker will conduct the activities necessary to cause harm to an organization and has been largely discussed [430] (in ref. [684], the authors identify some of the discussions regarding the application of the Cyber Kill Chain). Some authors [92, 350] have proposed the addition and removal of different stages in order to improve or adjust the original model, and the topic has also been discussed in technical conferences. Moreover, some efforts to unify models and variants, such as [489], have been made. Despite this, the original proposal has been widely used and applied to specific problems regarding advanced threat actors, such as those related to the modeling of the attack stages against critical infrastructures [112, 253, 379, 686].

## 2.3 Problem Statement

In the modeling of offensive operations, the establishment of an accurate sequence of actions and the identification of the tactics and techniques that threat actors perform in order to achieve their goals is a key requirement for the prevention, detection, and neutralization of the threat. The analysis of the initial foothold into a target infrastructure is a must for this modeling of hostile activities. The Cyber Kill Chain "Delivery" stage or the MITRE ATT&CK "Initial access" tactic both represent this point of initial contact between the threat actor and its target.

The delivery tactic is usually linked to the delivery of a malicious payload to the target, embedded into a weaponized such as a particular file type or web link. In fact, in many references delivery is just identified as "payload delivery" [119,179,581] or "malware delivery" [347,496]. The delivered has been generated on a previous stage of the operation, usually called weaponization [284], and, after a successful delivery, the payload is detonated, starting the next stage of the attack, which is commonly identified as exploitation.

This focus on the delivery of malware or malicious payloads has relevant limitations. It refers only to specific delivery techniques, while other ones, whose use is increasing on a daily basis, are not considered in this concept. Malwareless operations do not use malicious payloads, not only to achieve the delivery tactic, but none of the required tactics perform a successful operation.

Nowadays, threat actors' delivery techniques include not only the weaponization of a malicious payload, but also different approaches, some of them even without a malicious payload, that allow hostile actors to accomplish their goal: to break the target's security perimeter and to open the way for the exploitation stage.

We have analyzed different approaches for the identification of a suitable up-to-date classification scheme for delivery techniques, especially for the ones not linked to the use of malicious code and simply being the abuse of legit resources. In Section 2.4 we present a summary of these approaches. However, no suitable approximation for such a structure that allows analysts to detect an ongoing operation has been identified. It is mandatory to analyze, understand, and identify the different techniques for delivery used nowadays, in order to be able to detect and neutralize them. Such a structure would allow analysts to identify gaps in their security countermeasures and to discover new techniques deployed for the delivery tactic.

## 2.4 Techniques and Limitations

Until this moment, no valid complete taxonomy for delivery techniques has been identified. All approaches are partial, providing a relationship of delivery techniques but without a particular structure, focusing on an specific type of delivery, or analyzing just particular delivery techniques used by a given threat actor. We miss a global structure to classify delivery techniques, which allows not only this classification but also the identification of security gaps in the monitoring scheme for an organization. Such a taxonomy would allow analysts to detect the compromise of their infrastructures, as well as to identify those monitoring gaps and deploy countermeasures against previously unknown techniques.

As we have stated before, MITRE ATT&CK is the key reference for the identification of delivery techniques, identified as Initial Access in the framework. It focuses on the adversary tactics, techniques, and procedures (TTP) derived from real attacks [604]. However, this framework does not provide any classification for these techniques, exposing just a plain relationship for them and their subtechniques. Although, such as an approximation is useful for the identification of particular techniques used in offensive operations, it lacks a whole structure, so it can not be taken as a valid reference for the definition of a delivery techniques' taxonomy.

The Cyber Kill Chain represents a starting point that provides an intelligence framework for understanding the multistage attack. It identifies the "Delivery" stage of a hostile operation but it does not provide any information about adversarial tactics and techniques. Following the Cyber Kill Chain stages, and with different kill chain models, many works [43, 51, 166, 408, 489] provide specific examples of delivery techniques; however, few of them discuss these techniques from a threat modeling perspective. In addition to particular examples, no valid taxonomy approach has been identified in these works. In the case of specific operations related to APT activities, we face the same situation; different works [64, 439, 627] analyze the delivery techniques used by Advanced Persistent Threats in their campaigns, providing particular examples of techniques, but none of them present a taxonomy for the delivery tactic. With a more general approach, ref. [503] analyzes life cycles and models for Advanced Persistent Threat operations.

In [123], Ping Chen et al. identify two types of delivery techniques: direct and indirect delivery. In direct delivery, the hostile actors send exploits to their targets, while in indirect delivery, they compromise a third party that is trusted by the target, and then they use this compromised third party to indirectly serve exploits to the target. For each category, the authors propose spear phishing and the watering hole, respectively, as examples of techniques. In this work, the goal of the authors is not to provide a taxonomy for delivery techniques, but to present a general survey on Advanced Persistent Threats; for this reason, they do not focus on a particular taxonomy. In addition, as the paper is dated to 2014 we consider this a valid initial approach that needs to be enhanced; over the years, advanced threat actors have developed new techniques for the delivery tactic that have to be considered: for example, the exploiting of public facing applications. A similar approach is used in [570], although, in this case the provided examples do not reflect the used categories.

As specified in Section 2.3, most analyses are focused on the delivery of a malicious payload through multiple ways. Malware delivery campaigns have been analyzed in [687], in which Ziyun Zhu et al. adopt three stages from the STIX data model (exploitation, installation and command, and control) to represent malware delivery campaigns. As this approach is focused on modeling whole malicious campaigns, the delivery tactic is linked to other stages inside an operation, which is an approach inappropriate for our particular focus on the analysis of the delivery tactic.

Being a malware with high impact campaigns during the last few years, ransomware delivery has been specifically analyzed in different works. In [217], Keertika Gangwar et al. propose an analysis and detection approach of ransomware based on its delivery mechanisms. The authors provide a feature selection and extraction from different ransomware families, but they do not identify any classification for the delivery techniques used in each case, focusing only on URL and indicators of compromise linked to ransomware. In [515], Pratyush Raunak et al. focus on the detection of ransomware delivered through a specific technique, namely, exploit kits. In [166], Tooska Dargahi et al. discuss the delivery of ransomware by techniques linked to social engineering, malvertisement, and traffic distribution systems, providing an analysis for each of them. None of these works tries to establish a complete taxonomy for ransomware delivery techniques.

General delivery techniques have been analyzed in-depth, taking into account both their description and their countermeasures. General social engineering attacks, on which many delivery techniques are based, are modeled with graphs in [60]. Phishing is modeled in works such as [211, 301, 348] or [549]. Watering hole techniques are described in [342, 599]. Even less common delivery techniques, such as baiting [83, 127] or supply chain compromise [153, 516, 678] have been deeply analysed. Please note that although not widely used until now, supply chain compromise is an increasing trend for threat actors, as we will detail later in this paper.

As stated before, social engineering is the base of many delivery techniques; apart from its modeling, different taxonomies have been developed for social engineering attacks. In [295], Koteswara Ivaturi et al. divide social engineering techniques into two main categories: person to person and person to person via media. In [343], Katharina Krombholz et al. establish three parameters for a taxonomy definition: type, operator, and channel, while [267] presents a kill chain to classify social engineering attacks, with three mandatory steps: orchestration, exploitation, and execution. In [21], Hussain Aldawood et al. differentiate two parameters to establish an accurate taxonomy: who or what attacks are based on (human or technology) and how they are executed (physical, technical, social, and socio-technical). In this work, the authors also provide accurate examples of particular techniques for many social engineering attacks. These works present different taxonomies for social engineering attacks; in [199] we can find a summary of them, and a survey is presented in [541], where Fatima Salahdine et al. provide accurate classification schemes for this technique. However, not all delivery techniques include social engineering; therefore, we must generalize our taxonomy in order to provide a global valid approach for the delivery tactic.

As phishing is one of the most widely used techniques for the delivery tactic, it has been largely analyzed and its particular implementations have been classified. In [120], Junaid Ahsenali Chaudhry et al. identify four phishing techniques: spear phishing, clone phishing, malware-based phishing, and search engine phishing. This approach is purely focused on specific technical aspects in order to identify phishing countermeasures; therefore, it is not suitable to establish a taxonomy, even an initial one. In [513], Justinas Rastenis et al. define a taxonomy for e-mail based phishing attacks, based on different features of the malicious e-mail used in an operation. The authors highlight the lack of e-mail-based phishing attacks' taxonomy and propose a novel one, but only with a focus on this particular technique; thus, it can not be considered a general purpose approach that is valid for all delivery techniques. In [642], Gaurav Varshney et al. propose five categories for web phishing techniques (the authors define them as “Tactics”, an approximation that we consider inconsistent with the current definitions of tactics and techniques): Spoofing website text and images, web link manipulation, malicious use of scripting languages, Java Script popups, fake address bars, and utilizing browser vulnerabilities. However, this approach focuses only on web phishing, and does not consider other phishing techniques.

The literature review for delivery techniques and their classification, detailed in this section, can be grouped into six main families:

- General models, such as the MITRE ATT&CK or kill-chain models, where delivery is considered as a tactic that can be performed through different approaches.
- Threat actors reports, which analyze specific delivery techniques exploited in real operations.
- Initial classification approaches, which try to propose a classification scheme for delivery techniques but whose focus is not this scheme.
- Malware-focused analysis, which present the mechanisms used to deliver general or particular malware samples to a target.
- General techniques' description, which provide an analysis for well-known delivery techniques in a general context, typically a whole hostile operation, without delving into the delivery internals.

- Particular techniques analysis, which provide an in-depth dissection of specific techniques such as social engineering or phishing, or of particular elements such as malware.

In Table 2.1, a tabular comparative study featuring the main characteristics of the different analyzed approaches is shown.

Table 2.1: Comparative literature study.

| Family | References | Pros | Cons |
|---|---|---|---|
| General models | [43, 51, 166, 408, 489] | Industry standards<br>Specifications based on real attacks<br>Useful for the identification of particular techniques in offensive operations | Plain structure<br>Not an in-depth analysis<br>Not designed for the identification of security gaps |
| Threat actors reports | [64, 439, 503, 627] | Real world cases<br>In-depth analysis for each case<br>New delivery techniques are presented | Focus on specific delivery techniques<br>No structure proposal |
| Initial classifications | [123, 570] | Initial structure for techniques<br>Real techniques mapping | Not focused on delivery, but general approaches<br>Date of publication |
| Malware-focused | [166, 217, 515, 687] | In-depth analysis for malware delivery<br>Detection oriented<br>Malware as a relevant threat | Focus on specific approach: malware and malware related<br>Not designed to identify security gaps outside malware ecosystem |
| General techniques | [60, 83, 127, 153, 211, 301, 342, 348, 516, 549, 599, 678] | Delivery techniques analyzed with a global operation perspective: completeness<br>Real world cases | No classification proposal<br>Not in-depth analysis |
| Particular techniques | [21, 120, 199, 267, 295, 343, 513, 541, 642] | In-depth analysis of particular techniques<br>Structure proposals for these techniques | Focus on particular techniques, without considering a global delivery scheme |

## 2.5 Our Proposal

Initial access, or delivery, is just the compromise of the target security perimeter. This security perimeter is the boundary within security control measures, which are in effect to protect assets [13]. The compromise is always a break in from outside the target premises to inside them, no matter where the tactic is initiated from; please note that in our proposal, we identify inbound and outbound logical paths for delivery, but this classification refers only to the initial connection for a logical compromise, not to the whole tactic. To compromise the target perimeter, we identify three mandatory elements: an artifact, a transport vector from this artifact to the target's premises, and a path to break the perimeter. In this sense, in order to establish a taxonomy for delivery techniques, we provide the following definitions:

> Delivery Object, is the object used to break the target's perimeter. This object is usually a deliverable artifact generated by a weaponizer, in which the malicious payload is embedded, typically in the form of an application data file such as Adobe Portable Document Format or Microsoft Office. However, our concept of object includes not only ad hoc, malicious artifacts, such as links or files, but also points out the infrastructure to be abused by hostile actors, such as public facing services.
>
> Delivery Vector, is the transport used to deliver the artifact to its target. Examples of delivery vectors include USB memory drives, mail messages, hardware implants, or supply chains.
>
> Delivery Path, is the way the delivery vector breaks the target's perimeter. Although in some cases this delivery path is directly linked to the delivery vector, in other cases they are independent, as we well analyze later in this work. Examples of delivery paths include both physical and logical routes to the target.

To achieve the delivery tactic, an object is used or abused by a threat actor, who delivers it to its target by a delivery vector and breaks the perimeter through a specific path. The path identifies which point is compromised, the vector identifies how it is compromised, and the object identifies what is used to achieve the compromise. For all of these three concepts, we propose a taxonomy based both on the characterization of the techniques and on the related countermeasures to prevent and to detect them. We have not identified any delivery technique that does not have all of these items, so we consider all of them as mandatory for a successful delivery.

The delivery object is divided into two types: dynamic and static objects. This proposed division reflects the techniques exposed by threat actors in the use of malicious artifacts and in the abuse of infrastructure features to achieve the Initial Access. Dynamic objects are those that contain a malicious payload that detonates when they are used or accessed in some form. We identify two sub types of dynamic objects: weaponized and not weaponized, depending whether the malicious payload is coupled with another object. The first ones are weaponized by the threat actor, for example in the form of a link or file (logical weaponization) or directly into removable media (physical weaponization), and sent to the target in order to exploit a local vulnerability or weakness when they are accessed. On the other hand, non-weaponized dynamic objects are directly launched against the target without a couple, usually to exploit a remote vulnerability or weakness in order to break the security perimeter. We consider especially relevant this division of dynamic objects, as weaponization or its absence is a key factor for the technique; weaponized objects are usually linked to operations in which the end user is deceived, as the couple hides the malicious payload, while non-weaponized objects are usually related to the exploitation of technical vulnerabilities. Moreover, the countermeasures to be applied in each case are different: when dealing with weaponized objects, where the target is a human, awareness is a must, but when dealing with non-weaponized objects, where the target is an infrastructure, the main countermeasures are related to technical vulnerability management.

Static objects are those that do not contain a malicious payload, but they are abused by the threat actor in order to break the security perimeter of its target. In this case, we also identify two sub types of static objects: those that pre-exists in a legitimate form in the target's infrastructure, with independence from the threat actor and those not legitimate, which are generated by the actor. The first ones are just abused by the threat actor in an operation against a target, facing as a legitimate user, and the second ones are ad hoc generated, in the form of backdoor, to provide the threat actor with a delivery, and in some cases an exploitation and a persistence or capability; this generation is typically performed by exploiting a particular vulnerability or weakness of the target. Again, this division is relevant for the identification of the used techniques: legit objects are pre-existing ones, just abused by a threat actor, and the detection of these techniques are usually based on anomaly patterns, while illegitimate objects are created by the exploitation of a vulnerability or weakness, and their detection is mostly based on misuse patterns.

Typical examples of static objects include external-facing accounts, both legit abused by the actor (pre-existing), and generated for the delivery.

Delivery object-proposed taxonomy is shown in Figure 2.2. As we have previously stated, these objects are used or abused by a threat actor to get initial access to its target. We can identify the two main families of delivery objects: dynamic and static. Inside both of them, we differentiate the specific types of objects that we have detailed in our work.

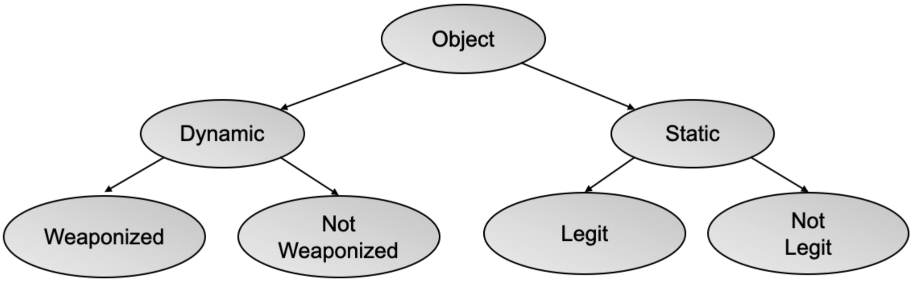

Figure 2.2: Delivery Objects' taxonomy.

Taking into account the delivery vector, we differentiate between those that compromise the target's infrastructure in first place and those that compromise a third party infrastructure to, as a second step, achieve the compromise of the target. This approach is linked to [123], in which Ping Chen et al. define direct and indirect delivery as a starting point for an initial classification, as stated before, and it is consistent with the identification of the first contact between the threat actor and its target. We define two sub types of direct vectors: those that are deliberate and those that are unintentional. In the first one, the threat actor abuses an insider from the target in order to use or abuse the delivery object. In unintentional delivery, the threat actor can abuse a non-malicious element in two ways: the syntactic (an actual vulnerability or weakness in the infrastructure) or semantic (a decoy to a user, for example in spear phishing attacks). Please note that syntactic or semantic refers only to the delivery, not to the exploitation: once delivered, a semantic attack can either exploit a vulnerability or not. This proposed division of direct delivery represents the main differences in the Initial Access techniques from the perspective of how the target is compromised, and provides an uncommonly in depth analyzed perspective: those of the insiders.

In indirect delivery, we find the compromise of trusted and untrusted third parties. A hostile actor can compromise an untrusted infrastructure that, indirectly, compromises the actual target on some interaction between them, typically through web navigation over a malicious web page. On the other hand, the compromise of a trusted third party and the exploiting of a trusted relationship with the actual target allow the threat actor to achieve delivery in a more concise way. This compromise of a trusted third party is usually in the form of compromise of a supply chain or in the compromise of a third party with an infrastructural trust point with the target. Please note that, from the compromised party's point of view, an indirect delivery vector must be seen as direct. The difference between trusted and untrusted parties is a key one, as a trusted party will have more available attack surface against the target than an untrusted one; thus, countermeasures against them must be more strict.

The delivery vectors' proposed taxonomy is shown in Figure 2.3. Those vectors represent the transport used to use or abuse the delivery object in order to get initial access to the target. These vectors can be exploited in their own target (Direct) or in a party which has some kind of relationship with the final target (Indirect). Both of them have particular vectors, and we consider especially relevant the deliberate vector, representing the use of insiders in a target organization that voluntarily provide initial access to a threat actor. These insiders are not generally considered in literature, as we will detail later in our work.

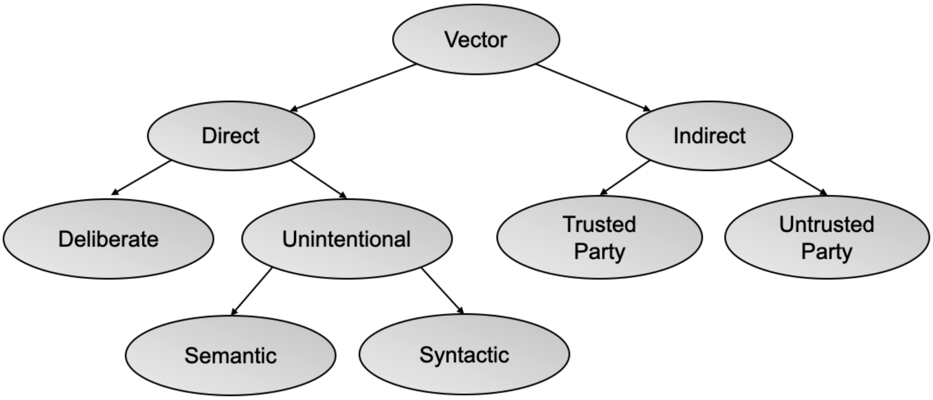

Figure 2.3: Delivery Vectors' taxonomy.

Finally, from the perspective of the delivery path, we identify two main views for an organization security perimeter: the logical and the physical view. The human fact is a key piece in security that has been largely analyzed [36, 147, 240], and some even works [44, 600, 639] define a human perimeter for organizations. However, we will focus on the logical and physical perimeters: although many delivery techniques, such as spear phishing, rely on human vulnerabilities, the delivery itself has to be performed, breaking a logical or a physical protection. Please note that the concept "perimeter" refers not only to the target's owned infrastructures, but also to the infrastructures related to the target and used while delivering, for example, to cloud services. So in this sense, we can define two main types of delivery path: the logical and the physical.

Inside the logical delivery path, we differentiate inbound and outbound paths, depending on from where the first malicious connection for the delivery is started. If this first connection is started from the organization's premises, such as in phishing techniques, we refer to the outbound logical path, while if it is started outside these premises, such as in the abuse of valid accounts, we refer to the inbound logical path. This distinction is specially relevant for our taxonomy, as it allows us to fine tune the identification of particular techniques and their identification, thus enabling organizations to establish better protection mechanisms against the threat. In both cases, it is mandatory to monitor external-facing systems to detect the compromise, but when dealing with inbound logical path techniques, the monitoring efforts must be directed to applications and services exposed to internet. If we deal with an outbound logical path, these monitoring efforts must be directed to the legit external services offered to their own organization.

Regarding the physical delivery path, we differentiate between connected and not connected approaches; in this case, a not connected physical delivery path is the one that breaks the physical security perimeter of the organization once, not connecting the delivery object to the target IT infrastructure. On the other hand, a connected physical delivery path represents a double security perimeter violation: the first one, the physical, gets into the target, and the second one, logical, connects the delivery object to the target infrastructure. Again, this difference is relevant for the identification of the particular technique used by a threat actor and thus for the identification and deployment of appropriate countermeasures. While dealing with connected approaches, in which we find a double perimeter compromise, it is possible to establish two layers of security countermeasures. On the other hand, in not connected approaches there is a single compromise, so from the perimeter's point of view only a checkpoint can be established, although other countermeasures, not specifically perimeter-related, can also be deployed, especially for the monitoring of the object activities in the target.

The full delivery path-proposed taxonomy is shown in Figure 2.4. As we have previously stated, the delivery path represents the compromised perimeter for the target, the physical and logical perimeters being the considered ones in our work. Inside both of them we identify the different approaches that detail this compromise, as these approaches allow analysts to detect known delivery techniques and to identify novel ones.

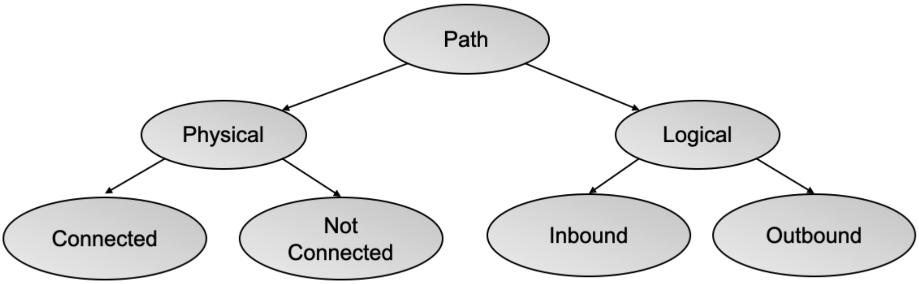

Figure 2.4: Delivery Paths' taxonomy.

By setting the object, the vector, and the path we can classify all delivery techniques into our taxonomy. These items provide a global view of the tactic, as they themselves define the delivery process: an object is sent through a specific vector to the target, physically or logically breaking its perimeter, to achieve the tactic goal. In Table 2.2, we summarize the proposed taxonomy.

Table 2.2: Proposed taxonomy for delivery techniques.

| | | | |
|---|---|---|---|
| Object | Dynamic | Weaponized | |
| | | Not Weaponized | |
| | Static | Legit | |
| | | Not legit | |
| Vector | Direct | Deliberate | |
| | | Unintentional | Syntactic |
| | | | Semantic |
| | Indirect | Trusted Party | |
| | | Untrusted Party | |
| Path | Physical | Connected | |
| | | Not Connected | |
| | Logical | Inbound | |
| | | Outbound | |

Following our proposed approach, we can classify all the delivery techniques and identify security gaps to establish countermeasures against them. We have selected two of the most widely used techniques for Initial Access, spear phishing, and the abuse of valid user accounts on external-facing infrastructure. We can classify them into the proposed categories for delivery object, vector, and path, as an example regarding how our proposed taxonomy can be applied.

Spear phishing is a particular type of phishing, in which the target and context are previously investigated so that the email is tailored to the receiver [99, 517]. These actions are executed by sending a malicious object (typically a file or a link) by e-mail to a particular target [474]. This object is specially crafted to detonate when it is accessed by the receiver, who opens it as a legitimate e-mail; when it detonates, the malicious payload is executed and the threat actor can continue with further steps of the cyber-kill chain. Following our proposed taxonomy, this technique uses a weaponized dynamic object, is based on a semantic unintentional direct vector, and follows an outbound logical path.

Regarding the abuse of valid user accounts on external-facing infrastructure, in this case the threat actor just obtains valid credentials and abuses them to remotely access the targeted infrastructure. For example, these credentials can be guessed by brute force or obtained from a data leak. Once the threat actor gets these credentials, it has access to the infrastructure and can start the execution of the rest of the cyber-kill chain actions. In our taxonomy, this technique uses a legit static object, is based on a semantic unintentional direct vector and follows an inbound logical path.

Table 2.3: MITRE ATT&CK Initial Access techniques.

| Technique ID | Name | Sub-Techniques |
|---|---|---|
| T1189 | Drive-by Compromise | N/A |
| T1190 | Exploit Public-Facing Application | N/A |
| T1133 | External Remote Services | N/A |
| T1200 | Hardware Additions | N/A |
| T1566 | Phishing | - Spearphishing Attachment<br>- Spearphishing Link<br>- Spearphishing via Service |
| T1091 | Replication Through Removable Media | N/A |
| T1195 | Supply Chain Compromise | - Compromise Software Dependencies and Development Tools<br>- Compromise Software Supply Chain<br>- Compromise Hardware Supply Chain |
| T1199 | Trusted Relationship | N/A |
| T1078 | Valid Accounts | - Default Accounts<br>- Domain Accounts<br>- Local Accounts<br>- Cloud Accounts |

### 2.5.1 Mapping to MITRE ATT&CK

As stated in this work, MITRE ATT&CK is the main public effort to establish a classification of TTP used by threat actors; for this reason, we have performed a mapping of the MITRE ATT&CK Enterprise "Initial Access" (the given name in the framework to the delivery) tactic onto our proposed structure.

At the time of this writing, MITRE ATT&CK Enterprise "Initial Access" tactic (last modified on 19 July 2019), identified as TA0001, consists of techniques that use various entry vectors to gain an initial foothold within a network. Techniques used to gain a foothold include targeted spear phishing and exploiting weaknesses on public-facing web servers. Footholds gained through initial access may allow for continued access, such as valid accounts and the use of external remote services, or may be limited-use, due to changing passwords. For this particular tactic, MITRE ATT&CK identifies the techniques shown in Table 2.3.

T1189, Drive-by Compromise, refers to the threat actor gaining access to a system through a user visiting a website over the normal course of browsing. In this technique, the object is a dynamic and weaponized; the vector is indirect, through an untrusted party; and the path is logical and outbound.

T1190, the Exploit Public-Facing Application, refers to the threat actor taking advantage of a weakness in an Internet-facing computer or program using software, data, or commands in order to cause unintended or unanticipated behavior. In this technique, the object is dynamic and not weaponized; the vector is direct, unintentional, and syntactic; and the path is logical and inbound.

T1133, External Remote Services, refers to the threat actor leveraging external-facing remote services to initially access and/or persist within a network. In this technique, the object is static and legit; the vector is direct, unintentional, and syntactic; and the path is logical and inbound.

T1200, Hardware Additions, refers to the threat actor introducing computer accessories, computers, or networking hardware into a system or network that can be used as a vector to gain access. In this technique, the object is static and legit; the vector is direct and unintentional; and the path is physical and connected.

T1566, Phishing, refers to the threat actor sending phishing messages to gain access to victim systems; all forms of phishing are electronically delivered social engineering. T1566 presents three sub-techniques, and in all of them the object is dynamic and weaponized; the vector is direct, unintentional, and semantic; and the path is logical and outbound.

T1091, Replication Through Removable Media, refers to the threat actor moving onto systems, possibly those on disconnected or air-gapped networks, by copying malware to removable media and taking advantage of Autorun features when the media is inserted into a system and executed. In this technique, the object is dynamic and weaponized into removable media; the vector is direct and unintentional; and the Path is physical and connected.

T1195, Supply Chain Compromise, refers to the threat actor manipulating products or product delivery vectors prior to receipt by a final consumer for the purpose of data or system compromise. In this technique, the object can be dynamic or static, depending on how the supply chain is abused; the vector is indirect, through a trusted party; and the Path can be physical or connected (sub-technique T1195.003) or logical and inbound (sub-techniques T1195.001 and T1195.002). This a clear case where MITRE ATT&CK does not delve into the particularities of technique.

T1199, Trusted Relationship, refers to the threat actor breaching or otherwise leveraging organizations who have access to intended victims. In this technique, the object is static, as it does not contain malware for the delivery and legit, as exploited trusted relationships pre-exist in the infrastructure. The vector is indirect through a trusted party and the Path is logical and outbound. As in T1195, again MITRE ATT&CK does not provide enough information to delve into a more specific classification of this particular technique.

T1078, Valid Accounts, refers to the threat actor obtaining and abusing credentials of existing accounts as a means of gaining Initial Access, Persistence, Privilege Escalation, or Defense Evasion. In this technique, the object is static and legit in all of the sub-techniques exposed by MITRE ATT&CK; the vector is direct, unintentional, and syntactic; and the path is logical and inbound.

Table 2.4: MITRE ATT&CK techniques classification by threat group.

| | | | |
|---|---|---|---|
| Object (T1195) | Dynamic | Weaponized (T1189, T1566, T1091) | |
| | | Not–Weaponized (T1190) | |
| | Static | Legit (T1078, T1199, T1133) | |
| | | Not legit (T1200) | |
| Vector | Direct | Deliberate | |
| | | Unintentional | Syntactic (T1190, T1133, T1200, T1091, T1078) |
| | | | Semantic (T1566) |
| | Indirect | Trusted Party (T1195, 1199) | |
| | | Untrusted Party (T1189) | |
| Path | Physical | Connected (T1200, T1091, T1195.003) | |
| | | Not Connected | |
| | Logical | Inbound (T1190, T1133, T1195.001, T1195.002, T1078) | |
| | | Outbound (T1189, T1566, T1199) | |

As we can see, all MITRE ATT&CK techniques for "Initial Access" can be mapped onto our taxonomy; this mapping is summarized in Table 2.4. If we analyze each of the elements from our taxonomy and their mapping to MITRE ATT&CK, we get interesting findings about the framework, especially those related to important gaps.

Regarding the delivery object, all techniques but one are classified into low-level nodes. The technique in upper nodes is T1195, Supply Chain compromise. MITRE ATT&CK does not provide enough information for its fine classification, although it is becoming a commonly used technique and the framework defines three particular sub techniques inside it. In fact, supply chain attacks increased in number and sophistication in the year 2020 and this trend has continued in 2021, posing an increasing risk for organization [18, 145, 194]. Therefore, the number of published research works related to supply chain cyber security is increasing in last years, especial since COVID–19 pandemic [354, 444]. We find it mandatory to provide a more exhaustive analysis of this technique from the perspective of the delivery object, identifying the different approaches through static and dynamic objects that this technique can be executed with. This analysis would provide organizations better countermeasures to prevent, detect and neutralize supply chain attacks.

Regarding the delivery vector, all techniques are classified into low-level nodes. In this case we find an important gap; as MITRE ATT&CK does not consider deliberate delivery, this is the abuse of an insider from the target organization voluntary helping the hostile actor to achieve its goals. Although not usually considered when dealing with delivery, the insider threat has been a relevant problem for years; thus, we find it mandatory to identify delivery techniques (including the human aspects) for this threat. Dealing with the delivery vector, it is also interesting that phishing, being the most-used delivery technique, is the only one identified as semantic, as most techniques are classified as syntactic. The framework should delve into semantic delivery vectors, as people have always been the weakest link in the security chain and there are different delivery techniques that benefit from social engineering, as we have discussed in Section 2.6.

Finally, regarding the delivery path, MITRE ATT&CK identifies delivery techniques both for a physical and a logical compromise, although most of the delivery techniques analyzed from hostile actors reports are based on a logical delivery. We must highlight the fact that the not connected physical delivery path has no linked techniques; this means that these kinds of delivery approaches are not considered in MITRE ATT&CK; thus, most defensive teams are also not considering them in their security countermeasures. The framework should delve into different techniques to physically break the security perimeter of an organization and deploy an autonomous, not connected implant.

If we delve into the particular delivery techniques that threat actors are performing, the data from MITRE ATT&CK allows the linking between groups and techniques. In Table 2.5, this relationship is shown, exposing the number of threat groups using each initial access technique or sub technique.

Table 2.5: Delivery techniques exploited by threat actors.

| **Technique ID** | **Name** | **Number of Groups** |
|---|---|---|
| T1189 | Drive-by Compromise | 24 |
| T1190 | Exploit Public-Facing Application | 16 |
| T1133 | External Remote Services | 20 |
| T1200 | Hardware Additions | 1 |
| T1566.001 | Spear phishing Attachment | 64 |
| T1566.002 | Spear phishing Link | 34 |
| T1566.003 | Spear phishing via Service | 7 |
| T1091 | Replication Through Removable Media | 4 |
| T1195.001 | Compromise Software Dependencies and Development Tools | 0 |
| T1195.002 | Compromise Software Supply Chain | 6 |
| T1195.003 | Compromise Hardware Supply Chain | 0 |
| T1199 | Trusted Relationship | 5 |
| T1078.001 | Valid Accounts: Default | 0 |
| T1078.002 | Valid Accounts: Domain | 10 |
| T1078.003 | Valid Accounts: Local | 8 |
| T1078.004 | Valid Accounts: Cloud | 2 |

Please note that a single threat group can be executing more than one initial access technique, and also that those techniques or sub-techniques with zero groups mean only that they have been observed in real operations but attribution has not been possible. With this information, we can conclude that different spear phishing techniques are the most widely used, while supply chain compromises can be only performed by specific threat actors; particularly, no identified threat actor is able to compromise the hardware or dependencies and developmental tool supply chains, although MITRE ATT&CK identifies these techniques as the ones used to gain initial access. Of course these results are consistent with the findings we have exposed in our paper, related to phishing as a key delivery technique and supply chain compromise as a growing trend. As we have previously performed with MITRE ATT&CK techniques, we can map the number of groups performing each technique onto our taxonomy, as shown in Table 2.6.

Table 2.6: MITRE ATT&CK groups classification.

| | | | |
|---|---|---|---|
| Object (6) | Dynamic | Weaponized (133) | |
| | | Not–Weaponized (16) | |
| | Static | Legit (45) | |
| | | Not legit (1) | |
| Vector | Direct | Deliberate | |
| | | Unintentional | Syntactic (61) |
| | | | Semantic (105) |
| | Indirect | Trusted Party (11) | |
| | | Untrusted Party (24) | |
| Path | Physical | Connected (5) | |
| | | Not Connected | |
| | Logical | Inbound (62) | |
| | | Outbound (134) | |

Regarding the delivery object, most groups are able to use techniques based on weaponized objects in order to gain initial access to their targets. Only one of them is able to gain access through not legit objects, in this case, through hardware additions. These data show the importance of protecting organizations against malware coupled with legitimate objects, as the techniques relying on weaponized objects are the most-used ones. This protection can be achieved through perimeter security elements such as web traffic inspection devices or sandboxes for mail attachments or links. While all delivery techniques are important, the probability of being compromised through Weaponized Objects is much higher than the probability of being compromised through a hardware addition; therefore, on a prior basis most organizations must allocate more resources to protect themselves against a malware compromise than to protect themselves against a hardware addition or a supply chain compromise. Finally, it is important to note that only six groups are able to execute different techniques based on undefined delivery objects, all of them through supply chain compromises. As we have stated in this paper, these compromises are not widely exploited, but supply chain attacks are a growing trend.

Regarding the delivery vector, Semantic delivery is the most widely used vector. This implies that most groups exploit social engineering techniques to gain access to the target infrastructure, a fact that highlights the importance of security awareness for most organizations. The use of untrusted parties to gain access to the target is also relevant. Threat actors rely on the trojanization of web sites visited by the users of the targeted organization to gain access to their victims. Although this technique can be considered a non-directed one (there is no guarantee about who is visiting the trojanized web site), advanced threat actors analyze their targets and identify the sites those targets will visit with high probability. This fact again highlights the special relevance of the security countermeasures that must be implemented in order to protect navigation traffic. Finally, the fact that no threat group is identified as being able to exploit Deliberate delivery, that is, that no threat group is exploiting insiders to gain access to a target, is especially relevant

Finally, regarding the delivery path, most initial access is gained through a logical delivery. As physical delivery is usually harder to exploit, most threat actors focus on logical paths to gain access to their targets. The protection of the logical perimeter, both of inbound and outbound connections, is highlighted once again. Finally, as we have previously stated, no threat group is identified to be able to exploit not connected delivery paths. As few countermeasures will be implemented to protect this kind of compromise, this fact leaves a potential window of opportunity for threat actors.

### 2.5.2 A Practical Example

In this section we provide a practical example for our proposed taxonomy and its usefulness. For this case study, we have chosen the supply chain compromise delivery technique. As we have previously stated, supply chain compromise is a growing trend in advanced hostile operations; thus, it is mandatory to design and implement security countermeasures to face this technique.

In order to protect an organization against supply chain techniques, the first step is to understand how they are implemented by threat actors. MITRE ATT&CK being the main framework for tactics and techniques, it is the first reference to analyze supply chain compromises. As we have stated, this framework provides three sub-techniques, two of them regarding software supply chain compromises and the last one regarding hardware supply chain compromises. With a main focus on software-related compromises, hardware supply chain attacks are in the background. This is consistent with an in-depth literature review, where most research is focusing nowadays on software supply chain attacks, that is, on the compromise of the software delivered to the target, both commercial [477] and open source [456]. On the other hand, hardware supply chain is, in general terms, less considered. The hardware supply chain security is analyzed and modeled in [254]. Hardware implants in supply chain attacks are analyzed in works such as [255], where the authors focus on the security of electronic devices, or [260], where Jacob Harrison et al. provide a review of Printed Circuit Boards' malicious hardware implants through the supply chain.

This initial literature review, whose aim is to identify the security countermeasures to be implemented in order to protect an organization, presents two main problems that can lead organizations to leave relevant security gaps that can be exploited by advanced threats. The first problem is related to the focus on software supply chain attacks. These compromises are the more common ones, as hardware compromises usually require much more effort, human and economic, and their scope is limited to a reduced group of targets. Different Advanced Persistent Threat groups, such as APT29 or Sandworm, are able to compromise the software supply chain, while the groups with hardware compromise capabilities are less common. However, hardware supply chain compromises are still a relevant threat to organizations; thus, they must be conveniently considered when designing and implementing a protection plan.

The second problem is related to the hardware versus software approach. This distinction is too simple in some cases and it usually refers to the way an artifact is delivered to the target, without considering other elements. In our proposed taxonomy, it focuses only on the delivery path. By adopting this simple hardware or software consideration, without evaluating other elements of the compromise, relevant protection gaps can be left by a security team.

To face these common problems, our proposed taxonomy helps defensive teams to cover security gaps by considering uncommon delivery approaches. In addition, our taxonomy also allows the identification of novel techniques, in order to evaluate their likelihood and impact and, if applicable, to establish countermeasures to face them. Regarding supply chain compromises as a delivery technique, our taxonomy considers not only the delivery path point of view, but also the vector and the object. If we map supply chain related techniques onto our proposed taxonomy, both the ones defined in MITRE ATT&CK and unusual techniques identified in the literature review, we can find gaps that can lead to security breaches through a supply chain compromise.

In relation to the Path, as we have stated, supply chain techniques are mostly linked to software and hardware delivery, not considering situations such as:

- The mix of both approaches: the delivery of trojanized software through physical means, for example, through any removable media. This delivery technique is identified in our taxonomy, and it is a key threat to be considered in air-gapped networks, where software is installed or updated through physical media.
- The physical delivery of a trojanized element that is not directly connected

to the target infrastructure, for example, a device to remotely listen to a conversation.

If these supply chain delivery techniques are not considered in a threat model, defensive teams will not be able to identify and establish appropriate countermeasures. Through the dissection of the delivery technique and the identification of its components, our taxonomy includes all the relevant elements to identify these approaches to the delivery.

In relation to the delivery vector, all supply chain compromises are Indirect, by their own definition. The abuse of untrusted parties is not usually considered in the literature, as major well-known supply chain attacks, such as SolarWinds [40, 477] or Kaseya [31] are based on a trusted party to compromise their final target. Hostile actors took advantage of this gap in 2021, as most organizations are not aware of these attacks. For example, the Lazarus group have developed techniques to compromise its targets through the trojanization of general-purpose software available on Internet [274]. This advanced threat actor is able to download the malicious software to the target infrastructure through social engineering techniques, thus successfully developing a supply chain compromise approach through untrusted parties. This approach, observed in 2021, would have been considered by applying our proposed taxonomy, as it identifies not only trusted Parties but also untrusted ones as indirect delivery vectors.

Finally, regarding the delivery object, none of the works we have analyzed focuse on the relevance of this element for supply chain delivery techniques. In fact, the MITRE ATT&CK framework does not provide enough detail for their exposed supply chain techniques to map them onto relevant categories of our taxonomy. This lack of analysis can result in security gaps for an organization. All supply chain compromises are based on the manipulation of the delivered product at any stage of the chain. However, the type of manipulation is relevant for its detection. For example, a trojanization through dynamic objects can be detected by malware analysis, while a trojanization through a static object would not be detected by that analysis, and would require a hardening check. If the object is not considered, organizations cannot identify and implement appropriate security countermeasures to detect supply chain compromises. Our taxonomy, in this case the delivery object, allows analysts to identify different types of manipulation and to deploy the relevant countermeasures in each case to face them.

In this practical example, we have applied our proposed taxonomy to the identification of novel and uncommon supply chain delivery techniques. We have shown how our work helps analysts to establish a model for delivery techniques that allows them to anticipate in hostile operations through the analysis of the delivery object, vector, and path. In this way, cyber security levels can be increased and global protection for an organization is enhanced. Although we have focused on a supply chain compromise, our findings and proposed taxonomy can be applied to all kind of delivery techniques used by advanced threat actors.

## 2.6 Discussion

In order to provide an initial taxonomy for delivery techniques, we have analyzed the different existing approaches to accomplish the Initial Access tactic performed by advanced threat actors. In this analysis, one major finding is that voluntary delivery is not considered among the analyzed techniques. A deeper classification for deliberate actions is processed in different works regarding the insider threat [73, 129, 225, 281, 494, 661], although such a fine structure is outside the scope of this paper. Being a MITRE ATT&CK, the commonly accepted framework for tactics, techniques, and procedures, we consider that it should identify such approaches for a deliberate initial access, thus providing mechanisms to mitigate these techniques.

A major finding during our research is that different well-known delivery techniques are not considered in current frameworks. This gap between currently used delivery techniques and those identified in key references, such as the MITRE ATT&CK, which is a primary source for security analysts, may lead to an opportunity window for threat actors, as defensive teams are not considering some hostile approaches; thus, they are not implementing countermeasures against them. In this sense, we consider our taxonomy to help analysts identify those lacking techniques that can be executed by threat actors, thus helping organizations to identify appropriate countermeasures against them.

In addition, also regarding this framework, we have identified an important lack of formal structure in MITRE ATT&CK techniques for different particular tactics. This framework being the main effort and the de facto standard to identify and analyze tactics and techniques from advanced threat actors, we consider that it should define a taxonomy, or at least a classification, for the identified techniques in each case, thus providing analysts more concise information about them and increasing defensive capabilities to prevent, detect, and neutralize threats. This work will be shared with MITRE in order to be considered to enhance the ATT&CK framework.

As identified research lines, we propose, in the first place, a deeper analysis of supply chain compromises as a key delivery vector. As we have stated before, supply chain attacks are not the most widely used technique, but their popularity among threat actors is increasing and their impact is high. We have provided a practical example to improve the identification and detection of supply chain compromises through our proposed taxonomy, but of course more research has to be performed. Ref. [405, 516] provide a framework and catalog of supply chain attack patterns identifying objects of the compromise, types, time frames, and, as a key element, points of attack within the supply chain. Ref. [677] proposes a threat model for supply chain attacks, and ref. [679] provides a threat analysis for these attacks. Much work has been performed in this line, but most of it is focused on the analysis of particular approaches, lacking a formalization and with an abstract, technology-agnostic structure for their classification. This kind of analysis should improve an organization security by providing a deeper knowledge about supply-chain-based initial access, thus enabling appropriate countermeasures against these kinds of hostile actions, so we miss a deeper work from this perspective.

As we have stated, another research line we identify is the analysis of the deliberate human factor in offensive operations; when dealing with advanced actors, they have capabilities to employ not only technological approaches for an initial access or for other tactics, but also to use insiders to achieve their goals. and a relevant challenge is to determine if an action delivered by an insider is a justifiable threat [553]. The detection of these internal hostile actors is mandatory in order to provide an adequate level of protection. Although in classical security the insider threat has been well studied over the years, we consider that these analysis must be carried to cyber operations, following a kill-chain approach [49, 371], where the mix of people and technology can lead to a high impact for an organization.

Finally, we identify a third research line related to the application of computational intelligence to the identification and classification of delivery techniques. Our taxonomy can be used as a general classification scheme and can help analysts to select relevant features to model delivery techniques. With these features, main computational intelligence approaches can be exploited to identify and classify techniques. Fuzzy logic systems or neural networks have been successfully applied in many different fields, such as fault detection [28, 502, 651] or smart cities [291, 459]. Related to delivery techniques, the role of fuzzy logic in supply chain management [26, 652], resilience [98, 344], or risk assessment [178, 216] are relevant research fields.

## 2.7 Conclusions

Our work provides an initial taxonomy for delivery techniques in order to better understand the global tactic and to protect organizations against threat actors achieving initial access into a target. Delivery is a key tactic for advanced threat operations; part of the global success of an offensive operation relies on the correct achievement of the delivery, as it is the first stage of the operation in which the threat actor interacts with its target in a hostile way. For this reason, an accurate classification and structure for the techniques linked to the delivery tactic is a must in order to identify capabilities, to profile advanced actors, and to develop, implement, and maintain security countermeasures against them. However, we have identified an absence of a suitable classification scheme for the techniques related to the delivery stage in hostile cyber operations by advanced threat actors. Even MITRE ATT&CK, the key reference in the subject, lacks a suitable approach to classify the different techniques inside the delivery tactic. As this tactic is mandatory in all kind of offensive operations, we consider it especially relevant to establish a suitable taxonomy for it, thus helping organizations to better understand this stage and enhancing their prevention, detection, and neutralization capabilities.

In this work, we have delved into how the delivery tactic is achieved. To establish such a taxonomy, we have dissected the tactic as a way to model it. We identify the different elements that define the delivery and we deploy them into a classification that provides this taxonomy. As stated before, as the MITRE ATT&CK is considered the key reference for tactics and techniques used by threat actors, we have aligned our approach with this framework and classified all the identified techniques into our proposed taxonomy. We consider our work to significantly contribute to improving, not only the threat model for hostile advanced actors, but also the detail organizations must consider for a suitable protection against them, in this case against the delivery tactic.

Tactics and techniques are one of the first key points to model advanced threat actors and to deploy capabilities in order to prevent, to detect, and to neutralize them. We consider our proposal as a starting point towards a commonly accepted taxonomy that helps research to better understand hostile actors, especially advanced ones. In future works, our proposal can be fine tuned to provide a more accurate approach to the Initial Access performed by advanced threat actors; we consider that all improvements must be aligned with industry standards, such as MITRE ATT&CK, in order to be useful to the security community. In addition, in this paper we have identified key research lines regarding delivery techniques whose relevance is growing or, is direct, which is not considered in main frameworks.

In this sense, we miss in main frameworks, such as MITRE ATT&CK, a deeper analysis of delivery techniques in three directions: supply chain attacks, as growing threats; deliberate delivery techniques, especially those regarding an insider; and physical perimeter breaking, with no further connection to the target infrastructure.

# Chapter 3

# A Taxonomy for Threat Actors' Persistence Techniques

*This chapter dissects the persistence of Advanced Threat Actors in their operations and proposes a novel taxonomy for persistence techniques. A novel concept is introduced and discussed: persistence point, the location within the compromised infrastructure where a persistence artifact is stored. This approach improves not only the modeling of advanced threat actors but, which is more important, it provides organizations a platform–agnostic model for the detection of persistence techniques. As the rest of this work, this proposal is aligned with MITRE ATT&CK, the main framework for the identification of threat actors' tactics and techniques.*

This chapter was published as: Villalón–Huerta, A.;Marco–Gisbert, H.;Ripoll-Ripoll, I. A Taxonomy for Threat Actors' Persistence Techniques. Computers & Security. Elsevier, 2022. https://doi.org/10.1016/j.cose.2022.102855

## Contents

The main contribution of this paper is to provide an accurate taxonomy for Persistence techniques, which allows the detection of novel techniques and the identification of appropriate countermeasures. Persistence is a key tactic for advanced offensive cyber operations. The techniques that achieve persistence have been largely analyzed in particular environments, but there is no suitable platform–agnostic model to structure persistence techniques. This lack causes a serious problem in the modeling of activities of advanced threat actors, hindering both their detection and the implementation of countermeasures against their activities. In this paper we analyze previous work in this field and propose a novel taxonomy for persistence techniques based on persistence points, a key concept we introduce in our work as the basis for the proposed taxonomy. Our work will help analysts to identify, classify and detect compromises, significantly reducing the amount of effort needed for these tasks. It follows a logical structure that can be easy to expand and adapt, and it can be directly used in commonly accepted industry standards such as MITRE ATT&CK.

## 3.1 Introduction

Persistence refers to the capability of a malware to survive to changes and interrupts, including reboots, in a compromised system. It is a widely used term in malware research, so most of its definition are malware–related; for example, Michael Sikorski et al. [574] define persistence as a behaviour of malware by which it tries to be in a compromised system for a long time. A more accurate definition is provided by Zane Gittins et al. [231], where the authors define persistence as a method by which malware survives a reboot of the victim operating system. Persistence is an important attribute that malware writers consider in its design; the reason is simple [660]: the longer the malware can stay in the victim's device, the more value it generates for the adversary.

However, persistence is not malware–specific: advanced threat actors do not use malware in all operations, but they also try to maintain their persistence in a targeted victim. In fact, persistence is a key tactic for these actors; when an Advanced Persistent Threat (APT) compromises a victim, one of the first steps that it will execute is to guarantee its foothold on the targeted infrastructure, maintaining the compromise upon system reboots.

Considering persistence as a key goal, persistence techniques are those that allow a threat actor to achieve it; they have been largely analyzed, but we have not found a clear structure for them, being most of the approaches architecture and operating system (OS) dependent. This fact hinders the identification of compromises out of the analyzed technologies. This paper provides a platform–agnostic taxonomy for persistence techniques, suitable for their identification and characterization. This taxonomy is based on a concept we introduce in our work, the persistence point: the location within the system where a persistence artifact is stored.

The contributions of this paper are the following ones:

- To provide an OS–independent taxonomy of persistence points, thus allowing analysts to identify persistence techniques in all platforms.
- To ease not only the detection of persistent artifacts but also the identification of the capabilities of threat actors.
- To identify the most used persistence points, then determining the most exploited system capabilities to achieve persistence and establishing probabilities in order to prioritize investigations, where applicable.

The rest of the paper is organized as follows. The background, section 3.2, provides an introduction to cyber operations tactics and techniques. In section 3.3 we assess the problem of the lack of a suitable structure for persistence techniques to help analysts in their investigations. Section 3.4 analyzes the prior work to identify persistence approaches, and in section 3.5 we propose a novel taxonomy for persistence points, as a direct way to structure techniques and to identify a compromised system. In section 3.6 we discuss the results of our work, comparing them with previous approaches and identifying improvements where applicable, as well as future research lines. Finally, section 3.7 summarizes the outcome of the overall work.

## 3.2 Background

From an abstract point of view, Robert Axelrod et al. define persistence of a resource [46] as the probability that if you refrain from using it now, it will still be usable in the next time period. When specifically dealing with malware, persistence is defined [330] as a process by which malware ensures continual execution on a system, independent of low-level system events such as shutdowns and reboots.

Persistent malware is the one [647] that makes permanent changes in memory and permanently changes the control flow within a system such that it can continue to achieve its objective. This feature allows the malware to be aware of and to react to changes in the compromised systems. Without persistence, malware is severely limited, and consequently, its impact is also limited; for this reason most threat actors, particularly advanced ones, will establish persistence in their victims in order to achieve an extended temporal effect for their operations.

In the context of advanced threat actors' operations, persistence is the tactic by which adversaries try to maintain their foothold on a targeted infrastructure; this tactic comprises different techniques that include any access, action, or configuration changes that allow the attacker to achieve the tactic, such as replacing or hijacking legitimate code or adding startup code. As we can see, this concept of persistence as a tactic exceeds all technological –including malware– considerations: persistence is defined as a key tactic for threat actors, regardless of malware.

To accomplish its goals, an advanced threat actor must guarantee its presence in a compromised infrastructure: that is, in one or more compromised systems.

## 3.3 Problem statement

Persistence is a key tactic for advanced threat actors and even for simple malware: the ability to control a compromised target over time, while remaining unnoticed, is an essential to guarantee the success of an offensive operation. However, when analyzing the capabilities that these threat actors develop in their operations, all research is focused on particular approaches related to specific technologies. In this sense, Microsoft Windows environments, as the most abused platform, have been largely analyzed in order to identify the locations where the artifacts that grant persistence can be located.

An analysis focused on specific platforms, or even on specific malware, has obvious limitations, as all the identification of persistence techniques are architecture and OS–dependent. The most relevant limitation is related to the identification of techniques in technologies outside of the scope of those specific platforms. This means that when faced with the compromise of a new platform and trying to identify if an attacker has enabled persistence on it, an analyst has only vague references for investigation. In this case, analysts do not have a common reference to identify the mechanisms that enable persistence, nor to identify the locations of the system where an intruder can achieve persistence. For an agile identification of persistence techniques in new environments, analysts need a common platform–agnostic reference that allows them to quickly determine which elements have to be examined and what to look for inside them.

In addition to the dependence on specific platforms, much research work about persistence focuses on specific malware and its capabilities to survive a system reboot. These approaches are useful for the analysis of particular malicious code, but they lack a global vision of the persistence. As we have stated before, persistence is a key tactic for an advanced threat, not a simple malware capability. In fact, different APT actors do not rely on malware to conduct their operations, the so called malwareless ones, while others do not use malware to achieve some specific tactics. Please note that although in computer virology the definition of malware is an open problem [340], when we refer to malware we refer to any code added, changed, or removed from a software system in order to intentionally cause harm or subvert the intended function of the system [396] [433] [410]. Attending to this definition, we do not consider elements such as legitimate system tools or remote credentials for legitimate services as malware, although they can be abused to maintain persistence.

When facing malwareless operations, detection becomes more complex. Atomic and computed indicators of compromise, such as hashes, do not provide the full capabilities for an accurate detection, so analysts must usually deal with behavioral indicators: for example, those that allow the contextualization of a legitimate system tool execution in order to classify it as anomalous. Due to their stealthiness, malware free approaches are interesting for advanced threat actors in different tactics of an operation, from intrusion [689] to data exfiltration [688] or lateral movement [627]. Regarding persistence, different techniques exploited by threat actors such as SANDWORM [582] or APT29 [432] do not rely on malware but on legitimate services and tools. Even some actors such as ALLANITE are able to conduct whole malwareless operations [583]. In this way, we argue that persistence is not only malware–related, but it is a tactic of advanced threat actors in offensive cyber operations that can be achieved through malware or through simple abuse of legitimate resources.

No suitable approximation for persistence techniques that would allow analysts the identification of compromised systems has been defined. In addition, no research has been performed regarding where persistence is stored in a targeted system: that is, which locations of a system should be analyzed in order to detect the compromise. Identifying where persistence is stored is a must when dealing with operations performed by advanced threat actors. Tactics are mandatory to identify what to look for, but detecting where the attacker has stored an artifact for persistence is mandatory to identify where to look for techniques that implement the persistence tactic. Without a common platform–agnostic reference for persistence techniques, the detection of persistence artifacts in a targeted system follows an unstructured approach that delays the defensive capabilities and opens a window of opportunity for hostile actors.

## 3.4 Approaches and limitations

MITRE ATT&CK (Adversarial Tactics, Techniques, and Common Knowledge) is a globally accessible knowledge base of adversary tactics and techniques based on real–world observations. This framework represents the biggest effort to identify tactics and techniques for advanced actors, and it presents a plain structure for Persistence techniques, which makes it difficult to establish a clear taxonomy or categorization for them. Some efforts have been made to align technical persistence capabilities with MITRE ATT&CK. [462] provides a link between Microsoft Windows malware and the framework by identifying the malware persistence capabilities. These relationships have also been analyzed in [231] and [356], where the authors map different persistence approaches used in Microsoft Windows malware samples to MITRE ATT&CK techniques. However, in the case of the Persistence tactic, this framework is too close to the most technical aspects of techniques, so it is not suitable for abstraction and for a modeling of techniques in a platform–agnostic way. Other malware–related frameworks from MITRE, such as MAEC (Malware Attribute Enumeration and Characterization) or CME (Common Malware Enumeration) do not provide even a categorization for persistence techniques. SHIELD, an active defense knowledge base that MITRE is developing to capture and organize what they are learning about active defense and adversary engagement, identifies defenses against the techniques stated in ATT&CK, but it does not provide a valid taxonomy for them.

Arushi Sharma et al. [561] characterize malware persistence into three categories: run–time, re–boot and trojanized system binaries; although this can be considered an interesting approach to consider behaviors of the malware in order to design a Fuzzy Inference System, it does not capture all the elements to build a taxonomy model.

Based on their basic approach to persistence, Matthew Webb [659] classifies techniques in five main categories: User Login Execution, System Startup Execution, Dynamic Linked Library (DLL) Injection, Execution Hijacking and Adversary Backdoors. This proposal provides some key aspects to a OS–independent approach, but as it is mainly focused on Microsoft Windows aspects, it still lacks a complete platform–agnostic view.

Jennifer Mankin [381] breaks persistence capabilities into three phases: installation, system boot and service load. Although this is an interesting approach for the modeling and detection of persistence, the author does not propose a classification for persistence mechanisms, but a common breakdown structure for all of them, with a focus on Windows services as a particular technique.

Most of the current approaches to the analysis of persistence techniques have been made to identify OS–dependent persistence methods. As we have stated before, the most analyzed persistence mechanisms are related to Microsoft Windows environments. Windows Auto–Start Extensibility Points (ASEP) were first introduced by Yi–Min Wang et al. in [653]; they were defined as the subset of operating system and application extensibility points that can be "hooked" to enable auto–starting of programs without explicit user invocation. ASEP are a key concept for malware persistence, as they define points that an attacker can abuse to maintain its foothold on a targeted system. In [623] Daniel Uroz et al. propose a taxonomy for Windows ASEP divided into four categories: system persistence mechanisms, program loader abuse, application abuse, and system behavior abuse; each of them is analyzed and its characteristics are extracted, identifying families of persistence points as shown in figure 3.1. Although this is a valid classification, in addition to being focused on Microsoft Windows it only defines the main four previous categories, without sub categories for them, so it should be detailed in order to specify a more complete taxonomy. A key resource to identify persistence points in Windows environments is Microsoft Autoruns [535] [534]. It is a utility that has the most comprehensive knowledge of auto–starting locations of any startup monitor, showing what programs are configured to run during system boot up or login, and when various built–in Windows applications are started.

The technical persistence capabilities that Microsoft Windows provides have been analyzed in different works. [413], [574] and [468] identify particular persistence mechanisms and locations for this operating system, while [412] provides a basic structure for persistence capabilities: startup shell directories, registry RUN, services, file infection, DLL hijacking, Winlogon and Task Scheduler. These approaches provide different proposals for the identification of persistence techniques in Microsoft Windows. However, none of these approaches, all of them focused on a particular environment, provides a platform–agnostic proposal suitable to be used in other technologies.

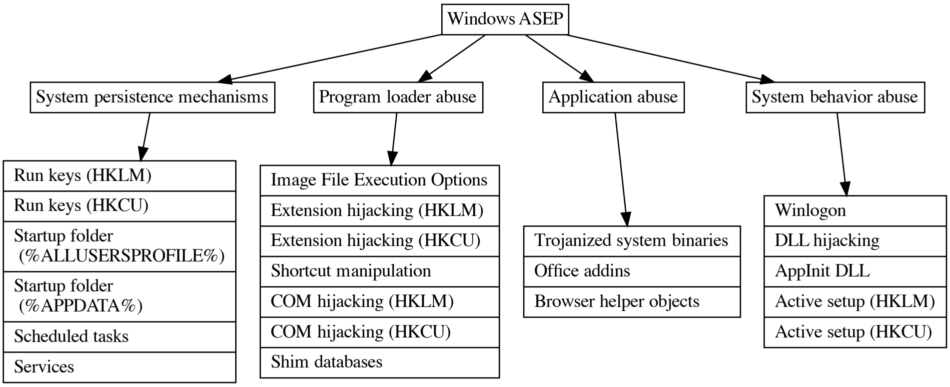

Figure 3.1: Windows Auto–Start Extensibility Points

Regarding other operating systems, in [286] Jun–ho Hwang et al. identify four methods for ELF malware persistence in Linux systems: subsystems initialization, time–based execution, file infection and replacement and user files alteration. Mike O'Leary [468] analyzes both Windows and Linux persistence mechanisms with practical examples, but without establishing a common framework for the classification of the identified techniques. [656] analyzes malware persistence mechanisms in Mac OS X, as well as the particular techniques used by different malware samples in this operating system. Also regarding Mac OS X, [656] provides an initial approach to technical capabilities and analyzes several malware samples and their persistence techniques. Again, as in Windows environments, none of these approaches delves into the definition of a general proposal for persistence techniques, considering only the identification of particular techniques for specific platforms.

Another key research line for persistence techniques is IoT–focused malware, as connected devices are growing in number and critical infrastructures are a clear target for hostile actors. Linux IoT malware persistence capabilities are analyzed in [86], where Calvin Brierley et al. identify, without providing a model or structure, six methods for persistence: modifying writable file systems, recreating read–only file systems, initrd/initramfs modification, "set writable flag" kernel module, update process exploitation and ubootkit. The authors defend that no universal method to gain persistence on IoT devices has been identified. In [102] Andrei Bytes et al. analyze techniques used in Programmable Logic Controllers (PLC), extending Linux generic mechanisms to particular embedded Linux devices but without providing a suitable structure for them. [440] analyzes rootkit persistence techniques in IoT devices, identifying Linux Kernel Modules, ramdisk–based and user space programs as main categories. Inside the last of them, the authors classify techniques into the following classes: service managers, job schedulers and user–specific startup files. Related to smart grid environments, in [189] Peter Eder–Neuhauser et al. identify persistence methods such as manipulation or anti malware tools, code obfuscation or encryption techniques. The authors analyze malware samples using each of the identified persistence mechanisms, but they do not provide a general structured classification suitable for its use outside smart grid environments.

In addition to different operating systems technologies, the persistence capabilities of malware families are also a key research focus. One of the most analyzed families is ransomware, due to the growing impact that this malicious software is causing in all kind of organizations. Yassine Lemmou et al. analyze [359] different ransomware samples and their persistence mechanisms, among other behaviors. Regarding banking malware, which is also a relevant issue, Paul Black et al. provide [74] an identification of the mechanisms for persistence in this kind of malicious programs: specific registry keys, such as registry or AppInit_DLLs, program trojanization, and bootkits.

Persistence capabilities from specific malware samples have been also analyzed in different works. [231] analyzes several samples to identify their persistence capabilities. In [490] the authors analyze WannaCry and Petya capabilities without focusing on an in–depth analysis, identifying the persistence capabilities of both malware samples. [17] and [318] provide an in–depth analysis of WannaCry, and [655] analyzes iWorm persistence, an OS X backdoor. Although these works analyze the persistence techniques of main malware samples or families, none of them provides a suitable structure for the abstraction of these techniques.

Malware that does not rely on the file system to run is referred to as fileless malware, and its persistence capabilities have been analyzed in different works. In [345] Sushil Kumar et al. classify fileless malware in three different categories in relation to its persistence techniques: memory–resident malware, Windows registry malware and rootkit fileless malware. In [544] the authors identify Windows registry, WMI store, SQL tables or Scheduled tasks as usual locations to achieve fileless persistence. In [647] Sebastian Vogl et al. analyze persistent data–only malware and discuss its challenges, presenting a proof of concept of this kind of malware and providing additional techniques to achieve persistence. Data only malware is malware that introduces specially crafted data into a system with the intent of manipulating the control flow without changing or introducing new code. [506] provides an initial classification for rootkits, differentiating between kernel–level and user–level rootkits, being the first ones the most analyzed in literature. In [310] Jestin Joy et al. expand this classification, identifying three types of rootkits: virtualization, kernel–level and library–level, being this last one a user–level rootkit. Nevertheless, none of the related work on malware persistence provides a taxonomy for the persistence techniques they explore, focusing once again only on the particular features of samples, families or malware persistence approaches.

## 3.5 Our proposal

After the analysis of the current approaches to define  a valid classification for persistence techniques, their strengths and weaknesses and, especially, their main lacks, we propose a platform–independent persistence taxonomy. This taxonomy will allow analysts to classify techniques, regardless of the technology used in each case, as well as to identify new persistence capabilities that threat actors might develop, thus being able to identify and implement counter measures against them.

To establish such a taxonomy we have to formally define the following concepts:

- Persistence point. The location within a compromised system where a persistence artifact is stored.
- Persistence technique. The actions that enable a persistence mechanism into a target, relying on a persistence point.

Please note that in this context, the term “artifact” refers not only to malicious software implanted in the targeted system, but to any software or configuration abused or manipulated by a threat actor in order to gain persistence. A persistence point is, as its definition highlights, a location; this location can be a hardware component, a file system point, a Windows registry entry, etc. As an example, in previous sections we have referred to fileless malware; regarding persistence capabilities, this could be considered a first classification for persistence points, although it is too general, as it comprises many types of persistence points with little relation between them.

Persistence techniques directly rely on persistence points to achieve their goal; from an analysts’ perspective, a persistence point defines where to look for persistence, while a persistence technique defines what to look for. No technique can be achieved without a persistence point. In this context, we can understand a persistence technique just as the abuse or manipulation of a persistence point. We refer to abuse when a threat actor does not modify the persistence point, but just exploits one or more of its standard capabilities. We refer to manipulation when the attacker alters the persistence point for his own benefit, by fabricating spurious data or by modifying or canceling legitimate data. For example, techniques relying on system accounts as a persistence point include those related to credential abuse, those related to the modification of legitimate users’ credentials and those related to the addition of non legitimate users to the system. The relationship between persistence techniques and persistence points is direct: all persistence techniques rely on at least one persistence point, and we can detect the artifacts that achieve persistence by inspecting those points.

In this work we establish a taxonomy for persistence points, which directly defines a taxonomy for persistence techniques. We propose four upper level categories:

- Pre–OS persistence points, those regarding hardware, firmware or initial sequences of a system boot, before a particular operating system is loaded.

- OS persistence points, those related to the boot of a particular operating system and to its native capabilities.
- Server–software persistence points, those related to remotely accessible software that is provided to users without a full access to the system.
- User persistence points, those related to particular user activities or configurations.

In the following sections we discuss each of the proposed categories for persistence points and we specify their particular structure, thus defining a taxonomy for persistence techniques.

### 3.5.1 Pre–OS persistence

The boot process for a computer system fully depends on the specific hardware and on the OS that will be running on it. Most research is focused on the boot process for BIOS or UEFI based x86 systems running Microsoft Windows OS flavours. In this particular case, the boot process can be described in a simple way as follows:

1. On startup, firmware (BIOS or UEFI) is called and transferred with execution.
2. This firmware initializes hardware devices and goes through storage devices to look for a bootable one.
3. When a bootable device is found, boot code is executed, loading the boot sector for a bootable partition.
4. Boot sector loads and executes the operating system boot loader.
5. Boot loader loads the OS kernel from the storage device.

In the case of UEFI systems, this firmware directly loads the operating system boot loader without relying on boot sectors, thus skipping step 3; in these systems, this boot loader is just an EFI file in the filesystem.

While a rootkit is malicious software that impacts its target at user and kernel levels [511], a bootkit is a particular rootkit that transfers its storage location from the file system to the hardware and activates itself while or even before the operating system kernel is loaded [363]. Bootkits are largely analyzed in [386]; as this malicious software is loaded before the operating system, it can tamper the whole computer system. According to the different stages in boot process, [363] classifies bootkit technologies into four categories: BIOS–based bootkit, MBR–based bootkit, NTLDRbased bootkit, and others. A similar approach is followed in [218], where Hongbo Gao et al. expose the same classification without the "others" category. As BIOS–based bootkits write their malicious payload directly into the BIOS, they typically have to target particular BIOS or hardware, so they are rarely seen in the wild [246]. MBR and NTLDR–based bootkits are the dominant types in real–world malware.

These works provide a key approach for an initial taxonomy of Pre–OS persistence points: those that are OS–independent, BIOS or MBR, and those that are based on the initial steps of a particular OS being loaded, such as NTLDR and others. However, this approach lacks a whole family of persistence points that we must consider: those related to hardware implants. In fact, hardware implants are not usually considered in persistence techniques families, such as those presented in MITRE ATT&CK.

To provide a platform–agnostic approach, we divide Pre–OS persistence points into three main families: hardware, firmware and software related, as shown in figure 3.2.

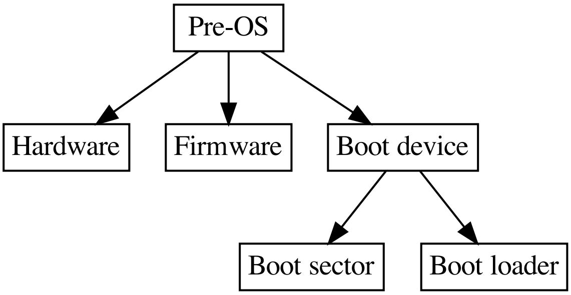

Figure 3.2: Pre–OS persistence points

Hardware persistence points are those related to hardware implants to maintain persistence on a specific targeted system. At the early stages of a system boot, this hardware and its linked firmware are initialized by a firmware component, just as UEFI or BIOS. These implants have few academic research, especially focusing on covert channels for air–gapped networks [252] [650] [418]. In [353] Austin Lasota provides a very brief introduction to different types of hardware implants in Apple's Mac hardware and their countermeasures. Until now, the most extensive information about these implants continues to be NSA's ANT catalog, which has been publicly exposed and which is detailed in works such as [115].

In addition to hardware, UEFI, BIOS or equivalent firmware is another key persistence point. In [385] Alex Matrosov proposes a classification of vulnerabilities and attack vectors for BIOS persistent infection, identifying persistent and non–persistent implants for UEFI firmware through post exploitation and supply chain vulnerabilities. Note that when referring to hardware and firmware persistence in our taxonomy we are considering not only the main system components, but also the peripheral devices, especially those that have direct memory access (DMA) to the main system runtime memory. In fact, DMA attacks have been largely analyzed, from a pure malware perspective in [593], where Patrick Stewin et al. introduce the concept of DMA malware, to a general threat actor capability [85].

Finally, firmware components go through a boot sequence that depends on the type of firmware being used; here we find the third family of Pre–OS persistence points, those related to the boot device. In this case we differentiate two families of persistence points: boot sector and boot loader ones. Legacy systems (i.e., BIOS) use boot sectors (for example, Master Boot Record, or MBR, which loads a Volume Boot Record, or VBR). Boot sectors are to call a boot loader (in those systems with UEFI this boot loader is just an EFI file in the filesystem, as stated before), which loads the OS kernel, which is another persistence point but in this case OS–dependent, as described in next section.

We want to highlight that when dealing with virtual machines running over hypervisors, our proposed taxonomy is completely valid. In this particular case, pre–OS persistence points are ones exposed in this section, but considering two main aspects. The first one is that these persistence points can be identified in the hypervisor systems, not only in the virtual machines. The second one regards these virtual machines: being the same persistence points, on virtual machines they are stored in software that emulates hardware components.

### 3.5.2 OS native persistence

Once the OS starts to boot, a second category for our taxonomy is defined by those persistence techniques based on different OS capabilities. In this case, the obtained privileges can be those of the operating system or those of a particular user, as system boot is executed with administrative privileges. In figure 3.3 our proposed OS native persistence points taxonomy is shown.

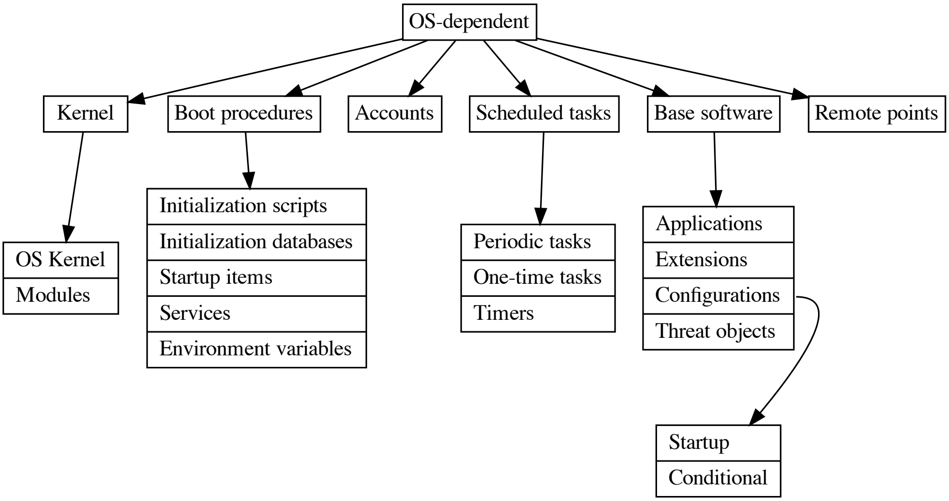

Figure 3.3: OS native persistence points

Within this category we find first of all the operating system kernel as a persistence point; when an OS boots, its kernel is loaded and a process is started to execute tasks such as software initialization or module loading. By subverting the operating system kernel, a kernel rootkit embeds itself into the compromised kernel and stealthily causes damage with full unrestricted access to the system's resources [525]. Kernel persistence can be achieved through the compromise of the OS kernel itself or through the compromise of kernel modules in any of their forms, such as kernel modules, kernel extensions or kernel drivers, loaded on startup or when certain conditions are met. Malware such as Drovorub exploits kernel persistence points, establishing persistence through kernel modules [202].

Once the kernel is loaded and the basic capabilities of the operating system are started, boot procedures are another OS–dependent persistence point. In this stage of a system boot we find clock synchronization, daemons launched or subsystems initialized; in addition, please note that many operating systems can execute or load user files when booting. In all of these tasks, code is executed and configuration files are loaded, so a hostile actor can modify them to achieve OS–dependent persistence; in the particular case of remote objects access, such as in clock synchronization, the threat actor can also compromise remote systems in order to establish persistence on a targeted one, but here we must refer to remote persistence points, as stated in this section.

Inside the boot procedures family, in addition to these initialization scripts, many operating systems provide a database, such as Windows Registry or AIX Object Data Manager, where stored data is accessed during system boot to initialize specific OS capabilities, such as network settings or program launch. We must differentiate this family of persistence points from the previous one, as they are stored on a different location for the OS and they are accessed and managed in most cases with specific native tools. In fact, malware that enables persistence by adding a specific entry into these databases is considered fileless, as opposed to malware that stores an artifact in the filesystem. For example, Windows Registry is a widely used persistence point; we can find different threat actors exploiting this registry to achieve persistence, such as APT32 [163] or APT37 [210], as well as specific malware such as WannaCry [17] or Bisonal [266] [275].

In addition to initialization databases, different operating systems such as Microsoft Windows or Apple Mac OSX also provide a startup items location, usually a folder with references to programs (in many cases in the form of symbolic links) that are executed during system boot. By simply adding the correct reference to a program in this persistence point, the program will be launched on startup; jRAT is a known cross–platform backdoor exploiting this kind of persistence points [315]. Please note that in these operating systems there is usually a similar persistence point for each particular user, where references are executed when the user logs in.

As a fourth category inside the boot procedures family, we find services, also called daemons, managed by the OS and launched to perform specific tasks in the background. Linux malware often exploits these services persistence points [157] [468]. The tasks launched by services may include starting clients to connect to remote services, enabling operating system capabilities such as the execution of scheduled tasks or the launching of long time running processes that will be up as the operating system is running. Inside this category we find standard, mandatory OS services and daemons, those related to non operating system native applications: although they can also be started as services in many cases, from a technical point of view, we classify their persistence points inside the "Server Software" category. For example, the compromise of a Microsoft Exchange Server in order to establish persistence by a threat actor relies on Microsoft Exchange (a server software) as a persistence point, not on a native system service. This point is clear, as server software has its own configurations, in many cases including access accounts, outside the OS configuration files. While a system service can not be stopped or uninstalled without introducing some kind of system instability, a server software can be completely stopped interfering only with its own availability.

Finally, as a fifth family inside this category we must consider environment variables; these variables are set during system boot and they affect global execution of applications in the system. We differentiate them from the base software categories, as in this case the persistence point is not related to a particular application, but to the whole system. In addition, please note that many of these variables can be overwritten by user–related ones; in this case we do not consider them as a separate persistence point, as they are set in the login scripts for a particular user.

One of the easiest points for an attacker to gain persistence is through the abuse of accounts that grant access to targeted systems. This technique has been exploited by threat groups such as APT29 [136]. When dealing with persistence points, we must differentiate accounts related persistence points from login–related ones; while all of them trigger execution when a user logs in the system, account–related ones are stored in the system's user table, while those related to user's login are stored in the user's home directory. In the same way, we must also differentiate system accounts, those that grant access to the system, both as privileged and as unprivileged users, and application accounts, those that grant access to a specific application served within the system, mainly to external users: for example, a web, database or e–mail account. An attacker can abuse both of them, but persistence points are different: in one case, the persistence point is located in the system's user table, while in the other one the persistence point is the particular user database or equivalent regarding the targeted application. For this reason we consider application accounts as a category for persistence points within software compromise, not within the accounts category.

The abuse of scheduled tasks capabilities that any OS provides comprises persistence techniques exploited in the wild by threat actors such as APT3 [51] or specific malware such as Emotet [346], and even frameworks such as Cobalt Strike provide this capability [641]. Unix `at` or `cron` utilities, or Windows `at.exe` or Task Scheduler are commonly used by threat actors to maintain persistence on compromised systems. In this case, we can find three families of persistence points, as they are stored in different locations in the operating system: those related to the execution of periodic tasks, those related to the execution of one–time tasks and those related to timers.

Base software represents another family of OS–related persistence points in our taxonomy. In this case, persistence is stored in software that is not initialized during system boot (as it is in the "Services" branch) but in software that is natively provided by the OS, stored at disk and that is not mandatory for the OS to boot, but it is to run properly. In this category of persistence points we propose in our taxonomy four families: the own application and its extensions (that is, the binaries launched on the execution), the software configuration and threat objects.

The binaries used by the base software can be trojanized by a threat actor in order to execute malicious code, and this can be done both in the main software binary and in its extensions (for example, libraries, plugins or code loaded under certain circumstances). In this context, we define trojanization [381] as the process to hijack an executable object that already exists on the system, patching it with malicious code that will be executed when the previously–benign program is loaded to run. An example of a threat actor achieving persistence through this persistence point is Gelsemium, a cyber espionage group that drops a trojanized DLL to be loaded by the spoolsv Windows service [182].

Software configuration is also a persistence point abused by attackers; in this case we refer to the configuration loaded on startup and to the configuration loaded under certain events (for example, when a condition is met or when a software extension is executed), named conditional configuration in this work. Please note that startup and conditional configuration persistence points are differentiated because they may be stored in different locations and they can also have different syntax and even different formats.

Finally, the last main category inside the base software persistence points is the one related to threat objects for the software. These malicious objects are especially crafted to execute certain actions on the targeted system when they are accessed in any form (for example, loaded or executed) by the software; until that moment, no malicious activity is performed, being this access the trigger that maintains persistence. Webshells are well–known examples of threat objects: malicious files added to the web contents that, when accessed in some form, can grant access to the system. Webshells are exploited by many threat actors such as Deep Panda [611] or OilRig [349].

In our proposed taxonomy, we consider remote persistence points as the last family of locations that persistence techniques rely on. In this case, the persistence point is located on a remote system and persistence is triggered when the remote object is accessed. Although a remote persistence point can be found for almost all the previous categories, it is important to differentiate them as they are not stored on the targeted system, so this system has no hostile activity until the remote point is accessed. Remote persistence points can be found from the own boot process, for example in netbooted systems where the operating system boots from a network image, but also when dealing with accounts, remote services or even remote threat objects opened by a server software. We highlight that we consider important to differentiate them from local persistence points because they are stored outside the targeted system, so the techniques for their location and identification are different. Apart from laboratory proofs of concept, we have not identified specific real–world malware relying on remote persistence; although technically possible, we consider it is not exploited on the wild.

### 3.5.3 Server software persistence

Server software is another key persistence point. This category is related to software that, when running on a system, provides services to remote users; we must differentiate this family from the category of base software because server software is not mandatory for the operating system to run properly. With the exception of appliances or dedicated systems, server software is installed apart from the OS, so it can be uninstalled without affecting the OS native capabilities. Inside this category we can find software such as mail or web servers, VPN hubs or terminal servers.

In our taxonomy we propose five families of persistence points for server software persistence. We are identifying persistence points related to software, so four of them are the same as in base software, and also as the ones in user software: the own application and its extensions, the software configuration and the threat objects. The fifth of these families is related to the accounts that grant access to the software. In figure 3.4 this taxonomy is shown.

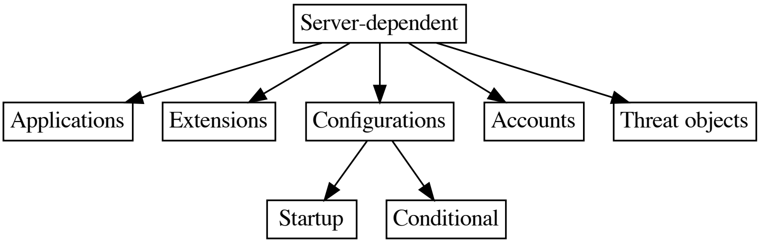

Figure 3.4: Server–dependent persistence points

We must focus on this fifth category for server software persistence points, the one related to Accounts. In OS base software or in user software we do not find a family for Accounts, as in these cases software is executed in a system by a previously authenticated entity. However, when dealing with server software, Accounts are a key persistence point, as they are remotely abused by attackers. In server software, persistence techniques include the abuse of legitimate accounts as well as their manipulation. Account abuse relies on valid software credentials that are used by a threat actor among time, granting direct access to the software and to the information. Known threat actors abusing accounts to gain persistence are APT28 [428], APT29 [221] or APT39 [265]. Account manipulation include the addition of accounts to be exploited by an attacker, as well as the modification of credentials that grant access to the software. Known threat actors performing account manipulation for persistence purposes include Sandworm [582].

### 3.5.4 User dependent persistence

Finally, a fourth category for persistence points, as for techniques, is the one based on specific user locations and actions. In these techniques, persistence is triggered after a user executes a particular action, and the persistence point is usually located in the home directory of the targeted user, with exception of remote persistence points, as we will describe later. The privileges of the hostile actor are those of the particular user that executed the action, being the capabilities to execute privileged commands restricted to the elevation of privileges through the exploitation of vulnerabilities. In figure 3.5 our proposed User–dependent persistence points taxonomy is shown. Please note that although an attacker can enable persistence relying on the user's scheduled tasks, this family of persistence points is considered OS–dependent, as it does not rely on a specific user action to be triggered, but on the complete boot of the OS.

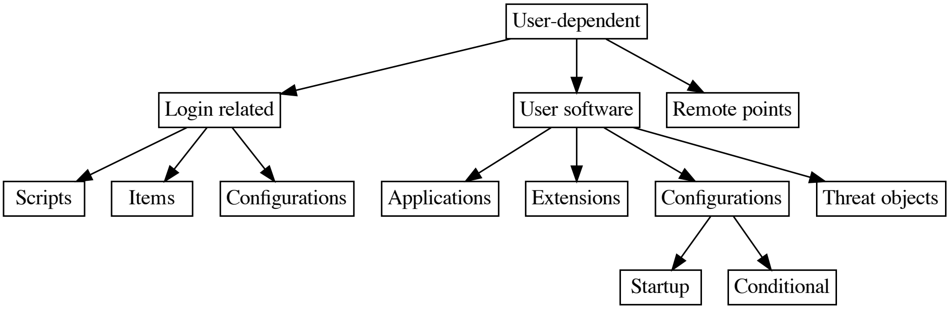

Figure 3.5: User–dependent persistence points

In first place, login–related techniques are those that group persistence points triggered when a user logs in a system. We must consider this category apart from the one exposed in the previous section, regarding accounts of a targeted systems, as the persistence point in this case is different and located in the user's own configuration, not in the system capabilities.

Inside the login–related persistence points we must differentiate between logon scripts, logon items and logon configurations. The first family includes those user–defined files that are executed when the user logs in the system, usually in the form of scripts. The second one refers, as when dealing when OS–booting related persistence, to the location where references to applications are stored to be automatically launched when a user logs in. Finally, the third family refers to specific software configurations that are loaded in the login process, not by user software but by server software that enables the login process. For example, if a threat actor achieves persistence by modifying a user SSH “authorized_keys” file, it is not altering a user software configuration, but a user configuration loaded by a server software. These persistence points could be considered as conditional configurations for server software, but we find it important to differentiate them, as they are located in a different place from global server software configurations, and they are also writable without privileged access to the system. When we are referring to login related persistence points, we must also consider logout as a persistence trigger. Most operating systems allow users to define scripts, items or configurations to be accessed not only while a user logs in, but also when a user logs out of the targeted system. APT28 is a threat actor which actively exploits Login related persistence points in Windows systems [106]; examples of malware which is also able to exploit these points include Attor [277] or Netwire [126].

Another main family of user–persistence techniques is the one based on user software. When we refer to user software we are dealing with applications that are executed by specific users, not by system capabilities, although in most cases, particularly in multi–user environments, the software itself and some of its extensions are not writable by a normal user, but only by a privileged one. For user software as a persistence point we propose a structure similar to the one regarding server software, with the exception of the "Accounts" category, as user software does not use this kind of persistence point. Following this approach, persistence points can be found in the own application executable, in its extensions, in its configuration parameters, both on startup or conditional, or in specific threat objects.

As in server software persistence points, the first category in user software ones is the own application or its extensions, regarding techniques consisting on their malicious manipulation in order to maintain persistence on a compromised system. Each time the compromised application is executed or its extensions are loaded by the user, the threat actor can execute malicious code on the target system. Naikon group is an example of an APT relying on these persistence points, as to maintain persistence it drops a malicious extension to be loaded by Microsoft Word at startup [121]. Examples of specific malware abusing these persistence points are Industroyer [582] [176], which trojanizes Windows Notepad to establish a backdoor persistence mechanism, or Kobalos, which replaces the SSH client with a trojanized version in order to steal credentials on compromised systems [361] [487]. In addition to the use of applications and their extensions as persistence points, a threat actor can rely on the particular configurations of these applications, both loaded at startup or under certain conditions. This approach is the same that we have stated in server software but, in this case, related to user software. For example, MuddyWater APT exploits Microsoft Office configurations to maintain persistence on a compromised target [613], while APT32 replaces Microsoft Outlook configuration files to implant a backdoor for persistence [163]. Finally, the last category inside user software persistence points, as in server software ones, is the one regarding malicious objects opened or loaded by an application. In this case, persistence is triggered when the threat object is accessed by the user. These persistence points are used regarding objects that are regularly accessed, even automatically loaded, by the user; in other case, persistence would be very weak for an advanced threat actor, as it would fully depend on the user manually opening a threat object. To our knowledge, this kind of threat objects without a guaranteed access are not commonly exploited, although it is technically possible. The only group we have identified relaying on these persistence points is Gamaredon, which inserts malicious macros into existing documents providing persistence when they are reopened [81].

Please note that threat objects, as in server software, can be both local (for example, malicious templates to be loaded) and remote; in fact, not only threat objects, but also software configurations, extensions or even login related persistence points can be both local and remote. In these cases, as we did in OS dependent taxonomy, we consider again the concept of remote persistence point, as they are located outside the targeted system, so persistence relies on a third party, also compromised, system.

### 3.5.5 Summary

In our proposal we present a novel taxonomy for persistence points, those that store artifacts that are abused to maintain persistence in a compromised system. This model provides a direct taxonomy for techniques exploited by threat actors. We propose four high level families for these points: those regarding locations prior to the OS boot, those regarding locations directly linked to OS capabilities, those regarding locations related to server software as an addendum for the system and, finally, those related to user–related locations of the system. Each of them is divided into different families to provide an accurate persistence points taxonomy.

The proposed taxonomy for persistence points we have developed in this work is shown in table 3.1, where we summarize the different persistence points families for all of the main stated categories.

Table 3.1: Persistence points proposed taxonomy.

| | | |
|---|---|---|
| Pre–OS | Hardware | |
| | Firmware | |
| | Boot device | Boot sector |
| | | Boot loader |
| OS–dependent | Kernel | OS kernel |
| | | Modules |
| | Boot procedures | Initialization scripts |
| | | Initialization databases |
| | | Startup items |
| | | Services |
| | | Environment variables |
| | Accounts | |
| | Scheduled tasks | Periodic tasks |
| | | One–time tasks |

| | | | |
|---|---|---|---|
| | | Timers | |
| | Base software | Applications | |
| | | Extensions | |
| | | Configurations | Startup |
| | | | Conditional |
| | | Threat objects | |
| | Remote points | | |
| Server software | Applications | | |
| | Extensions | | |
| | Configurations | Startup | |
| | | Conditional | |
| | Accounts | | |
| | Threat objects | | |
| User–dependent | Login related | Scripts | |
| | | Items | |
| | | Configurations | |
| | User software | Applications | |
| | | Extensions | |
| | | Configurations | Startup |
| | | | Conditional |
| | | Threat objects | |
| | Remote points | | |

## 3.6 Discussion

We have identified the absence of a suitable taxonomy for the techniques commonly used by advanced threat actors to achieve persistence. As with the rest of tactics MITRE ATT&CK defines, this framework provides a plain relationship for persistence techniques. Such a plain structure hinders the analysis and, most important, the detection, of these techniques. As persistence is mandatory for most advanced threat actor's operations, it is important to establish a suitable approach for persistence techniques that allows analysts to identify the persistence in a potentially compromised system.

Persistence is not only a key tactic for advanced threat actors, but also a key feature for malware. Persistence has been analyzed in three main research lines: Microsoft Windows techniques, specific malware capabilities and, during the last years, persistence in IoT–related infrastructures. None of these lines provides a suitable taxonomy for persistence techniques, but only weak classifications schemes for them, linked to specific operating systems or even malware capabilities. Without a global, platform–independent approach, a relevant problem for analysts is to identify persistence in environments that have not been previously explored.

In this paper we define the concept of persistence point as the location within a compromised system where a persistence artifact has been stored. Dealing with persistence as a global tactic for advanced threat actors, these locations are classified into a novel taxonomy for persistence points, thus establishing the locations of a system that have to be analyzed to identify persistence mechanisms. In this way, our approach provides a common reference for the identification of persistence techniques, as the relationship between persistence techniques and persistence points is direct. All persistence techniques rely on at least one persistence point. For analysts, by inspecting these points it is possible to detect the artifacts that achieve persistence, even when facing compromises in new environments or technologies.

To discuss the completeness and correctness of our work, we have mapped MITRE ATT&CK persistence techniques to our proposed taxonomy. This framework is the main public effort to establish a classification for tactics and techniques used by threat actors. As on May, 2022, MITRE ATT&CK "Persistence" tactic (last modified on 19th July 2019), identified as TA0003, consists of techniques that adversaries use to keep access to systems across restarts, changed credentials, and other interruptions that could cut off their access. MITRE ATT&CK provides no structure for techniques inside the "Persistence" tactic; the framework places all of these techniques at the same level, providing in some cases specific sub techniques. Although this approach is followed in all ATT&CK tactics and techniques, we advocate that it is important to provide a fine classification for all tactics, and in this case, for the "Persistence" one, by dividing its techniques at least in the first–level classification we provide in this work, based on persistence points.

The mapping of MITRE ATT&CK persistence techniques to the persistence points we propose in our work is shown in table 3.2.

Table 3.2: MITRE ATT&CK techniques mapping.

| | | |
|---|---|---|
| Pre–OS | Hardware | |
| | Firmware<br>T1543.001<br>T1543.002<br>T1543.004 | |
| | Boot device | Boot sector<br>T1543.003 |
| | | Boot loader<br>T1543.003 |

| | | |
|---|---|---|
| OS–dependent | Kernel | OS kernel<br>T1547.006 |
| | | Modules<br>T1547.006 |
| | Boot procedures | Initialization scripts<br>T1037.004 |
| | | Initialization databases<br>T1547.001<br>T1547.003<br>T1547.004<br>T1547.005<br>T1547.010<br>T1037.001<br>T1546.001<br>T1546.002<br>T1546.007<br>T1546.008<br>T1546.009<br>T1546.010<br>T1546.011<br>T1546.012<br>T1546.015<br>T1574.011<br>T1137.002 |
| | | Startup items<br>T1547.001<br>T1547.011<br>T1037.002<br>T1543.001<br>T1543.004 |
| | | Services<br>T1097<br>T1547.002<br>T1547.008<br>T1547.010<br>T1547.012<br>T1543.002<br>T1543.003<br>T1546.003<br>T1546.014<br>T1574.010<br>T1053.004 |

| | | |
|---|---|---|
| | Environment variables<br>T1574.001<br>T1574.004<br>T1574.006<br>T1574.007<br>T1574.012 | |
| Accounts<br>T1098.004<br>T1136.001<br>T1136.003<br>T1078.001<br>T1078.003<br>T1078.004 | | |
| Scheduled tasks | Periodic tasks<br>T1053.003<br>T1053.005 | |
| | One–time tasks<br>T1053.001<br>T1053.002 | |
| | Timers<br>T1053.006 | |
| T1205.001<br>Base software | Applications<br>T1546.005<br>T1546.006<br>T1546.008<br>T1574.005 | |
| | Extensions<br>T1574.001 | |
| | Configurations | Startup<br>T1546.011<br>T1546.013<br>T1574.002<br>T1574.006 |
| | | Conditional |
| | Threat objects<br>T1574.008<br>T1574.009 | |
| Remote points<br>T1037.003<br>T1136.002<br>T1574.001<br>T1543.005 | T1078.002 | |

| | | | |
|---|---|---|---|
| Server software | Applications<br>T1546.005<br>T1546.006 | | |
| | Extensions | | |
| | Configurations | Startup | |
| | | Conditional | |
| | Accounts<br>T1098.001<br>T1098.002<br>T1098.003<br>T1133 | | |
| | Threat objects<br>T1525<br>T1505.001<br>T1505.002<br>T1505.003 | | |
| User–dependent | Login related | Scripts<br>T1546.004 | |
| | | Items<br>T1547.001<br>T1547.007<br>T1547.009<br>T1547.011<br>T1543.001 | |
| | | Configurations<br>T1098.004 | |
| | User software | Applications<br>T1554<br>T1546.005<br>T1546.006 | |
| | | Extensions<br>T1176<br>T1137.001<br>T1137.006 | |
| | | Configurations | Startup<br>T1574.002<br>T1137.003<br>T1137.004<br>T1137.005 |
| | | | Conditional |
| | | Threat objects<br>T1574.008<br>T1574.009 | |

| | |
|---|---|
| | Remote points<br>`T1137.004` |

As we can confirm, all techniques and subtechniques identified by MITRE ATT&CK for the “Persistence” tactic can be mapped to our proposed taxonomy. Based on persistence points, our approach provides not only this full coverage, but also a platform–agnostic structure for all these techniques, improving the detail that MITRE ATT&CK defines. Our structure also considers techniques not identified or partially identified in the MITRE ATT&CK framework: for example, our proposal extends T1137, Office Application Startup, to generic applications startup, considering not only Microsoft Office but any application a user can execute to achieve persistence.

Analyzing this mapping, it draws our attention that Hardware is a persistence point in the second classification level without MITRE ATT&CK associated techniques. This fact highlights the absence of techniques relying on hardware implants as persistence points. As we have stated in this work, these implants are expensive and highly platform–dependent, so threat actors do not use them on the wild. The other persistence point in the second classification level without linked techniques is server software extensions and configurations. In this case, this fact highlights that when a hostile actor uses server software to achieve persistence, it mostly relies on accounts, threat objects and even the own application binary as persistence points. This is a normal finding, particularly with the accounts and threat objects persistence points, as for a threat actor it is usually easier to abuse or manipulate such objects than to rely on extensions, not used in all server software deployments, or configurations, in many cases customized for each particular deployment and thus harder to iterate among multiple victims.

Another key finding is that Pre–OS persistence points are the less exploited ones by hostile actors. As we have stated in our work, although these persistence points are the hardest ones to detect and eradicate, their exploitation is usually expensive and difficult to achieve. In front of this situation, we can confirm that the abuse and manipulation of persistence points linked to the operating system boot procedures comprises most of the techniques from the framework. Particularly, an initialization database such as Windows Registry is the main persistence point for all analyzed MITRE ATT&CK techniques. Our taxonomy also extends this specific platform–dependent approach to a generic family of Initialization databases for OS–dependent persistence points.

Our approach significantly improves the analysis of persistence linked to specific operating systems or technologies. Platform–dependent approaches are only useful for the particular environments they are designed for, but they are not able to establish a common reference to be used in all platforms. None of the previous approaches we have analyzed defines a platform–agnostic classification, although some of them try to establish such a classification for the particular environments they study. These particular approaches have provided the common basis for our novel proposal of a general taxonomy suitable for all platforms. Our novel platform–agnostic taxonomy is a useful tool for defenders to face the detection of persistence techniques, as well as for the planning and execution of offensive operations such as cyberspace exploitation or cyberspace attack. The application of our proposal to both of these perspectives improves an organizations' security.

From a practical point of view, persistence detection is usually a complex task when facing advanced threat actors. These actors exploit uncommon techniques in a broad range of platforms, from standard operating systems such as Microsoft Windows or Linux to closed appliances, and even legacy systems. When facing incidents, MITRE ATT&CK framework is the common starting point for defenders. However, this framework is technology–dependent, so it can not be exploited to face persistence in new environments, and even to hunt not previously identified techniques in common platforms. In this sense, our taxonomy provides a general model that can be used to structure knowledge about persistence points, thus allowing analysts the identification of novel, or at least uncommon, ones. As all techniques rely on at least one persistence point, the identification of these persistence points directly implies the discovery of persistence techniques.

Relying on a closed framework for persistence analysis, we are limited to the previously identified persistence techniques. For example, MITRE ATT&CK identifies persistence techniques linked to different user software compromise; they include techniques linked to specific user software such as Microsoft Office or web browsers. Without a platform–agnostic model, persistence detection is mostly limited to these specific applications. By using our taxonomy, analysts can not only deal with persistence points linked to Microsoft Office or web browsers, but they can extrapolate them to other user applications, thus being able to identify new persistence points. Through this identification, security analysts can determine all the possible locations where an artifact can be stored for persistence purposes: i.e., the analysts are identifying where to look for, or where to implant (in an offensive operation), persistence artifacts. This is especially relevant not only in common platforms or technologies, but also when facing the identification of persistence techniques in novel, not previously analyzed, environments, such as legacy systems or proprietary appliances.

Once persistence points have been identified, from a defensive perspective these points must be analyzed in order to find traces of abuse or manipulation. Please note that, as we have previously stated, a persistence point can store not only malware, but also a full range of configurations or legitimate tools to enable persistence through them. This analysis can be performed through intrusion detection techniques, out of the scope of our proposal, such as misuse or anomaly detection. In addition, from an offensive perspective, the identification of persistence points will help the red team to determine the different points where an artifact can be stored in a target. Those points will range from the well–known ones to the less used ones, this is, to the less monitored ones. On a prior basis, the exploitation of uncommon persistence points will increase the probability of success for an offensive cyberspace operation.

Finally, in this section we identify machine learning as an especially interesting research line. Machine learning approaches can be applied not only to identify intrusions, including persistence, against infrastructure, but also to classify them into a suitable taxonomy of persistence points. Different researches have been conducted to analyze malware capabilities with machine learning approaches [546] [435] [683] [52]. A summary of them can be found in [620] or [578]. Nevertheless, although some of these approaches establish a suitable classification for malware [229] [500], none has focused on the persistence mechanisms used in each case, neither in a potential classification for these mechanisms.

## 3.7 Conclusions

Persistence, the ability to keep presence in a targeted system for a long time, is a key tactic for the operations of advanced threat actors. These operations are expensive for an actor, so once a target has been compromised it is a common approach to keep control of this target as long as possible.

Persistence techniques have been largely analyzed in particular technologies, from specific operating systems to malware samples or ICS (Industrial Control Systems) environments. However, these analysis lack a global perspective, thus making it difficult to identify general capabilities that can be extrapolated from one environment to another. This is a relevant problem for security analysts when facing the potential compromise of new systems or technologies, as there is not a common reference to check for the identification of techniques in these environments.

In this work we provide a global platform–agnostic taxonomy for persistence techniques that allows the analysis of compromised systems regardless of their technology, thus easing the security analysts' work. This novel approach identifies persistence points as the locations within a system where persistence artifacts can be located. These points represent the components of a targeted infrastructure that are abused to maintain persistence, so they define the locations where analysts must check the presence of artifacts. The relationship between persistence points and persistence techniques is direct.

Our taxonomy is based on four main persistence points families, regarding those that are located before the OS boots, those that are located on OS–dependent capabilities, those that are located on server software and, finally, those that are user–dependent, mainly located in a user's particular files. All of these families are divided into different categories to specify where persistence artifacts can be located. As a first and novel taxonomy, our proposal can be used as a fundamental basis for new and more specific approaches.

# Chapter 4

# CNA tactics and techniques: a structure proposal

*In this chapter we delve into destructive cyberspace operations, known as Computer Network Attack. As they are less common, these operations are less analyzed than cyber espionage ones. However, they are a major threat for all infrastructures. We propose a taxonomy for tactics and techniques linked to Computer Network Attack, thus improving the capabilities for their detection and neutralization.*

This chapter was published as: Villalón-Huerta, A.; Ripoll-Ripoll, I.; Marco-Gisbert, H. CNA Tactics and Techniques: A Structure Proposal. Journal of Sensor and Actuator Networks. 2021, 10, 14. https://doi.org/10.3390/jsan10010014

## Contents

Destructive and control operations are today a major threat for cyber physical systems. These operations, known as Computer Network Attack (CNA), and usually linked to state-sponsored actors, are much less analyzed than Computer Network Exploitation activities (CNE), those related to intelligence gathering. While in CNE operations the main tactics and techniques are defined and well structured, in CNA there is a lack of such consensuated approaches. This situation hinders the modeling of threat actors, which prevents an accurate definition of control to identify and to neutralize malicious activities. In this paper, we propose the first global approach for CNA operations that can be used to map real-world activities. The proposal significantly reduces the amount of effort need to identify, analyze, and neutralize advanced threat actors targeting cyber physical systems. It follows a logical structure that can be easy to expand and adapt.

## 4.1 Introduction

An important threat against cyber physical systems is operations focused on its disruption or control. Although these activities can be directed against pure IT environments, the impact they can produce in a cyber physical system is usually much bigger, as it exceeds the cyberspace and materialize into the real-world, thus causing, for example, human losses. These activities are identified as Computer Network Attack (CNA), operations taken via computer networks to disrupt, deny, degrade, or destroy the information within computers and computer networks and/or the computers/networks themselves, as we will define later, and are performed by advanced threat actors, specially state-sponsored ones. Please note that when we refer to an *attack*, we are referring to disruption or manipulation operations, not just to any kind of attack (or cyber attack); this is an important point, as in many works the authors consider *attack* any operation against a technological infrastructure, no matter the objective of the attacker is.

CNA operations are much less analyzed than Computer Network Exploitation ones, or CNE, those related to cyber espionage; as an example, the main de facto standard to identify tactics and techniques—and to link them to advanced threat actors—MITRE ATT&CK, focuses mainly on CNE tactics and techniques, leaving CNA ones as a secondary aspect. This is due to the prevalence of CNE activities among advanced threat actors, which is a confirmation of the lack of studies to structure those attacks and its modus operandi. Although CNA operations are not as usual as those related to intelligence gathering, their importance is increasing with the expansion of cyber physical systems. An operation targeting these environments to destroy or to control them can cause damage not only to a specific industrial process, but to the society on the whole.

This work provides a structure of tactics and techniques for CNA operations and actors aligned with ATT&CK, thus complementing the effort that MITRE has done in this framework and helping to improve it. We have followed the

MITRE ATT&CK structure as the reference to develop those tactics and techniques, as it is the main public effort to establish a classification of Tactics, Techniques, and Procedures (TTP) used by threat actors, and we also propose an initial mapping for our approach to this classification. The tactics—and its associated techniques—linked to CNA activities will be submitted to MITRE to be included in the *Enterprise Tactics* section at MITRE ATT&CK. It is important to note that this work is the first proposal towards a global approach for CNA operations that can be used for establish a definitive taxonomy of TTP in those operations that nowadays have become a major threat for cyber physical systems. Moreover, a key point is that our proposal is not linked to specific components of a cyber physical system (for example, to improve specific detection mechanisms, those usually based on attack signatures in an expert system), but provides an upper approach to model and categorize threat activities against infrastructures.

One of the most important tactics is the manipulation one, as its effects are usually not noticed immediately, and so the damage on the cyber physical system can be larger; apart from that, it is the less analyzed tactic until now. In this tactic, we have identified a new research line that has to do with the identification of subcategories for manipulation families that are stated here, in order to refine the taxonomy and to cover all the relevant particular techniques; for example, inside the falsification family we could mention spoofing, a family of techniques in which an actor successfully masquerades as another by falsifying data, thereby gaining an illegitimate advantage [191].

The contributions of this paper are as follows.

- Identify the tactics linked to CNA operations.
- Discuss and establish techniques for each of the tactics identified.
- Define a structure for CNA tactics and techniques compatible with standards accepted among the community and suitable for its improvement.
- Identify a key research line for the structure and analysis of the manipulation tactic.

The rest of the paper is organized as follows. The background Section 4.2 provides a brief introduction to Information Operations—where CNA is located—and to the MITRE ATT&CK framework, as the main reference for the development of tactics and techniques. In Section 4.3, we assess the problem of the lack of a unified structure for CNA tactics and techniques, and its importance for the modeling of advanced threat actors. Section 4.4 analyzes the prior work in this field, stating that little research has been done, specially for the manipulation tactic. In Section 4.5, we propose a novel taxonomy for CNA tactics in which we identify four specific ones, and for each one of them we establish a classification for its associated techniques. In Section 4.6, we discuss the results of our work, comparing them with other approaches and identifying improvements. Finally, Section 4.7 summarizes the outcome of the overall work and identify future research lines.

## 4.2 Background

### 4.2.1 Computer Network Attack

Information Operations (IO) is defined as the integrated employment, during military operations, of information-related capabilities in concert with other lines of operation to influence, disrupt, corrupt, or usurp the decision-making of adversaries and potential adversaries while protecting our own [447]. One of the core capabilities inside IO is Computer Network Operations (CNO), which can be described as the *actions taken through the use of computers and networks to gain information superiority or to deny the adversary this enabling capability.* CNO is an umbrella term that comprises three main activities [414]: (1) Computer Network Attack (CNA), (2) Computer Network Defense (CND) , defensive measures to protect and defend information, computers, and networks from disruption, denial, degradation, or destruction [172], and (3) Computer Network Exploitation (CNE). While CND is about computer and network protection, whereas CNE is focused on information gathering, that is, in espionage (or cyber espionage).

CNA operations are [447] *those taken via computer networks to disrupt, deny, degrade, or destroy the information within computers and computer networks and/or the computers/networks themselves.* As we can see from its definition, CNA has been usually linked to the so-called 4D: disrupt, deny, degrade, and destroy [392], referring to destructive—with more or less impact—actions performed through computer networks. This one, *through computer networks*, is a key point: a missile launched against a data center is not CNA, despite the fact that it shall destroy computers and networks.

The CNA concept is currently under revision and may be soon replaced by a more generic one: Cyberspace Attack (CA), a capability inside Cyberspace Operations (CO) [450], defined as [625] *the employment of cyberspace capabilities where the primary purpose is to achieve objectives in or through cyberspace.* In this context, CA is a broader term than CNA: in fact, an interesting point in [625] is the relationship between CO and Electronic Warfare (EW). CA includes not only the 4D, but also manipulation, for example, in operations associated with deception, corruption, or usurpation. A compilation of these doctrine and terminology (and its use and history) can be found in [111, 657]. Every threat actor that performs CNE or CNA activities develops Tactics, Techniques, and Procedures (TTP) to achieve its goals [308]. Table 4.1 briefly describes the TTP definitions.

Tactics specify what a threat actor is doing, at the highest level of description, to accomplish a certain mission, and techniques specify how tactics are implemented and procedures—outside of the scope of this work—describe a particular implementation. These Tactics, Techniques, and Procedures represent the behavior of the actor, very similar to what we usually call its modus operandi, from the highest level description (tactic) to the lowest level one (procedure) [304].

Table 4.1: Tactics, Techniques, and Procedures (TTP) definitions.

| Definition | Description |
|---|---|
| Tactics | The employment and ordered arrangement of forces in relation to each other. |
| Techniques | Non–prescriptive ways or methods used to perform missions, functions, or tasks. |
| Procedures | Standard, detailed steps that prescribe how to perform specific tasks. |

### 4.2.2 MITRE ATT&CK

MITRE ATT&CK (*Adversarial Tactics, Techniques, and Common Knowledge*) is a globally accessible knowledge base of adversary tactics and techniques based on real-world observations. This knowledge, contributed by analysts worldwide, can be used as the base for the development of specific threat models and methodologies. Started in 2013 and published in 2015, ATT&CK develops a process for modeling an adversary's post-compromise behavior at a fine level. An excellent description of the framework and the work done can be found at [598].

MITRE ATT&CK framework is today's de facto standard to structure tactics and techniques of advanced threat actors. As of March 2019, ATT&CK had defined 11 enterprise tactics—those related to the activities of an attacker into its victim—and 223 enterprise techniques associated with those tactics. Apart from that, ATT&CK defines 15 pre-attack tactics—related to the activities of an attacker before compromising its victim—and 174 pre-attack techniques linked to them, as well as 13 mobile tactics—related to the compromise of mobile devices—and 66 mobile techniques. Beside tactics and techniques, ATT&CK identifies software—a generic term for tools, artifacts, malware, etc.—that can be used to implement one or more of the techniques, and which is out of the scope of this work.

MITRE ATT&CK also links Advanced Persistent Threat groups (APT) to tactics, techniques, and software. With 78 identified groups at the time of this writing, everyone of them is named, aliased, described, and linked to specific techniques (including pre-attack and mobile) and software. In this way, an analyst can establish relations between those entities to model an adversary and its activities against a target and, most important, to establish defense mechanisms to prevent, detect and respond to a threat.

Without any doubt, ATT&CK is an enormous effort to provide to the community an unified framework to identify the activities of advanced threat actors, from their TTP to the software they use, correlate information among those entities and improve not only the knowledge about APT, but also the defense mechanisms required to counter them. It constitutes a framework that, as usual, has to be improved with continuous work and contributions; in this sense, we miss in ATT&CK

a deeper approach to the tactics, the techniques, and even to the software related to CNA activities.

Until 2019 MITRE ATT&CK was focused on CNE rather than in destructive operations or threat actors. In April 2019, a new enterprise tactic called "Impact" was added, where they specify the techniques whose primary objective is to reduce the availability or integrity of a system, service, or network, including manipulation of data to impact a business or operational process. This Impact tactic is directly related to CNA activities, and includes fourteen techniques, such as defacements, data manipulation, or data destruction. Although this tactic from MITRE ATT&CK is a good starting point, it is mandatory to develop it in depth.

## 4.3 The issue

The impact of CNA operations in cyber physical systems has been largely discussed [533], especially since the discovery of Stuxnet [352]. This malware, Stuxnet, is a key piece in one of the best known manipulation operations. Although it was not the first one or the most noxious, it was the most reported [414]. Stuxnet is a cyber weapon, probably developed by United States and Israel, that compromised industrial systems to manipulate the centrifuges that enriched Uranium, so slowing down the Iranian nuclear program and directly impacting on a critical cyber physical system [417].

As stated before, CNA operations, those destructive—in more or less grade—or control-oriented, are much less analyzed than CNE ones; although many of a threat actor's TTP are common to CNE and CNA operations, in some cases, especially when dealing with actions on the target—or, depending of the tactic used, with actions against the target—they differ. As TTP are critical to identify and model threat actors, much work has been done to structure tactics and techniques, but most of the efforts in this sense are related to CNE activities. This focus on CNE may be due to two main facts:

- Many threat actors are engaged in intelligence gathering more than in destructive attack campaigns.
- Most of CNA activities require a previous CNE operation to know the attacked target, so CNE is almost always present.

During the last years, CNA operations that are publicly known are arising and we can face different threat actors engaged in both CNE and CNA operations; a good example is APT28, a group linked to the GRU, the Russian Military Intelligence Service, working not only in cyber espionage, but also in destructive campaigns against its targets. This is an important threat not only to pure IT environments but also to OT ones, where their impact exceeds the cyberspace by causing denial effects on cyber physical infrastructures, especially on critical ones. In fact, it is considered the fastest growing threat for cyber physical systems [403].

In this work, we address an issue that has not been largely approached and whose importance is increasing during the last years. Threat modeling, including the modeling of specific cyber attacks, has been largely discussed and many approaches exist nowadays to face the problem. However, when dealing with tactics and techniques for CNA operations no global view has been defined, a situation that hinders the modeling of advanced threat actors Most of the work regarding the structure of tactics and techniques for these actors have been focused on CNE operations, while CNA ones are not so structured. This situation has an obvious reason: as stated before, CNE is much common, and all advanced CNA operations require a previous CNE one. However, destruction and manipulation operations have increased during the last decade, as well as the actors performing them. They have become a major threat to cyber physical systems, especially against critical infrastructures, where the damage exceeds the cyber world and can impact in the real one, including a potential loss of life.

Therefore, we are facing an increasing problem, and the lack of a structure suitable for the modeling of these actors and its operations causes a weak protection against them. Without accurate capabilities to model CNA operations, starting with a structured view of its tactics and techniques as key elements, infrastructures are not well protected, especially cyber physical systems.

## 4.4 Approaches and limitations

Some works have established a taxonomy for computer network attacks; the authors of [637] provide an ontology whose goal is to automate the classification of a network attack during its early stages. In this ontology, the “Attack goal class” refers to the purpose of the attack, which could be considered equivalent to the tactic. The authors identify five goals in CNA operations: to change data, to destroy data, to disrupt data, to steal data, and a springboard to other goals. While this last goal cannot be considered a tactic itself and the theft of information is not inside CNA but inside CNE operations (although as stated before all CNA serious operations require a previous CNE one), the change, destruction, and manipulation of data are all considered in the taxonomy we propose; not only referring to data, but referring to all pieces of a cyber physical systems that can impact in the process they support.

As stated before, when dealing with CNA we usually refer to 4D: disrupt, deny, degrade, or destroy the information or the systems or networks themselves; this initial set of actions, expressed in many works since the 1990s [241, 293, 571, 612] has been superseded by two approaches: to consider denial as an umbrella term for the other 3D and to include manipulation as a form of attack, beside the denial tactics.

The first of these approaches considers denial as the actions to prevent access to, operation of, or availability of a target function by a specified level for a specified time [450]; in other words, denial is not a tactic, but an effect that can be achieved by degradation, disruption, or destruction. This approach is used in many works, and seems specially consolidated in modern doctrines [621], [78, 625], thus considering denial as a goal, but not as a tactic itself.

Apart from denial tactics, inside CNA operations we must consider a tactic called manipulation [450, 625], whose goal is to control or change the target's information, information systems, and/or networks in a manner that supports the attacker's objectives. Those objectives, in CNA operations, are clear: to cause denial effects against access or operation, goals that can be achieved by tactics such as degradation, disruption, destruction, and manipulation. However, following this approach, what is the difference between a manipulation that achieves a degradation or a direct degradation as a tactic? It is a very subtle one: mainly, manipulation refers to a manner that is not immediate apparent or detected: a DDoS (degradation or disruption) or a ransomware attack (destruction) are immediately identified by the victim. If the tactic was manipulation, the attack would not have been immediately detected, and would extend in time, so impact would have been higher in advance.

It is important to differentiate CNA from Electronic Attack (EA), a branch of Electronic Warfare (EW), another IO core capability. EA relies on the use of electromagnetic spectrum, while CNA relies on the data stream to execute the attack [668], although both capabilities are converging inside the cyber field; a good reference to identify EW and CNO relationships can be found in [586]. However, EA tactics and techniques could be equivalent to CNA ones in some cases, so we must consider them in our work.

Attack modeling methodologies have been largely discussed. A summary and analysis of some of these methodologies can be found in [399, 565] and of course in [567], in which it is still one of the main references in the topic. Particular approaches to model an attack have been developed:

- The Cyber Kill Chain®, developed by Lockheed Martin, is part of the Intelligence-Driven Defense® model for identification and prevention of cyber intrusions activity, specifying what a threat actor must complete in order to achieve their objective. This model represents an industry accepted methodology for understanding how an attacker will conduct the activities necessary to cause harm to the organizations and has been largely discussed [430].

- The Diamond Model of Intrusion Analysis (DMIA) [105] establishes a formal method applying scientific principles to intrusion analysis, providing a simple, formal, and comprehensive method of activity documentation, synthesis, and correlation. The model represents an *adversary* deploying a *capability* over some *infrastructure* against a *victim*; these activities, called *events*, are the atomic features of the model, and they define one step in a series that a threat actor must execute to achieve its goal.

- MITRE ATT&CK, previously discussed in this work.

- The Detection Maturity Level (DML) [594] is a model to assess the maturity of an organization to detect cyber attacks in terms of its capabilities to consume and act upon given threat information. The DML model is composed of nine levels of maturity, from the most technical ones—level 0 represents no information about the threat—to the highest level ones—goals and strategy of the attacker. This hierarchical model and its improvements [88, 387] can give an approach not only to evaluate the detection capabilities of an organization, but also to the semantic modeling of an advanced threat actor: from a group to specific indicators left after an operation, thus helping the analysts to model the interests and behavior of the actor and its modus operandi in specific operations.

With the exception of MITRE ATT&CK, none of these approaches focuses on the tactics and techniques of an attacker. The Cyber Kill Chain® tries to specify the steps an attacker has to perform to achieve its objectives, without discussing how these steps can performed (a work that is done in MITRE ATT&CK). The Diamond Model provides a framework to identify adversary operations and to discover new capabilities, infrastructure, and victims, but once again it does not focus on what and how a threat actor works. An adversarial approach like DML does refer to tactics, techniques, and procedures of an actor, but not in depth.

As tactics and techniques are not widely discussed in threat models, and the only approach focused on these aspects (MITRE ATT&CK) does not cover CNA techniques in a structured way, it is mandatory to consider particular approaches for specific techniques and to contextualize them in a global structure, as well as to analyze other kind of attacks related to cyber activities (this is, EA), trying to provide a global, common structure for CNA operations.

The work in [624] states that EA tactics are disruption, delaying, deviation, and denial, and refers to techniques such as deception, jamming, masking, and directed energy, in no particular structure. The work in [447] considers degradation, disruption, and deception as techniques to accomplish the denial tactic, together with destruction. At this point deception emerges, being unclear in literature review where is located in a tactics and techniques structure. Apart from the "4D" (deny, degrade, disrupt, and destruct) linked to CNA, some authors [625, 626] directly advocate the existence of a fifth "D" while dealing with EA tactics, referring to deception. In this work, we will consider deception as a particular manipulation with two possible goals: dissimulation, to hide the real, or simulation, to show the false; later we will discuss the role of equivalence between deception and manipulation.

Most of the literature reviewed is focused on degradation techniques, and inside them, in two main directions: the first one is the characterization of electronic attacks against WiFi networks; for example, in [222, 562] the authors propose a classification of DoS attacks in wireless infrastructures, based on the network layer in which the attack is performed, from the physical layer to the application one. As stated before, this kind of techniques is outside of the scope of this paper because they are closer to EA than to CNA.

The other research field has been DDoS attacks taxonomies, that is, taxonomies of attacks engaging multiple hostile actors to degrade or disrupt a victim through a network connection. In spite of the fact that some authors state that there is not a DDoS classification model [70] and try to define a general and simple scheme that differentiates between attacks on bandwidth, host resources, and weaknesses, other studies try to define a taxonomy for DDoS attacks; for example, early works like [180, 588] established a classification for DDoS attacks that has been improved during those years [128]. In this model, the authors distinguish between active and passive attacks (in this case, the only established category is packet dropping). While referring to active attacks, the second difference is based on the depleted object: bandwidth or an object mandatory to access the targeted system. First, bandwidth can be depleted by flood attacks (those that require a bandwidth usage larger in the attacker than in the target) or by amplification attacks, in which a simple request is amplified in the target system thus degrading its performance when many requests are made; a typical example of an amplification attack is DNS amplification. Related to resource depletion, it can be accomplished by exploiting flaws in network protocols or by malformed packets that fool those protocol's implementation.

As seen, depletion is a key technique inside degradation; some approaches try to establish a taxonomy based on the depleted resource. The authors of [180, 588] identify bandwidth depletion and resource depletion; they defend that the first one can be achieved by flooding and amplification techniques, while the second one, resource depletion, can be achieved by techniques based on exploiting vulnerabilities. Our approach is different in the sense that we do not differentiate between bandwidth or resource depletion—at the end, bandwidth is just a resource to be depleted, just like CPU or storage, joining all the techniques above depletion: for example, an application can be flooded by legit queries—not exploiting—to deplete CPU.

The less-studied tactic we are addressing deals with the manipulation of cyber physical systems; we are facing a complex and novel task, because there are no works on the subject, and the few literatures we found are confusing. Although there are many taxonomies and ontologies for attacks [8,577,636] (a good summary can be found in [635]), none of them focus on techniques, but most on the complete modeling of an attack (and mainly understanding attack not only as a denial or manipulation operation). Even a good work on terminology and concepts referred below [450] includes deception, decoying, conditioning, spoofing, and falsification as examples of techniques to perform the manipulation tactic; after analyzing literature and practical cases, we cannot agree with these examples: apart from the fact that this relation is not exhaustive, it mixes techniques that should be considered as particular instances of an umbrella category (spoofing is just a family of techniques linked to falsification). Other, less doctrinal studies [325], identify data modification and infrastructure manipulation as the tactics related to what we—and other references—have called manipulation and that should be considered as a single tactic.

Other relevant work on manipulation is CAPEC (*Common Attack Pattern Enumeration and Classification*), from MITRE, which defines different mechanisms of attack that include infrastructure manipulation, file manipulation, configuration or environment manipulation, software integrity attack, modification during manufacture, manipulation during distribution, hardware integrity attack, malicious logic insertion, resource contamination, and obstruction; as its name implies, this resource focuses on attack patterns, not on techniques performed to accomplish manipulation. For example, a pattern like file manipulation can be performed by many techniques but techniques define how, no matter if it is applied to configuration, file, infrastructure, or manipulation.

As we can see, there is no global approach to define a structure for CNA tactics and techniques. Network attack is a common problem to face nowadays, and some approaches have been developed to classify them from different (objectives, mechanisms ...). However, when dealing with TTP only partial approaches have been addressed, mainly focused on degradation techniques (and inside them, particularly inside DDoS techniques). These partial approaches do not cover the full spectrum of CNA activities and they are neither aligned with today key references. Even MITRE, which has developed standards such as CAPEC (focusing on patterns of attack, not in TTP) and, specially, ATT&CK, do not cover the full range of CNA operations in a suitable manner.

As far as we know, our work is the first global approach for CNA operations TTP structure; all work done can be considered partial, focused on specific techniques like some DDoS types, and without a common framework where to structure CNA TTP. The lack of an unified, well-accepted taxonomy of CNA tactics and techniques impacts in the security of cyber physical systems, which nowadays are one of the main targets of state-sponsored threat actors. We propose a novel approach for a structure suitable for the modeling of such actors and its activities, that can be used to prevent, identify, and neutralize operations against technological infrastructures. We cannot compare our approach against existing ones, because as we have stated, all the reviewed research focuses on particular techniques without a common framework for CNA tactic and techniques. Our approach follows MITRE ATT&CK structure, so it can be easily used in real world scenarios and can improve MITRE's effort to maintain a global shared knowledge about threat actors' modus operandi.

## 4.5 Proposed Approach: A Novel CNA Tactics Taxonomy

In this section, we present a novel structure for tactics and techniques linked to destructive and manipulative operations against cyber physical systems. Our approach includes manipulation as a tactic in CNA operations, but we will not consider denial as a tactic itself but a goal or a meta tactic that can be performed by degradation, disruption, and destruction—in fact, also by manipulation. Further-more, it is not differentiated degradation from disruption techniques: disruption is a particular case of degradation where degradation level is 100%, so all of the techniques families expressed in degradation directly apply to the disruption tactic. Therefore, although we can face four main tactics for CNA operations (degradation, disruption, destruction, and manipulation) as Table 4.2 shows, only for three of them (all but disruption, seen as a particular case of degradation) are we going to define categories and subcategories, where applicable, in order to classify specific techniques in each of them.

Table 4.2: CNA tactics.

| **Name** | **Description** |
|---|---|
| Degradation | Degradation consists [450] of techniques used to deny access to, or operation of, a target to a level represented as a percentage of capacity; if this percentage is 100%, we refer to disruption. |
| Disruption | Disruption consists of techniques used to completely but temporarily deny access to, or operation of, a target for a period of time; it is a degradation whose level is 100%. |
| Destruction | Destruction consists of techniques used to completely and irreparably deny access to, or operation of, a target. |
| Manipulation | Manipulation consists of techniques used to control or change the target's information, information systems, and/or networks in a manner that supports the attacker's objectives. |

Our criteria for the proposed structure of TTP are based primarily (tactics) on what an attacker is trying to achieve against its target to accomplish its objectives: that is, on the effects of the operation or attack. In this way, CNA operations try to degrade, disrupt, destruct, or manipulate. Once the tactics are stated, inside each of them we identify the different techniques the attacker can develop to get the desired effect (as stated before, techniques refer to *how* a tactic is accomplished). Here, we identify some techniques that can be seen similar to others (for example, alteration vs. modification or deletion vs. cancellation); the key difference between them is the tactic they are linked to, which reflects what the attacker is trying to achieve. At this point, it is also important to note that the same technique can be used to perform different tactics: for example, manipulation techniques can also be executed both to degrade or to destroy a target, depending on the particular assets attacked in the target or on the particular kind of the manipulation performed. When we consider deletion or encryption, linked to destruction, we are not seeing the destruction of particular files of a cyber physical environment, but the destruction of its functionality.

The different tactics identified and analyzed in this work can also be classified from an impact point of view. In this sense, although it is hard to establish a global classification, and the final impact will depend on many factors (as on the complexity of a particular campaign, or on the early detection of an operation in a target), the manipulation tactic can considered by far the most dangerous for cyber physical processes. For advanced threat actors, manipulation implies not only knowledge about particular industrial processes, but also a high control level of the targeted infrastructure. Following manipulation, destruction attacks are the most impactful ones for the target, as they imply an irreversible damage for cyber physical systems, that can only be restored by entirely rebuilding them. Disruption, as a fully degradation level, and degradation tactics can be considered the ones with less impact on the target, as the normal behavior of the affected cyber physical systems is usually recovered once the degradation operation is contained and the affected environments are turned again to a stable status. Please note that, in spite of its impact level, all CNA tactics can lead to fatal consequences, especially when the target is a cyber physical system for a critical infrastructure.

### 4.5.1 Degradation

Degradation is a tactic whose goal is to reduce the effectiveness or efficiency of command and control systems and information collections efforts or means [625], and can be implemented using various techniques, including what we usually call Denial of Service (DoS), an explicit attempt to avoid legitimate users of a service make use of it [602], or Distributed DoS (DDoS), depending on the number of attackers.

We propose a classification for degradation techniques based not only on depletion attacks, but considering other techniques to accomplish the tactic; we differentiate between depletion, interference, and alteration approaches, and identify techniques in all of them as shown in Figure 4.1. This proposal includes all the main techniques and its subtechniques, dealing with all the different mechanisms an attacker can use to degrade a cyber physical system, thus providing a full coverage of the degradation tactic.

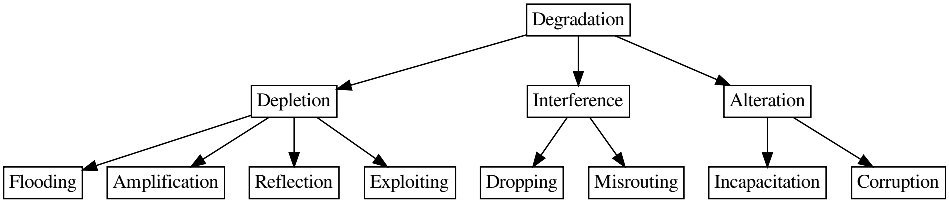

Figure 4.1: Degradation techniques classification.

#### Depletion

While dealing with degradation, the family of most studied techniques are those based on depletion—the degradation of an element whose work is mandatory to provide a service; depletion can be performed in any point between the target system and its legitimate users: the typically attacked points are the system itself or network devices whose degradation impacts in most users (usually, an attacker will want to maximize impact thus denying service to as many users as possible). In the special case of cyber physical systems, depletion attacks can also target physical elements of an infrastructure, although if its impact is not immediately noticed then we are facing not a degradation, but a manipulation tactic. Depletion can be achieved through different mechanisms, presented in this section.

Flooding techniques aim to degrade services by sending vast amounts of valid service requests, trying to degrade a key resource for the provision of the service: it may be the target system itself or an intermediate infrastructure mandatory to provide service to legit users, typically a network device. Usually, flooding is performed as an N-to-1 attack, where a lot of zombies send valid requests to a destination thus achieving a service degradation—known as Distributed Denial of Service (DDoS); it is a simple and cheap attack.

Amplification techniques are those that benefit from sending small requests to services that produce responses orders of magnitude larger in size, in such a way that with a reduced number of queries an attacker can exhaust the service; inside amplification we find techniques such like DNS or NTP amplification.

A third commonly accepted technique nowadays [226] is reflection; reflection techniques are those in which the hostile actor sends requests, spoofing its source address pretending to be the victim, to a reflector server. This reflector, unable to distinguish legitimate from spoofed requests, replies directly to the victim [42, 79]; launching a reflection attack from a botnet against a particular target will cause the target to be answered, thus degrading at least its bandwidth.

Finally, inside depletion, we can discuss exploiting techniques, those which are based on the exploitation of one or more flaws in a policy or in the mechanism that enforces the policy, or a bug in the software that implements the target system, and aims to consume excessive amount of resources of the target by sending it a few carefully crafted requests [6]. Inside exploiting we can consider techniques related to protocol exploiting or violation—including all layers, from infrastructure to application level—and also techniques related to application exploiting: malicious queries to a web application with a poor database structure can lead to CPU depletion on the web or database server, for example.

### Interference

Interference attacks in CNA activities work by intentionally inducing noise or injecting false data into the target, thus degrading its service. While regarding these techniques, it is mandatory to mention jamming, a subset of denial of service attacks in which malicious nodes block legitimate communication by causing intentional interference in networks [248] and which is a key threat to cyber physical systems [299, 364, 465, 674]; while jamming could be included inside an interference category for degradation or disruption operations, especially in WiFi sensor networks, we cannot consider jamming inside CNA techniques as it is related to Electronic Warfare. Anyway, interference techniques exist in CNA operations, with most well-known instances of this category relying upon features of the TCP protocol [57]; inside this category we can include packet dropping [118, 685], an attack whose goal is to make the source and the destination perceive disconnection or degradation of path quality. Of course, while referring to degradation not all packets are to be dropped—as it could be easily detected, and will fall into the disruption tactic—but only a subset of the packets is dropped, thereby making it more difficult to detect [685]. Although packet dropping is usually performed in wireless networks by EA techniques, it can also be used in wired infrastructures, mainly in network devices (just as routers). As stated, this kind of attack can be seen as techniques inside interference, as well as routing attacks: those related to the modification of routes that interconnect source and destination of a communication, from poisoning to black holing.

### Alteration

Another family of techniques inside degradation is alteration of system components. In this category, we can mention incapacitation, when the attacker disables one or more of key components, and change, when the attacker modifies key functions or data of the target (please note that the term “target” not only refers to a final infrastructure, but also to any point of the service delivery that can be attacked to degrade the service).

A particular tactic inside change techniques to achieve degradation is the corruption of system memory; it is a technique analyzed in many works [124, 540, 634]. Although it could be usually considered inside the destruction tactic—memory corruption attacks destroy the data, we identify it inside alteration techniques because it is performed against volatile memories and rebooting the system usually recovers its functionality (the goal of the attacker is not to destroy, but to degrade).

It is important to state that what we have called alteration differs from the manipulation tactic, although both imply the modification of key system components: in this context, alteration techniques must be seen inside the denial meta tactic, this is, they prevent access to, operation of, or availability of a target, while manipulation tactic tries to control or change information in a manner that supports the attacker’s objectives. For example, defacement can be seen as a technique inside alteration (change in particular): it directly denies the access to a legit resource, and although web defacement is usually a simple attack but it can be also performed by advanced actors—for example, for political purposes as we saw in Georgia 2008 attacks. Some authors distinguish between sophisticated and unsophisticated attacks (see in [571] for references).

### 4.5.2 Destruction

Although destruction can be performed through hardware, firmware, software, data, or network destruction, in most cases both logical and physical, while dealing with it in CNA operations targeting cyber physical systems we can identify techniques that delete firmware, software or data, that make them unusable—corruption—or that make them unusable unless a condition is met—encryption, unless the encryption key is known. All of them, when successfully used, damage its target in an irreversible manner: the target cannot perform any function or be restored to a usable condition without being entirely rebuilt [625]. Other destruction techniques, such as physical destruction, degaussing, or physical shredding, are considered outside CNA operations (some attacks, especially those against cyber physical systems, result in the physical destruction of one or more components of an infrastructure, but we will consider them inside the manipulation tactic: the attacker manipulates an industrial process, thus causing physical destruction), although they can also impact cyber physical systems.

Destruction techniques performed by advanced threat actors are usually executed against the components of the cyber physical system that most impact can cause (it is important to note that no destruction tactic can be executed without the destruction of one of these components); for example, file shredding would not be a regular technique because the file could be easily recovered from a backup, while disk wiping or cryptographic erasure are more common techniques in this context.

We propose three techniques to accomplish destruction, no matter which component they are performed against, as shown in Figure 4.2: deletion, the removal of key components; encryption, the encoding of those components thus rendering them or the system unusable; and corruption, the modification of key components with the same objective. In addition, the alteration techniques shown before can also be used to cause destruction of a target, although they can be included in the corruption family.

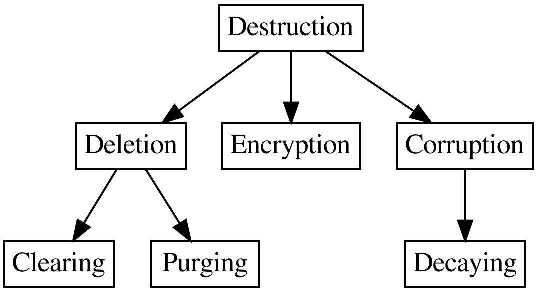

Figure 4.2: Destruction techniques classification.

### Deletion

Deletion is the removal of files on a target's system in order to interrupt availability in its services. In its simplest form (clearing), it applies logical techniques to sanitize data in all user addressable storage locations [331], protecting against simple noninvasive data recovery techniques. Clearing is usually performed through system commands or methods that do not really destroy the information—it could be retrieved by a forensic analysis, but the references, in the form of file system pointers, to it.

Of course, advanced actors do not usually perform clearing techniques on CNA operations, but erasure or purging, the logical removal of data from a storage so that it can no longer be read using state of the art laboratory techniques. Unlike clearing, purging is a real deletion of data. If it is performed against a storage structure, such as a hard disk or a file system, it is usually called wiping, while if it is performed against individual files or folders, it is called shredding, which destroys data by overwriting the space used by the object with a random pattern.

### Encryption

Encryption is a technique used to achieve destruction by encoding, in an irreversible way for the victim, data stored into a system. This non-reversible way implies that the victim does not have access to the decryption key; if access to this key is granted, data can be recovered. In some contexts, the use of this kind of cryptographic techniques when the decryption key has been destroyed is called cryptographic erasure and is an accepted technique for legit data sanitization [331]; if the decryption key does not even exist it shall be considered as corruption.

Although in some cases it is been possible for the analysts to recover the encrypted data, due to weaknesses on the encryption algorithm or in its implementation, this fact has been seen in some general, non-directed, ransomware, but no case linked to advanced actors in CNA operations is known.

### Corruption

Finally, in the destruction context, corruption can be seen as the deliberate modification of information—firmware, software, or data—to render it unusable; of course, while considering corruption in CNA operations context, we are referring to intentional corruption, not to unintended changes to data caused by errors. A good example of corruption in computer network attacks, although not performed by advanced actors, is CIH virus, which in the late 1990s was able to replace boot-time code in particular BIOS with junk, rendering systems unusable until their BIOS was replaced.

A particular case inside corruption is decaying, techniques linked to a gradual corruption of data; although decaying is usually caused by failures in computer systems, with factors like media type, file systems design or preallocation strategies [197], it also can be considered a progressive, not easily detected, CNA technique.

### 4.5.3 Manipulation

Manipulation is an attack against integrity and, while in CNA, its goal is clear: to create denial effects [450]. Manipulation alters its target not to enable intelligence gathering—as in CNE operations—but to damage it in an manner that is not immediately apparent or detected and, in many cases, that manifests in physical domains [625]. This one, not to be immediately apparent, is a key concept in manipulation attacks; as opposite to other tactics inside CNA, manipulation degrades or destroys its target without being detected. While all CNA tactics have become an important threat to cyber physical systems, manipulation is perhaps the most dangerous one, as it extends over time, so its impact is usually higher than these of other CNA tactics. For example, Stuxnet, previously referred in this work, influenced the development of cyber weapons not only from a technical perspective, and marked a milestone on the security of cyber physical systems [173].

As previously stated, we are addressing a novel task when establishing a categorization for manipulation techniques; our position is to keep it as simple as possible, and following this approach, we can subdivide manipulation techniques into three families: fabrication, modification, and cancellation, as stated in [314, 483]. These three families are capable of including different threats classification, for example, those presented in [443] against integrity, which refer to substitution, change (both considered inside modification), removal (cancellation), and addition (fabrication).

All of these types of attack can be applied against a target from the lower part of a cyber physical infrastructure [174, 559] to the higher one, and can be performed during the transmission, storage, runtime, as well as during manufacturing or supply of information. In Figure 4.3, we show the structure for this simple manipulation techniques classification.

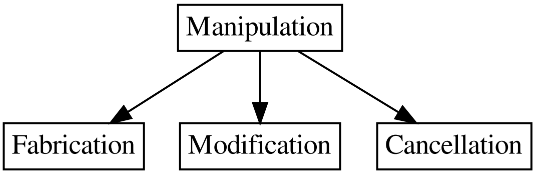

Figure 4.3: Manipulation techniques classification.

The definitions we could state for these techniques are obvious, as shown in Table 4.3, always remembering the goal of the tactic (denial) against cyber physical systems:

Fabrication techniques are those that include additional data into a system; these data can range from a single parameter in a configuration file to a piece of malware that causes an incorrect behavior of the target. Modification techniques alter existing legit data, changing it with logic, inputs, outputs, etc. that will cause a malfunction on the target. Finally, cancellation deals with the removal of certain data which is critical to the correct behavior of the target system; although cancellation is just a synonym for deletion, the techniques linked to destruction and those linked to manipulation are not similar: while in manipulation, we are referring to techniques that are not immediately detected, so their goal is not the destruction of data itself but a stealth removal of key elements in order to cause a malfunction of a process.

Table 4.3: Manipulation techniques.

| Name | Description |
|---|---|
| Fabrication | Inclusion of spurious data into a key element (or elements) in order to cause a malfunction of the target system |
| Modification | Replacement of legitimate data with malicious one, also in order to achieve a malfunction |
| Cancellation | Removal of key data in the target, with the same proposal than previous techniques |

Under this simple structure we can cover all techniques introduced in works cited before, like in [450], and it is also consistent with the techniques stated in MITRE ATT&CK for the Impact tactic regarding data manipulation (runtime, stored, and transmitted). In the same way, all CAPEC mechanisms for manipulation (data structures, system resources, and timing and state) can be linked to the proposed manipulation techniques structure.

### 4.5.4 Summary

In our proposal, we state that the tactics associated to CNA operations are degradation, disruption, destruction, and manipulation (we shall no longer mention 4D, as denial can be considered just a meta tactic); for all four tactics ("what") we have structured techniques ("how") families, with more or less detail, trying to define a common base to classify particular techniques in each of these families or categories. For example, a SYN Flooding is considered a technique to achieve degradation or disruption, so obviously it should be classified inside these tactics; more specifically, and following our proposed structure, it should be classified inside the Depletion category and the Flooding subcategory.

The proposed structure that we have developed in this paper is summarized in Table 4.4. We summarize the different identified techniques for each CNA tactic and, if applicable, we also propose subtechniques inside techniques. This structure also follow MITRE ATT&CK that, as we have stated, is today's de facto standard.

Table 4.4: CNA techniques structure proposal.

| Tactic | Layer 1 Techniques | Layer 2 Techniques |
|---|---|---|
| Degradation and Disruption | Depletion | Flooding<br>Amplification<br>Reflection<br>Exploiting |
| | Interference | Dropping<br>Misrouting |
| | Alteration | Incapacitation<br>Change |
| Destruction | Deletion | Clearing<br>Purging |
| | Encryption | |
| | Corruption | Decaying |
| Manipulation | Fabrication<br>Modification<br>Cancellation | |

### 4.5.5 Mapping to MITRE ATT&CK

As stated in this work, MITRE ATT&CK is the main public effort to establish a classification of TTP used by threat actors; for this reason, we have developed our approach following this standard, and as interesting exercise we propose a mapping of the MITRE ATT&CK "Impact" tactic (where the standard places the techniques oriented to manipulation and destruction) to our proposed structure.

At the time of this writing, MITRE ATT&CK "Impact" tactic (last modified on 25th July 2019), identified as TA0040, consists of techniques that adversaries use to disrupt availability or compromise integrity by manipulating business and operational processes. For this particular tactic, MITRE identifies the techniques shown in Table 4.5.

As we can see, MITRE ATT&CK provides no structure for techniques inside the "Impact" tactic, but places all of them at the same level under the tactic. Although this approach is followed in all ATT&CK tactics and techniques, given the importance of CNA activities, and its trends during last years, we consider it is mandatory to provide a wider detail, at least by splitting “Impact” tactic into the three main targets we propose in this work.

T1531, Account Access Removal, refers to the inhibition of access to accounts utilized by legitimate users; this is obviously not any destruction, but a “Degradation” or disruption tactic, and it maps to the “Change” subtechnique in our approach, stated as modification of key functions or data of the target.

T1485, Data Destruction, refers to render stored data irrecoverable by forensic techniques; it is a technique inside the “Destruction” tactic, mapping directly to “Deletion” technique in our approach, and more specifically T1485 maps to the “Purging” subtechniques inside deletion.

T1486, Data Encrypted for Impact, refers to the non-recoverable encryption of systems data or information. Like T1485, it is linked to the “Destruction” tactic and it maps directly to the “Encryption” technique in our approach.

T1565, Data Manipulation, refers to the insertion, deletion, or manipulation of data in order to hide activities. This technique is considered a tactic in our approach, mapping obviously to “Manipulation”; depending on the type of manipulation performed in each case, it could be mapped to specific techniques (fabrication, modification, or cancellation). This example shows why MITRE ATT&CK “Impact” should be detailed more in depth, as it is considering a whole tactic for CNA operations as a specific technique.

T1491, Defacement, refers to the modification of visual content; MITRE ATT&CK considers two types of defacement, internal and external, based on from where the modified content can be accessed. Following our approach, T1491 is considered inside “Degradation” tactic, particularly mapping to the “Change” subtechniques, inside the “Alteration” technique, as we have stated previously in our work.

T1561, Disk Wipe, refers to the destruction of data (from content to structure) stored in disks. Like T1485, it is linked to the “Destruction” tactic and maps to the “Purging” subtechniques, inside the “Deletion” technique. Unlike MITRE ATT&CK, in our approach we do not differentiate what type of data is purged, being in this way independent from specific approaches, close to particular technologies but not suitable for a high-level classification.

T1499, Endpoint Denial of Service, refers to the degradation or disruption of service to legitimate users. It is obviously linked to the “Degradation” (or disruption) tactic, more particularly mapped to the “Depletion” technique. Depending on how this Denial of Service is performed by the attacker, it can be mapped to particular subtechniques; MITRE ATT&CK establishes different subtechniques, where all the “Flood” ones can be mapped to “Flooding” in our approach and “Exploitation” maps directly to the “Exploiting” subtechnique in our approach.

T1495, Firmware Corruption, refers to the corruption of firmware in devices attached to a system in order to render them inoperable. As the specific firmware is targeted in an irreversible manner, we map this technique directly to the “Corruption” technique inside the “Destruction” tactic.

T1490, Inhibit System Recovery, refers to the removing of built-in capabilities designed to aid in the recovery of a corrupted system. Depending on how this technique is performed, it can be mapped to different techniques in our approach. MITRE ATT&CK specifies two possible actions to impact on system recovery features: to disable or to delete them. In the first case, this technique would map to "Incapacitation", a particular subtechnique of "Alteration" inside the "Degradation" tactic. In the second one, that referring to deletion, the technique maps obviously to the "Deletion" technique inside the "Destruction" tactic.

T1498, Network Denial of Service, refers to the degradation or disruption of the availability of targeted resources to legitimate users. As T1499, it is obviously linked to the "Degradation" (or disruption) tactic, more particularly mapped to the "Depletion" technique. Moreover, as in T1499, depending on how this Denial of Service is performed by the attacker, it can be mapped to particular subtechniques; in this case, MITRE ATT&CK establishes two subtechniques, the first one regarding a flooding approach and the second one regarding a mix of reflection and amplification techniques. Inside our proposal, these subtechniques can also be directly mapped to "Flooding", "Reflection", or "Amplification". Please note that MITRE ATT&CK does not establish a subtechnique for "Exploitation", as proposed in our work.

T1496, Resource Hijacking, refers to the leverage of resources impacting in availability. Although it could be considered a technique inside "Degradation", a particular case of a Denial of Service (in MITRE ATT&CK approach, T1499, Endpoint DoS), we cannot consider it a valid technique in a CNA approach. Of course, it is an action could lead to the degradation or disruption of particular services (those hosted on the targeted system), what attackers are trying to achieve is not this tactic, but simply an economic benefit from the earning of cryptocurrencies (as MITRE ATT&CK) states as most common. In this case, degradation is just a secondary effect from the operation, not what the attacker is achieving; if degradation was the desired effect, inside our approach this one would be considered a "Depletion" technique in most cases.

T1489, Service Stop, refers to the stopping of services on a system to render them unavailable to legitimate users. As no destruction is considered, this technique is linked to the "Degradation" or "Disruption" tactic; more specifically T1489 is mapped to the "Incapacitation" subtechnique inside the "Alteration" technique.

Finally, T1529, System Shutdown/Reboot, as it name implies refers to the shutdown or reboot systems to interrupt access to, or aid in the destruction of, them. In this particular case, our approach would link this technique to the "Degradation" tactic, specifically to "Alteration" techniques, particularly to the "Change" subtechnique: the attacker is modifying key functions of its target. In no way would T1529 be linked to the "Destruction" tactic: in spite of the case a reboot can be used to aid in the destruction of a target, as MITRE ATT&CK states, it is not a destructive technique by itself but a support action.

All identified MITRE ATT&CK techniques for "Impact" tactic can be mapped to our proposed structure; apart from that, different techniques and subtechniques showed in our work do not appear in MITRE ATT&CK actual specification. In Table 4.6, we summarize the mapping of MITRE ATT&CK techniques to the structure we propose in our work.

At this point, it is important to note that we are using MITRE ATT&CK for references, not ATT&CK ICS. This one is a knowledge base useful for describing the actions an adversary may take while operating within cyber physical systems, but we consider it an ongoing effort focused on the goals of the attacker, not in its tactics and techniques. For example, the techniques identified in this knowledge base are not techniques (objectives) from a pure point of view, but global goals of the attacker: a sample one, "Damage to property", cannot be considered a technique but an impact the threat is trying to cause.

Table 4.5: MITRE ATT&CK Impact techniques.

| **Techn. ID** | **Name** | **Sub Techniques** |
|---|---|---|
| T1531 | Account Access Removal | N/A |
| T1585 | Data Destruction | N/A |
| T1486 | Data Encrypted for Impact | N/A |
| T1486 | Data Manipulation | Stored Data Manipulation<br>Transmitted Data Manipulation<br>Runtime Data Manipulation |
| T1491 | Defacement | Internal Defacement<br>External Defacement |
| T1561 | Disk Wipe | Disk Content Wipe<br>Disk Structure Wipe |
| T1499 | Endpoint Denial of Service | OS Exhaustion Flood<br>Service Exhaustion Flood<br>Application Exhaustion Flood<br>Application or System Exploitation |
| T1495 | Firmware Corruption | N/A |
| T1490 | Inhibit System Recovery | N/A |
| T1498 | Network Denial of Service | Direct Network Flood<br>Reflection Amplification |
| T1496 | Resource Hijacking | N/A |
| T1489 | Service Stop | N/A |
| T1529 | System Shutdown/Reboot | N/A |

Table 4.6: MITRE ATT&CK techniques mapping.

| Tactic | Layer 1 Techniques | Layer 2 Techniques |
|---|---|---|
| Degradation and Disruption | Depletion (T1499, T1498) | Flooding |
| | | Amplification |
| | | Reflection |
| | | Exploiting |
| | Interference | Dropping |
| | | Misrouting |
| | Alteration | Incapacitation (T1490, T1489) |
| | | Change (T1531, T1491, T1529) |
| Destruction | Deletion (T1490) | Clearing |
| | | Purging (T1485, T1561) |
| | Encryption (T1486) | |
| | Corruption (T1495) | Decaying |
| Manipulation (T1565) | Fabrication | |
| | Modification | |
| | Cancellation | |

### 4.5.6 Practical Example

The specification and structuring of CNA tactics and techniques provide organizations a higher capability to identify and analyze threats and, most important, to map defensive controls to mitigate these threats. Our proposal does not focus on specific attack detection, but on the modeling of threats regarding its objectives. Using abstract models for threat modeling allows analysts to be independent of specific technologies or systems, thus facilitating a global definition of security requirements and the implementation of defense mechanisms [429].

Focusing on a specific example, we can consider the most common techniques for degradation and disruption: those related to pure Denial of Service (DoS). In the MITRE ATT&CK approach, they identify two techniques for performing DoS: those based on the endpoint and those based on the network. For Network DoS, MITRE ATT&CK defines two subtechniques, Direct Network Flood (0.001) and reflection and amplification (0.002), and for endpoint DoS they define four subtechniques, being three of them based on flooding of resources (OS, service and application, 0.001, 0.002, and 0.003) and the last one (0.004) on the exploitation of vulnerabilities in application or systems. For all of these techniques and subtechniques, MITRE ATT&CK identifies a single mitigation countermeasure: to filter network traffic.

By including endpoint and network, this approach identifies flooding, reflection and amplification, and exploitation; for all of them, as stated before, there is a single mitigation technique based on the filtering of network traffic. We cannot agree with this approach because it mixes different techniques and does not identify appropriate countermeasures in each case. Our proposal provides a more specific classification of tactics and techniques, so it can be used to identify more suitable countermeasures in each case.

In first place, amplification is not considered as a specific technique, in spite of the given name; amplification is not a subset neither a specific subtechnique inside reflection, because both techniques differ in how they are deployed, so they also differ in how they are mitigated. Countermeasures against amplification rely in many cases in the implementation and configuration of specific applications, while the ones regarding reflection do not rely on these aspects. In this way, we cannot deal with protection against amplification the same way we deal with protection against reflection. By considering them at the same technique level, countermeasures against them will not be appropriate.

The approach followed by MITRE ATT&CK also considers network denial of service only by exhausting the network bandwidth services rely on, while endpoint DoS are considered those that denies the availability of a service without saturating the network used to provide access to the service. This simple approach does not cover the possibility of interference in the network, that degrades its performance not by a depletion technique, so a potential threat actor performing interference operations (as such seen on Electronic Warfare) would not be considered, so no countermeasures would be applied to mitigate those attacks. This is an interesting point: as shown in Table 4.6, interference is the only technique not even considered in the MITRE ATT&CK approach, once again reflecting the absence of a unified common structure for tactics and techniques in this field of operations.

Apart from this two simple examples relating the mapping of defensive controls to specific attack techniques, a valid, complete, structure for CNA tactics and techniques allows organizations to improve the identification of threat offensive capabilities in a threat modeling approach as well as the attribution of specific operations. For a threat actor who can execute encryption techniques for destruction, is easiest to perform corruption or deletion techniques than to perform manipulation ones, on a prior basis.

Finally, tactics specification in three main categories provides potential information about threats not only regarding their objectives (what they are trying to do) and intentions, but also about their capabilities; in general terms, tactics can be seen from less (degradation and disruption) to more complex (manipulation). Manipulation attacks usually require specific knowledge about the target, its processes and its technologies, while a degradation attack, for example, based on depletion techniques, does not require these skills.

## 4.6 Discussion

We have identified an absence of a suitable structure of the tactics and techniques commonly used by advanced threat actors; even MITRE ATT&CK, as a key reference in the subject, lacks an approach for disruptive and manipulative operations, and we can assess that it has not been defined until now. As the number and impact of these operations are increasing in the last years, until they have become a major, significant threat for cyber physical systems, we consider it mandatory to establish such an structure that allows the prevention, detection, and neutralization of these operations.

Our work states an initial classification for that structure, following commonly accepted frameworks such as MITRE ATT&CK, thus allowing the identification of operations and the early adoption of appropriate countermeasures. We identify the main tactics specific techniques linked to this structure; of course, we do not try to provide an exhaustive compendium of particular techniques, but to identify the most relevant ones and to classify them into the defined categories. This provides a novel approach to the problem of structuring CNA operations, a mandatory requirement for the modeling of advanced threat actors' activities.

Until now, there was no doctrine providing this base model, a fact that directly impacts on the security of cyber physical systems. Most of the previous research has been focused on the study of particular techniques, such as Denial of Service, or in the analysis of particular operations, such as Stuxnet. No global structure for CNA threat actors' activities or operations has been previously stated, a fact that hinders the knowledge of those activities and the modeling of the threat actors behind them. As we have shown in previous sections, this lack of a global approach can lead to improper countermeasures against threat actions, thus degrading the security of cyber physical systems.

To specify an initial approach in a field where no doctrine has been established is a complex task. In the case of CNA tactics and techniques, it has been mandatory to analyze other disciplines inside Information Operations, such as Electronic Attack, in order for us to compare methods and concepts. It the evaluation of disruptive approaches has also been relevant for the less analyzed tactic, manipulation, where no previous work, neither partial, has been identified.

We provide the basis for the identification of techniques in CNA operations, generating a practical structure ready to be used in commonly accepted standards, such as MITRE ATT&CK. We have provided a proposal mapping to this standard, and it would be also interesting to expand and to improve the information provided by MITRE ATT&CK regarding particular techniques by identifying activities from advanced actors that perform them; for example, APT28, linked to Russian GRU, used this technique during 2017 attacks against Ukraine (NotPetya, BadRabbit), encrypting hard drives with a non-reversible algorithm and rendering them inoperable [438].

As we have previously stated, the manipulation tactic is the less studied and structured. We have proposed a taxonomy for techniques inside this particular tactic but, in every case, a future research line is to compare manipulation techniques as those stated while referring to deception in classic references like [63]. If we look at the definition we have stated in this paper, and we replace the technical aspects with cognitive ones we could consider that (technical) manipulation is just (human) deception; in fact, deceit is just an active manipulation of reality in order to manufacture it, alter it, or hide it [300]. A key reference, such as the work in [450], refers to some techniques directly linked to deception. Under this approach, manipulation (deception) techniques should include those to hide the real, or dissimulation, and those to show the false, or simulation [24, 63, 285]. Under dissimulation the authors identify masking (hiding the real by making it invisible), repackaging (hiding the real by disguising) and dazzling (hiding the real by confusion), and under simulation it is included mimicking (showing the false through imitation), inventing (showing the false by displaying a different reality) and decoying (showing the false by diverting attention) [143]. Although deception is considered a usual tactic in CND operations [23, 257, 692], their study from an offensive point of view is not as usual, so we consider it as a key research line.

Applying a valid up-to-date taxonomy may guarantee a significant advantage in terms of appropriateness and fitness for attack prevention; in [545], the authors present a taxonomy that is tested against the behavior of sensors modeled as agents. Other relevant attack taxonomies are those presented in [268, 278], but none of them is based on TTP for those attacks. Such contextualization should provide a global framework for the prevention, identification, and neutralization of attacks against cyber physical systems.

Finally, the use of machine learning approaches to detect intrusions against cyber physical systems, and its contextualization in a taxonomy of CNA tactics and techniques may be also a relevant research line. In [375, 532, 670], the authors provide relevant approaches, while in [311] the authors provide approaches not only in the detection, but also in the classification of such attacks. Loukas et al. [374] provides a taxonomy focused not in tactics or techniques regarding attacks, but in intrusion detection systems characteristics and architectures designed for vehicles. Extending those approaches for general taxonomies relevant to cyber physical systems and aligning them with a suitable classification, as the one provided in this work, could be an interesting research line.

## 4.7 Conclusions

Destructive and manipulation activities against all kind of targets, but especially against cyber physical systems, have increased in the last years. A classification and structure for the tactics and techniques linked to these operations is a must in order to identify capabilities, to profile advanced actors and to implement security controls to counteract them. However, few researches have been done in this sense, so analysts have almost no references to establish capabilities, families, etc. For destructive or manipulation activities performed by advanced actors, not even military doctrine has been found with the required depth level.

In this work, we propose a novel approach to identify and structure the techniques followed to perform each of the tactics linked to CNA operations against cyber physical systems. We have identified the main tactics accepted nowadays (degradation, disruption, destruction, and manipulation) and, for each of them, we have discussed and proposed the different techniques suitable to accomplish each of the tactics. Where applicable, we have also identified subtechniques. The proposed structure is aligned with MITRE ATT&CK, the main effort and the de facto standard to identify and analyze tactics and techniques from advanced threat actors. This proposed novel structure of tactics and techniques significantly contributes to improve the threat model of CNA actors.

Tactics and techniques are one of the first key points to model those actors and to deploy capabilities in order to prevent, to detect and to neutralize them, thus increasing the security not only of cyber physical systems, but of all technological infrastructures. We consider our proposal as the first, and therefore the starting point towards a commonly accepted taxonomy that helps researches to better know hostile actors, especially advanced ones, performing CNA operations.

In future works, our proposal can be used as a fundamental basis for new and more refined approaches. In this sense, manipulation techniques are the less structured ones until now—in fact, manipulation is not always considered as a tactic inside CNA operations—so in this particular case we identify an important research line.

# Chapter 5

# SOC Critical Path: a defensive kill–chain approach

*In this chapter a novel defensive kill–chain model for the detection of hostile activities is proposed. Although kill–chain approaches have been largely used from an offensive point of view, no defensive proposal existed until now. With this model, defensive teams increase their performance through a common working framework, thus allowing them to identify acquisition and analysis gaps.*

This chapter was published as: Villalón-Huerta, A.; Marco-Gisbert, H; Ripoll-Ripoll, I. SOC Critical Path: A Defensive Kill Chain Model. IEEE Access, vol. 10, pp. 13570-13581, 2022. https://doi.org/10.1109/ACCESS.2022.3145029

## Contents

Different kill chain models have been defined and analyzed to provide a common sequence of actions followed in offensive cyber operations. These models allow analysts to identify these operations and to understand how they are executed. However, there is a lack of an equivalent model from a defensive point of view:

this is, there is no common sequence of actions for the detection of threats and their accurate response. This lack causes not only problems such as unstructured approaches and conceptual errors but, what is most important, inefficiency in the detection and response to threats, as defensive tactics are not well identified. For this reason, in this work we present a defensive kill chain approach where tactics for teams in charge of cyber defense activities are structured and arranged. We introduce the concept of SOC Critical Path (SCP), a novel kill chain model to detect and neutralize threats. SCP is a technology–independent model that provides an arrangement of mandatory steps, in the form of tactics, to be executed by Computer Network Defense teams to detect hostile cyber operations. By adopting this novel model, these teams increase the performance and the effectiveness of their capabilities through a common framework that formalizes the steps to follow for the detection and neutralization of threats. In this way, our work can be used not only to identify detection and response gaps, but also to implement a continuous improvement cycle over time.

## 5.1 Introduction

The high benefits technology has contributed to are questionless, but so are the risks it introduces on a daily basis for individuals, organizations and countries. Hostile actors such as foreign countries, terrorist groups and organized crime are well aware of these risks and they take advantage of them, from cyber crime to cyber war. In this context, Computer Network Operations (CNO), defined [425] as those capabilities used to attack adversary computer networks, defend our own and exploit enemy computers to enable intelligence gathering play a prominent role, both in offensive operations, such as attack or exploitation, and in defensive operations.

The defensive CNO discipline is called Computer Network Defense (CND) and it is defined as [425] those actions taken through the use of computer networks to protect, monitor, analyze, detect and respond to unauthorized activity in own information systems and computer networks. CND activities are provided by a Security Operations Center (SOC), a center focused on the prevention, detection and response to security incidents [335]. Defensive centers for cyber security adopt different names [690] [89] [646], such as SOC, CSIRT, CERT, CSOC, ISIRT, etc., according to its specific functions, capabilities or even licenses. In this paper, we will refer to all of them as SOC.

In this context, a SOC has a clear goal of preventing, detecting and neutralizing cyber threats. Of course, the achievement of this high–level simple goal, is a complex task. Leaving aside the prevention activities, such as those related to systems hardening or user awareness, and focusing on the detection and neutralization activities, a SOC detects and responds to incidents. However, no single common definition of the mandatory tactics and their arrangement to perform this task have been identified among literature. While offensive kill–chain models are well defined and discussed, no defensive equivalent has been ever proposed. This situation leads SOCs to work by following non–structured approaches, a fact that impacts not only in their detection and response capabilities, but also in their effectiveness and improvement over time.

We define the SOC Critical Path (SCP) as the sequence of mandatory activities to detect and to neutralize a threat. SCP is a kill chain model, which is a specific arrangement of actions, identified as tactics, mandatory to achieve a goal. In this case this goal is the SOC goal, as stated before. Such a kill chain model formalizes SOC activities and provides defensive centers with the mandatory tactics they have to sequentially implement so as to be successful; as in offensive kill chain approaches, it not only helps to structure activities, but also to identify gaps and improvements in this implementation. In most cases SOC activities are done in an unstructured, not fail–safe way, thus giving hostile actors the opportunity to succeed in their activities.

In this work, we propose SCP as a model to detect and enable the neutralization, of threats. We consider this model to provide a platform–agnostic SOC kill chain, identifying the mandatory, sequential actions a SOC must perform. Each of the presented tactics can be deployed by different techniques, out of the scope of this paper, and some of them can also can be divided into sub–tactics, not mandatory but relevant to the detection and neutralization.

The contributions of this paper are as follows:

- To identify the mandatory tactics to detect security incidents, thus enabling its appropriate handling by defensive teams.
- To provide the basis for a global SOC detection and response process, not only linked to pure incident response, establishing a whole continuous improvement cycle.
- To establish the proper arrangement of the identified tactics in the form of a kill chain model. As this arrangement is mandatory for a SOC, it defines the correct sequence of tasks to be performed to detect and respond to incidents, i.e., the defensive kill chain.
- To identify the most relevant sub–tactics for each of the main established first–level tactics in order to help SOC analysts to develop particular techniques to achieve them.

The rest of this paper is organized as follows. The background in Section 5.2 provides concepts regarding the identified problem in SOC processes and emphasizes kill chain approaches, which are always defined from an offensive perspective. In Section 5.3 we assess the problem of the lack of a unified structure for CND operations and its importance in the identification of situations on the infrastructure that can lead to a security incident. Section 5.4 analyzes prior work on this issue. In Section 5.5 we propose the SCP model as a global sequence of tactics to be performed by blue teams to identify incidents, as well as an example regarding the practical application of our model. Section 5.6 discusses the results and compares them with those of other approaches. Finally, Section 5.7 concludes the paper and identifies future research directions.

## 5.2 Background

In this section we provide the necessary background related to SOC processes and kill chain model approaches that will be the pillars not only to understand the identified issue in Section 5.3 but also to properly follow the proposed SCP model in Section 5.5.

### 5.2.1 SOC concepts

As in any cyber operation, to be able to achieve its defensive goals, a SOC has to develop tactics, techniques and procedures. Tactics specify what to do, at the highest level of description, to accomplish a certain mission, while techniques specify how tactics are implemented; procedures, outside of the scope of this work, describe a particular implementation.

These tactics and techniques enable an effective threat detection and neutralization in a SOC. A threat is defined [603] as any circumstance or event with the potential to adversely impact organizational operations (including mission, functions, image, or reputation), organizational assets, individuals, other organizations, or the Nation through an information system via unauthorized access, destruction, disclosure, or modification of information, and/or denial of service; in this reference the authors identify four types of threat sources:

- Adversarial. Individuals, groups, organizations, or states that seek to exploit the organization's dependence on cyber resources.
- Accidental. Erroneous actions taken by individuals during the course of executing their everyday responsibilities.
- Structural. Failures of equipment, environmental controls, or software due to aging, resource depletion, or other circumstances that exceed the expected operating parameters.
- Environmental. Natural disasters and failures of critical infrastructures on which the organization depends, but which are outside the control of the organization.

To detect and neutralize a threat, from the adversarial to the environmental ones, any SOC has to manage a high volume of initially unstructured information: to acquire and analyze data from multiple sources, and to turn this high volume of information into an actionable, reduced, set of data. This is done through a set of technologies and processes represented in the so–called SOC funnel: the conversion of millions of inputs in a few actionable outputs suitable for management by a human team (the blue team).

The definitions for these sets of data, from millions of inputs to the reduced set of actionable outputs, are not clear among professionals [436]. In this work, we use the key definitions presented in this section.

A log message, or simply a log, is the minimum information unit generated by an information system [436], such as an application, operating system or database.

Log data are linked to and stored in the information systems that generate it. Some or all of the generated logs are sent to a SIEM and also called events. SIEM systems were designed [69] to collect events from different sources, normalize them to a common following a common syntax and structure, and store them once normalized.

An **event** is defined as a relevant, from a security point of view, contextualized information that reflects an observed change in the normal behavior of an object, such as an information system or a person. All events are stored in a SIEM, and three types are identified:

1. Raw events, those that automatically arrive to the SIEM from different data sources.

2. Aggregated events are those generated by the aggregation of raw events for a more efficient processing or analysis.

3. Correlated events, those generated by the correlation of raw or aggregated events.

An **alert** is an interesting event that requires spawning actions; alerts are usually managed on a ticketing system, where they are sent from the SIEM (if they are not part of the same product). Please note that on this ticketing platform not all alerts will have the SIEM as a source, as they can also arrive from other sources: from a user phone call to a manually introduced ticket on the platform. These alerts are also called actionable events, as they require specific actions to be executed in response to the alert.

Alerts are processed automatically and manually; when analyzed, this analysis can raise an **incident**. The incident concept is not well defined among cyber security researchers [232], although we accept the definition stated in [269], which refers to incidents, or cyber incidents, as any occurrence that has an impact on any of the components of the cyber space or on the functioning of the cyber space, regardless of whether it is natural or human–made, malicious or non–malicious intent, deliberate, accidental or due to incompetence, due to development or due to operational interactions. In this context, an incident can be seen as an alert or set of alerts that can impact in cyber operations. A key objective for a SOC would be a 1:1 ratio between alerts and incidents: as all alerts require an action that can be either manual or automated, those alerts that are not real incidents involve many negative factors, such as economic loss, productivity decrease or analyst burn out.

### 5.2.2 Kill chain models

A kill chain can be defined as a sequence of actions that a hostile actor has to perform in a particular arrangement to achieve their goals. Kill chain models have been developed to describe threat actors' campaign phases [404], as they describe the structure of the intrusion. Kill chain models help analysts to describe phases of intrusions, map adversary kill chain indicators to defender courses of action or identify patterns that link individual intrusions into broader campaigns, among others [284]. Kill chain models have been applied to the detection and understanding of potentially unwanted codes such as remote access tools [276], ransomware features [166] or banking Trojans [332], as well as to protect industrial control systems from advanced threat actors [43] [686] or to model the operations of these actors [51].

The most used kill chain model is the Cyber Kill Chain® framework [284], developed by Lockheed Martin, as a part of the Intelligence Driven Defense® model for identification and prevention of cyber intrusion activity, identifying what a threat actor must complete in order to achieve its objective. It was first described in [284] as a seven–step process suitable for CNA or CNE operations, as shown in figure 5.1.

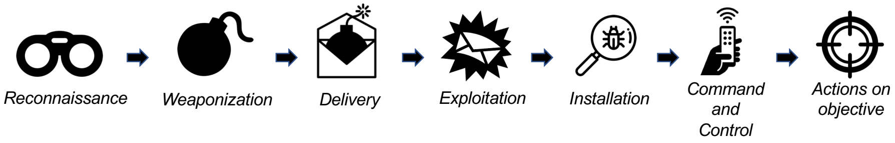

Figure 5.1: Cyber Kill Chain

These seven steps are defined as follows [284] [219]:

1. **Reconnaissance.** Research, identification and selection of targets.
2. **Weaponization.** Before attacking a target, the threat actor has to couple a remote access Trojan with an exploit into a deliverable payload.
3. **Delivery.** Transmission of weapons to the targeted environment to launch a particular operation.
4. **Exploitation.** After the weapon is delivered, exploitation triggers intruders' code.
5. **Installation.** Installation of an implant, just as a remote access Trojan, a backdoor or any kind of malicious software, on the victim system allows the adversary to maintain persistence inside the environment.
6. **Command and Control.** Compromised hosts must beacon outbound to an Internet controller server to establish a C2 channel, thus allowing the threat actor to control its target remotely.

7. **Actions on Objectives.** After progressing through the first six phases, intruders can take actions to achieve their original objectives, such as information theft, denial or hop to a third–party infrastructure.

The Cyber Kill Chain represents an industry-accepted methodology for understanding how an attacker will conduct the activities necessary to cause harm to an organization and has been largely discussed [430] (in [684] the authors identify some of the discussions regarding the application of the Cyber Kill Chain). Some authors [350] [92] have proposed the addition and removal of different stages in order to improve or adjust the original model, and the topic has also been discussed in technical conferences. Also, some efforts to unify models and variants, such as [489], have been made. Despite this, the original proposal has been widely used and applied to specific problems regarding advanced threat actors, such as those related to the modeling of the attack stages against critical structures [253] [112] [686].

Some critics of the CKC are related to its approach as a linear progression that does not accurately represent the actions of an actor; Mandiant (FireEye) presented in [398] the Mandiant Attack Life Cycle, a kill chain model in which the Weaponization stage is removed and introduces a loop to represent the continuous activities of internal recon, lateral movement and persistence performed by a hostile actor, as shown in figure 5.2. A summary of critics and improvements to CKC can be found in [327], and the proposed variants of the cyber kill chain model can be found on [329] or [684].

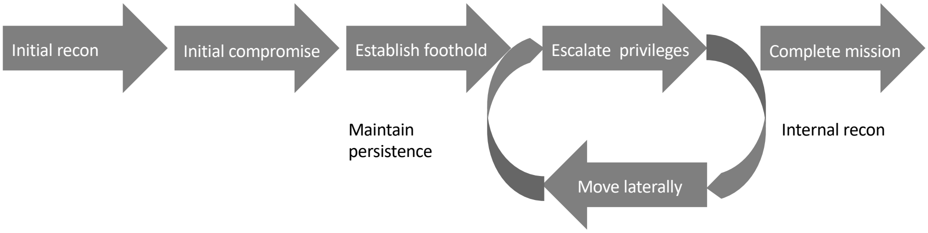

Figure 5.2: Mandiant's Attack Life cycle Model

## 5.3 The Issue: A Lack of Defensive Kill Chain

The main issue we have identified is the lack of a suitable, arranged set of tactics, in the form of a kill chain (this is, in the form of mandatory steps to accomplish a goal). There is not a single high–level, platform–agnostic kill chain, or even identification of tactics to achieve the SOC goals: this is, in order to establish the mandatory steps to detect and neutralize threats. This lack of such a model may be due to the fact that defensive centers focus on the alleged attacker activities as the first and main input, so they start their work always trying to detect, in order to respond, to such activities. With this focus on the incident response against hostile operations, subjects such as the correct data acquisition or processing are not properly covered, thus leaving windows of opportunity for an attacker to succeed.

Many approaches to model a SOC, including its processes perspective, have been proposed and developed in the academic and industrial literature. Nevertheless, as stated before, we have not identified a single high–level, platform–agnostic kill chain or even identification of tactics to achieve the SOC goals, in order to establish the mandatory steps to detect and neutralize threats. Of course, this is considered a relevant problem. The same way hostile operations are modeled in order to improve our knowledge about them and about the threat groups that perform these operations, we consider it mandatory to establish an equivalent model, in the form of a kill chain, for the defensive operations that enable the detection and neutralization of hostile activities. Such a model can improve aspects such as SOC classification, services, capabilities and technologies; however, most importantly, it can help defenders to analyze and to establish the requirements for an effective detection and neutralization of threats, to implement suitable techniques for each tactic and to arrange its activities.

The identification of hostile activities has been largely analyzed; both the particular tactics, techniques and procedures that a hostile actor has to perform in order to achieve its goals, as well as the particular arrangement of these tactics, are now well structured and accepted through the community. Focusing on tactics, techniques and procedures, from an offensive point of view, different approaches have been identified and analyzed, without establishing a particular arrangement, identifying commonly accepted models such as MITRE ATT&CK tactics and techniques matrix. In addition, focusing on the arrangement of these elements, an offensive point of view has been largely developed, such as Cyber Kill Chain. Different kill–chain approaches have been defined and used for the modeling of hostile activities [92].

Surprisingly, this state of the art, while dealing with the execution of offensive operations has no equivalent, commonly accepted approach for defensive operations. We find an important lack of the identification and definition of tactics and techniques for a SOC to run, as well as on its correct arrangement. In other words, there is no model for the mandatory activities a SOC has to perform. A model is defined [76] as an abstract representation of some domain of human experience, used to structure knowledge, to provide a common language for discussing that knowledge, and to perform analyses in that domain. Models are necessary in order to better understand and discuss abstract entities representing and structuring common knowledge and experiences, and allowing analysts to profile attackers from their goals to their TTP.

To successfully perform CND operations we consider that it is necessary to define a global SOC kill chain model regarding both the mandatory tactics to be executed and their arrangement. Therefore, we propose a kill chain model to allow a SOC to detect and neutralize threats. Of course, this model would have to be independent from aspects such as the budget or the technology; this is, no matter which technologies or how much money a defensive center has, the model has to be similar. However, most importantly, it has to be independent from the hostile activities they face, and regardless of the type of threat a SOC is handling, the model must always be the same. Please note that most of the threats a SOC has to deal with will be adversarial, but accidental, structural or even environmental threats must also be detected and neutralized.

In this work we address an issue that has not been largely addressed and that is a must for any defensive center. The lack of a suitable model for CND operations implies not only the absence of a homogeneous work flow across different SOCs, but also heterogeneous approaches to incident detection that cause security flaws and security monitoring deficiencies among SOC customers, resulting in unprotected infrastructure assets and undetected security breaches. This problem is especially important as the first mandatory actions, those related to planning and acquisition of data, are in many cases not specially considered. This situation, with many processes focused on pure incident response, leads SOCs to perform a right processing and even a right analysis among incomplete data, thus resulting on an incorrect monitoring that leads to a late incident detection and response. In addition, they do not face the fact that not all SOC responses have to follow a well–defined incident response process, but in many cases these responses have to be simpler, and so they are in practice.

## 5.4 Techniques and limitations

Defensive centers have been largely analyzed from different high–level perspectives, but none of them establishes a set of common tactics and their particular arrangement to provide a suitable SOC model. Many approaches are focused on the incident response process, which in our opinion is not correct, as this response is the last step of a set of activities a SOC has to successfully perform to achieve its goals. In addition, in some cases models are presented from a generic perspective, without identifying the activities a SOC has to develop in order to successfully detect and neutralize threats.

In [68] the authors propose the people, processes and technologies (PPT) model, later used specifically for a SOC in works such as [504], [638], [368] or [646]; in this triad, the processes branch is not uniform among the literature [646], providing an incomplete picture of the actions performed on the SOC daily basis.

Regarding this process definition for a SOC, many approaches are linked to incident handling processes [461] [427], leaving aspects as the acquisition or processing of data, although referred, in a secondary role. With this focus, critical aspects for incident detection are not considered, thus providing valid models once an incident is identified (from the response point of view) but lacking a unified approach to the previous mandatory tasks. In addition, a key concept is not considered in the approaches focused on incident response: the fact that in a SOC not all response activities are linked to this process. Response actions can vary depending on many factors, both technical (for example, those related to the priority or potential impact of an incident) and non–technical (for example, those related to signed contracts for a particular customer).

In 2013 Forrester introduced [205] the concept of the SOC funnel as a graphical simplification for SOC work. The SOC funnel represents the reduction process that must be performed by a SOC in order to obtain actionable alerts from billions or millions of raw events, as shown in figure 5.3. Although this model is conceptually correct, the authors did not specify the different tactics or the mandatory steps to accomplish this goal. More specific academic works, such as [75], have used this conceptual model, but also without deepening its tactics. Although this is a valid high–level approach, it does not specify what a SOC has to execute to achieve it.

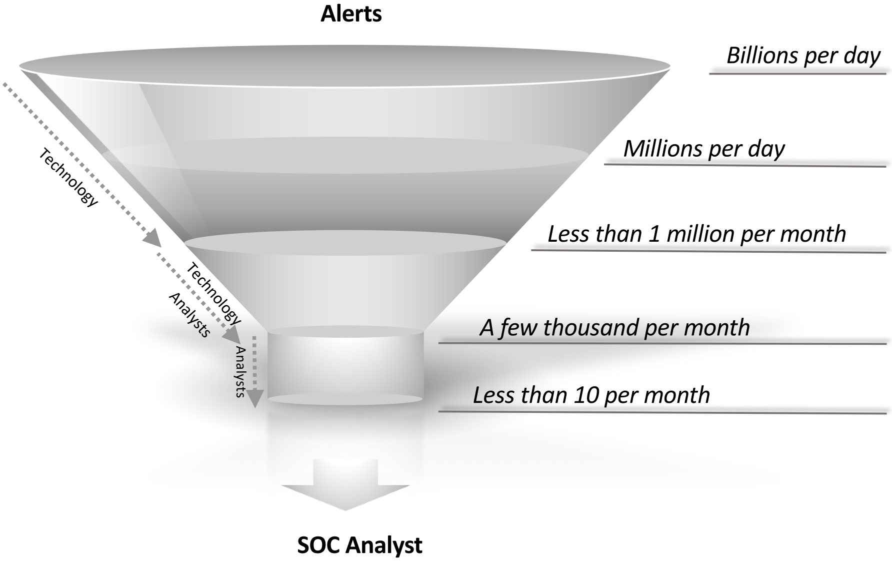

Figure 5.3: SOC Funnel

In [682] the authors proposed a SOC architecture based on four layers: data acquisition, data processing (which includes filtering, merging and formatting), correlation analysis and visualization. As the work is focused on correlation, this model does not approach the tasks after analysis, simplifying them as a global visualization step. We consider visualization as not a key tactic for a SOC, but a technique for human analysis. Especially when dealing with big data environments, situational awareness in any of its forms is an important decision–support mechanism [171] [188] for Computer Network Defense. From an architectural point of view, the layers of a SOC have also been analyzed in works such as [504] or [646].

These approaches identify the mandatory layers for a SOC to run: generation, acquisition, data manipulation and output or presentation layers. Although these layered–approaches can be linked to the tactics the SOC has to execute, they do not consider a real kill chain model but layers inside the SOC architecture, including people, processes and technologies. In addition, they do not represent the mandatory feedback that this chain must have, as the output of the process has to provide enhancements to the initial point.

In 2014 Ryan Stillions wrote a blog post [594] in which he presented the Detection Maturity Level (DML), a model to assess the maturity of an organization to detect cyber attacks in terms of its capabilities to consume and act upon given threat information. The DML model is composed of nine levels of maturity, from the most technical ones –really, level 0 represents no information about the threat– to the highest level ones –goals and strategy of the attacker–. The DML model has been improved [88] [387] by adding a tenth level of abstraction, DML-9, regarding identity of the threat actor, a useful information to connect multiple attacks to the same actor in order to predict strategy, tactics, techniques and procedures expected to be used in an operation. This hierarchical model can provide an approach not only to evaluate the detection capabilities of an organization, but also to the semantic modeling of an advanced threat actor, from a group to specific indicators left after an operation, thus helping the analysts to model the interests and behavior of the actor and its modus operandi in specific operations. However, neither the DML model nor its improvements detail the activities a SOC has to execute in order to provide these detection capabilities.

In [424] the authors presented a process for incident detection in a SOC inside a global incident response timeline. This process is based on four steps to identify the sources responsible for detecting and reporting incidents, the available channels to do so, the steps to accept inputs and, finally, the requirements on people and technology for this process to work. This is a valid approach, although it lacks an established arrangement and we identify a gap between the source identification and how this source could use the defined channels for notification.

In 2017 Matt Swann introduces, at Microsoft BlueHat Israel conference, the concept of a blue team cyber kill chain as a defender–centric version of the standard, offensive, cyber kill chain, as shown in figure 5.4; this approach defines the chain of actions a defender needs to go through to find and evict attackers. The author discusses the proposed stages and its window of opportunity in relation to the offensive kill chain, and it also presents a “response pyramid” which goes from the protected assets inventory to the ability to collaborate with third parties to disrupt campaigns. Although this approach relies mainly on incident response capabilities, it is an interesting starting point for the identification of SOC tactics. However, it has not been improved in other works, so it has not evolved since 2017; for example, the last step, contain, would have to be adapted to a more generic response approach.

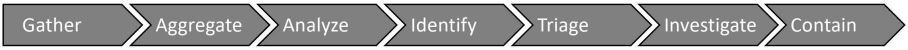

Figure 5.4: Blue Team Cyber Kill Chain

MITRE Shield is an active defense knowledge base that captures and organizes what they are learning about active defense and adversary engagement. While MITRE ATT&CK provides the attacker's tactics and techniques, MITRE Shield provides tactics and techniques available to defenders. These TTP are linked to the active defense concept, understood as [452] the employment of limited offensive action and counterattacks to deny a contested area or position to the enemy. MITRE Shield focuses on the tactics to respond to a hostile operation once it has been detected, not on the previous, mandatory steps to detect this operation. In this sense, we miss in the framework the identification of tactics for a SOC to monitor infrastructure and raise alerts.

Finally, as we have stated before, Mandiant Attack Life cycle is a relevant approach and reference for our work. It focuses on the detection of the activities of a hostile actor against an infrastructure. This approach, although useful for defensive teams (it focuses on what can be detected) is not accurate for the global modeling of threat actors; it is again focused on activities performed by the attacker, and so has been used in many works in both IT [227] [407] and OT environments [149], [264]. Nevertheless, it does not focus on what the defensive team must do to detect an operation.

None of the identified approaches among literature defines a formal arrangement of tactics, in the form of a kill chain model, to be performed by a SOC. This is, as we have stated before, no defensive kill chain model has been ever proposed. Most of the analyzed proposals focus on pure incident handling methodologies, which should be considered particular response techniques not suitable for all alerts a SOC handles daily. Some of them do not even propose tactics to be executed, but only an abstract model without delving into what activities a SOC has to perform. The closest approach we have found is Matt Swann's blue team cyber kill chain, which has not evolved in last few years and, as we will defend in section 5.6 lacks mandatory tactics and also focuses on incident handling, not on the whole cycle a SOC must follow to achieve its goals.

## 5.5 Our proposal

We define the SOC Critical Path (SCP) as the sequence of mandatory actions to detect and enable the neutralization, of a threat, as shown in figure 5.5. We can see SCP as a defensive version of a cyber kill chain. The SCP starts with the generation of a log record in a particular system and ends with the appropriate response to a detected incident. This incident can be raised by any threat, not only by adversarial ones.

As stated before, in our approach we propose the SCP steps as shown in figure 5.5. These steps are partially equivalent to those defined in the Intelligence Cycle [454] and we have adopted a similar nomenclature for them where applicable. Each of these steps represents a tactic that specifies what the SOC must perform to achieve its goals, and each of them comprises different techniques to achieve the

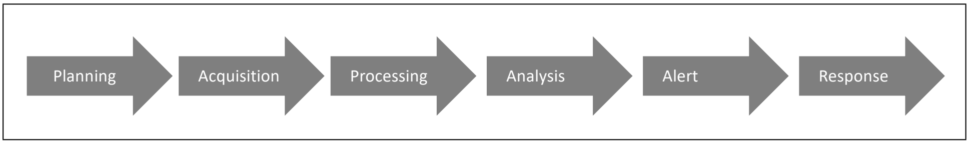

Figure 5.5: SOC Critical Path

particular tactic. Please note that the proposed arrangement is mandatory: it is not possible to raise an alert if there is no data analysis, there is no data analysis without basic or complex processing, and there is no processing if we are not able to acquire relevant data. In addition, for some of the identified tactics, we have proposed specific sub–tactics; these sub–tactics are not mandatory in all cases, but they are recommended and, in today's SOCs, their execution is a common approach. We have not gone into them in depth, as their details are outside of the scope of this work.

The SCP starts with the planning tactic, which will identify what needs to be monitored for alert raising and how these data must be processed and analyzed in order to detect and respond to security incidents. In this step, it is mandatory to analyze potential hostile operations and techniques against our protected assets, to understand how threat actors will try to damage us, to identify attack surfaces in our infrastructures, and to establish guidelines on what and how monitoring will be performed. The planning tactic, as the first step of our approach, will guide all the activities executed by a SOC to detect and neutralize threats, and it will receive feedback from the rest of the tactics of our model, especially from the Response one.

Once the global monitoring and response strategies are defined, data Acquisition is the next SCP step, where relevant data is acquired from the monitored infrastructure and sent to a central repository, typically the SIEM platform. Monitored infrastructure generates logs, and some or all of them are selected for detection purposes. In order to detect malicious activities, it is mandatory not only to send those logs regarding special actions, but also those related to usual activities on the monitored infrastructure. In this tactic we identify two particular sub–tactics: Extraction, regarding what information is mandatory to acquire, and Transportation, regarding what mechanisms must be used to send the acquired data from its data source to the central repository.

When received by the SIEM, logs are processed and converted into events. This processing includes, at least, some kind of standardization and the storage and retention of the normalized data in the SIEM. Depending on the SIEM technology, it can range from simple to complex processing mechanisms, but in any case it is a must, as without proper processing, the events cannot be analyzed, which is the next step of the SCP. Most technologies include a normalized format for information (logs) received from multiple, hybrid sources, such as firewalls, endpoints or proxies, as well as predefined retention capabilities that can range from a few hours to months or years, with deletion and the cold storage being the last step of the data processing. Please note that, in this context, the deletion of data does not necessarily mean its real removal, but its elimination from a warm site and its storage in a cold site that cannot be exploited for detection but for forensic purposes. In this tactic we identify three sub–tactics: Reception, regarding the mechanisms that enable the correct receiving of the data, Normalization, regarding what approach is followed to convert hybrid data to a common format, and Storage, regarding what strategies the SOC must follow to store the received and normalized data in a way that allows the next step of the SCP, the Analysis.

Once processed, SIEM technologies also provide the capability to analyze events, which is the next step in SCP. This analysis can be performed both automatically and manually. SIEM can usually perform automatic reduction, aggregation and correlation. They also provide specific capabilities for manual analyses, ranging from simple queries to the stored data to particular, platform–dependent languages. As in classical intrusion detection schemes, the goal of this analysis is to identify misuses and anomalies that can lead to an incident, so in the Analysis tactic we identify two mandatory sub–tactics: misuse analysis and anomaly analysis. Please note that the particular techniques that compose these sub–tactics are those regarding intrusion detection approaches, such as expert systems, neural networks or statistical anomaly detection.

When a misuse or an anomaly is detected, an alert is raised, usually on a ticketing platform. The alert generation can be automatic, from pre–defined use cases in the SIEM platform, but also manually, when an analyst defines a new hypothesis, customizes it and identifies a misuse or an anomaly not known before. Although alerting could be included in a global analysis tactic, we consider it apart, as for us it is not pure analysis but its result; it is equivalent to the dissemination step of the intelligence cycle stated before. In addition, please note that on the ticketing platform, related to these manual alerts, we can manage both alerts from SIEM data but also alerts from particular situations outside the SIEM scope, such as user calls. In this platform, the SOC must centralize all incident–related actionable events. We do not identify any sub–tactics for the Alert tactic, as it is a simple one, but only particular techniques to achieve it in an effective way, such as alert numbering schemes or alert data enrichment.

Through the ticketing platform, analysts respond to the alert that has been raised. All alerts require an appropriate response that can be automatic or manual, and given that all alerts on the ticketing platform require a response, a SOC objective is that all of them are real incidents. This response starts with the identification of the incident and, in case it is a real one (true positive), it continues with specific actions that are performed in order to neutralize the threat. Inside the incident response process these actions can range from deception to containment, and from simple actions to complex methodologies. The incident response process is well defined among many methodologies [1] [134]; a summary can be found in [614]. Many works define the particular sub–tactics, after the incident identification, such as containment, eradication, recovery and lessons learned [35], with little variation from this approach [2]. In any case, as we have stated before, SOC particular response activities depend on multiple factors, so these sub–tactics do not always apply; for example, a particular agreement with a customer may define the response to a specific incident type just as the notification of the action, or by automatic network block without further investigation or activities. For this reason, these sub–tactics are outside of the scope of this work, and cannot be considered mandatory in all cases.

To perform the response tactic, analysts may have to acquire and contextualize data from different sources, including their own ticketing platform, the SIEM, particular relevant systems of the organization or third–party repositories. If the incident is not a real one (false positive), the response tactic will not consider all the steps related to pure incident response, but the SOC will close the related ticket; in these cases, that a SOC has to minimize, it will also be an improvement to the process by refining correlation rules, use cases or acquisition exceptions.

Finally, please note that as in many processes SCP will improve global detection and response capabilities by giving appropriate feedback to the planning stage from all of the steps of the SCP, which will be produced by the analysis of all of the tactics role in detecting and neutralizing a threat, from the acquisition to the response one; in the same way, the improvements achieved will be applicable to all the tactics of the SCP.

### 5.5.1 A practical example

To provide a practical result for our proposal, we have analyzed a particular technique performed by threat actors in their operations and how our model helps a SOC to detect and neutralize it. We have chosen the Command and Control (C2) tactic stated by MITRE ATT&CK, in particular the T1071.001 technique, related to command and control through web traffic protocols. While using this technique, adversaries may communicate using application layer protocols associated with web traffic to avoid detection and network filtering by blending in with existing traffic. Commands to the remote system, and often the results of those commands, will be embedded within the protocol traffic between the client and server, as MITRE ATT&CK framework states. Please note that choosing any other tactic or particular technique for our example would follow the same structure and provide the same results.

In order to be able to detect this C2 technique, first of all the SOC team must analyze how threat actors perform it. As stated before, they communicate with the command and control system through web traffic, this is, through the legitimate navigation of the users inside the compromised infrastructure. In this way, a malicious HTTP/S hit is hidden inside the whole, legitimate web traffic. As the amount of data to be analyzed in order to detect this malicious hit is very large, this technique allow the threat actor to go easily unnoticed. In addition, this technique evades network traffic filtering mechanisms, as outgoing web navigation protocols are usually allowed in all organizations.

Once the technique is analyzed, following our proposed kill chain, the SOC team must approach the first step: Planning. This is when the SOC plans the rest of the relevant tactics and techniques that will be later executed, so defining all the mandatory tasks for all of the SCP steps. First of all, the SOC must plan how to acquire relevant information for the detection and neutralization of the threat actor. In this case, as we are dealing with HTTP/S command and control, the main data source will be this kind of traffic, which can be acquired by different techniques, such as passive acquisition (for example, sniffing network traffic and dissecting the particular protocols) or through the parsing of proxy logs. The SOC team must analyze the pros and cons of the different techniques that can be applied to the acquisition, in order to identify and put into practice the most suitable approximation. This planning not only has to consider technical aspects, but also economic, legal or human ones.

In relation to the processing, the SOC team must consider elements such as the storage of the acquired data, its retention period and, most importantly, the usability and agility of the SIEM platform where this data is processed. As it is mandatory to process all navigation traffic, not just specific alerts such as blacklist navigation attempts, the parameters to estimate the processing capabilities, such as the amount of needed storage, the computing power or the events per second rate, must be adjusted in order to provide SOC analysts a suitable environment for their work. The processing requirements will depend on the acquired data, so the proper arrangement of the proposed tactics we are defending is mandatory.

The next step in our model is related to the analysis of the data in the SIEM platform. This analysis can be performed through different techniques, both automated and manual, and its goal is to identify misuses and anomalies that can lead to an incident. Following our model, the SOC team will identify and put into practice analysis techniques suitable for the detection of web command and control channels, such as the use of black lists (misuse) or statistical or knowledge based anomalies, the so called “hunts” in threat hunting terminology. Please note that without a suitable processing, analysis can not be accomplished, confirming once again the mandatory arrangement for the tactics of the SCP.

Once data has been analyzed and particular conditions are met, the SOC will raise an alert if a suspicious activity has been detected. This alert will be sent to a ticketing platform, global for the SOC incident management processes, and it is mandatory to define how the alert will be generated in this platform, taking into account parameters such as classification, criticality or service level agreements for each type of alert.

Finally, when an alert is generated, some action from the SOC is required, reaching the last tactic of the SCP: the Response. The SOC team will respond, in one way or another, to the potential incident. In this last step, the SOC will follow the established operating procedures for each particular case. These procedures can range from a simple notification to the affected organization to a whole incident response deployment, following in this last case the usual sub–techniques: identification, containment, eradication and recovery. Again, the alerting, and thus the response tactics, must be executed after the analysis of the acquired and processed data, so we must stress one more time the relevance of the arrangement of the proposed tactics.

As we have stated, each of the tactics of the SOC Critical Path will provide feedback to the planning step, starting again the SCP in order to improve over time the SOC detection and response capabilities. In our example, this feedback is especially based on the analysis of the incidents that have not been properly detected, identifying which tactic or tactics have not been fully accomplished and improving them. This continuous improvement is mandatory for a SOC, as hostile actors modify their techniques over time and their detection and neutralization will always be the SOC's goal.

In this practical example, we have followed our proposed model to provide a suitable detection and response to a particular technique inside a hostile operation. We have identified particular tasks inside each step and justified the arrangement of our proposed tactics, thus providing SOC teams an example on how to apply the SCP to a real detection case. In order to present the improvements of our approach over previous techniques, we can compare the SCP to an equivalent kill chain model. But as we have stated in this work, it is hard to provide this comparison, as no direct, equivalent model has been identified during our research. SOC Critical Path is a novel proposal, being Matt Swann's blue team cyber kill chain the closer approach, but not having a direct equivalence.

Comparing our model to the blue team cyber kill chain, in Swann's work we find in first place the Gather step, which is equivalent to our acquisition, but without a proper, previous planning. This absence of a Planning tactic as a primary mandatory task can lead the SOC to execute subsequent steps lacking a clear, defined goal. Without identifying the relevant data sources that enable the identification of a particular offensive technique, without planning which processing requirements are mandatory and without a proper identification of analysis techniques, it is not possible for a SOC to provide an accurate detection capability.

The Blue Team Cyber Kill Chain's next step is the Aggregation one, a phase in which the defensive team joins the gathered data from multiple sources. In our model, we consider it a specific step for processing in which aggregation is set. But in opposition to the blue team cyber kill chain, inside the SCP Processing tactic, aggregation is just one of the particular techniques that a SOC can consider, together with other approaches just as reduction or, simply, storage of the acquired data. A processing tactic is much more suitable for a kill chain model, and inside this tactic, aggregation is a specific technique, but not the only one.

Next, the blue team cyber kill chain defines analysis and identification, which are equivalent to the analysis and alert steps in our model. The rest of the blue team cyber kill chain defined steps are triage, investigation and containment. We consider these tactics incorrect for a general model, as they focus on particular tasks not performed by a SOC in all cases. Triage and investigation can be considered particular techniques for our response tactic. If this response is a global incident response, these steps should be included in a general identification step, as they define tasks to perform this identification. In addition, containment is a tactic suitable for this global incident response, but in this case it should be considered together not only with identification, but also with eradication and recovery, as most incident response strategies define [306] [135].

Being hard to compare the SOC Critical Path to an equivalent model, as we have presented a novel approach, differences and benefits over the blue team cyber kill chain are clear. Our proposal provides a technology–agnostic arrangement of tactics for a defensive team to detect and respond to threats. Comparing this arrangement to the blue team cyber kill chain, we advocate that our model not only covers the whole cycle for a SOC to perform its task, but provides a homogeneous point of view of the mandatory tactics without considering particular techniques for any of them. This abstraction allows analysts to follow our model and, when needed, to go down into the techniques to perform their task. We consider this fact increases global detection and response performance, as the difference between what the SOC has to execute (the tactic) and how it has to be executed (the technique) is clear at all times. Also, giving feedback to the Planning tactic, which in the SCP is mandatory to enhance all the SOC work, closes the cycle and provides a continuous improvement element to our model.

## 5.6 Discussion

We have proposed a model for the SOC Critical Path, the sequence of mandatory actions to detect and enable the neutralization, of a threat. This model provides the arrangement of tactics that a SOC must perform to achieve its goals. In this sense, our approach gives analysts not only a kill chain equivalent for defensive cyber operations, but also a set of mandatory tactics to be considered in a center that provides cyber defense capabilities. Each of the defined tactics can be more or less complex, depending on the SOC maturity, but we defend all of them to be mandatory for a SOC to accomplish its goals. In addition, please note that this model is suitable not only for the detection and neutralization of adversarial threats, but is a common approach for all types of threats.

For the proposed tactics of the SCP we have chosen terms close to those used in the Intelligence Cycle, as we have stated before. We defend that SOC activities are in fact related to counter intelligence activities, as the SOC goal is the prevention, detection and neutralization of threats. Of course, these terms could be largely discussed (for example, the Planning tactic could be called the Preparation tactic), but we consider this is pure nomenclature. We have actively avoided terms used by the Cyber Kill Chain® as we consider it especially relevant to distinguish both approaches, one with an offensive perspective and the SCP with a defensive one.

As we have stated, all reviewed approaches for kill chain models are focused on the offensive point of view; these models provide the mandatory, arranged tactics for an attacker to achieve their goals. However, unlike in the offensive perspective, previous defensive proposals do not focus on tactics and techniques, neither on a critical path (kill chain in offensive jargon) to achieve a defensive actor's goals. Different studies analyze SOC processes with special emphasis on the incident response capabilities, thus not giving the mandatory importance to tactics to provide analysis or acquisition capabilities, and also considering a homogeneous, not real, work for a SOC day to day. As we defend, a global incident response is only a particular kind of response of all of the possibilities a SOC can provide to its customers.

The absence of a defensive kill chain model makes it hard to compare our approach with equivalent proposals. In fact, this absence is a real problem, as SOC activities are in many cases unstructured, not correctly formalized nor arranged, making it difficult to improve capabilities over time and thus giving adversaries relevant advantages to succeed. Different proposals have been analyzed in this work, none being suitable for the definition of mandatory tactics and their arrangement in the form of a defensive kill chain. The closest approach we have found, the blue team cyber kill chain, lacks important tactics and, as most of the frameworks, focuses on global incident response without taking into account relevant aspects such as the planning or proper acquisition of data, as we have stated in the practical example shown in section 5.5.1. For this reason we have proposed a novel kill chain model which sets all the relevant steps and their correct arrangement for a successful detection capability.

Future research lines would include to deepening into each of the presented tactics, in order to define sub–tactics inside the defined tactics. Those sub–tactics would not be mandatory to accomplish the global SOC goal, but recommended. In the cases where these sub–tactics can be identified in a specific order, they could be considered a more specific approach to the SCP. In our approach, we have provided sub–tactics for the main identified tactics, but we consider this to be an ongoing work. It is also important to note that in some cases the internals of some of the proposed tactics are well structured in the literature and have been discussed in our work, for example regarding the response tactic, while in other cases the potential sub–tactics are not so well structured, such as in the acquisition one.

In addition, an interesting future research line would be the specification of techniques for each tactic we have presented in this work. Although we provide a high–level description for SCP, we consider our research as the first proposal that has to be improved. These techniques would define how a specific tactic can be executed, and its structure could be similar to the MITRE ATT&CK approach, but considering the defensive point of view.

## 5.7 Conclusions

In this work we have presented the SOC Critical Path (SCP), a sequence of mandatory actions to appropriately detect and respond to security incidents, which is the main goal of Computer Network Defense, executed by centers such as a SOC. The SCP is equivalent, from a defensive point of view, to models such as the Cyber Kill Chain, which are focused on the attacker's perspective. Although this offensive point of view has been discussed in many works, the defensive one presents an important lack of research, so we have addressed an issue not largely analyzed in spite of being a must for a SOC to accomplish its goals: to prevent but, especially, to detect and to neutralize security threats. In this sense, our approach provides a kill chain equivalent to a SOC team.

We have proposed an approach based on the definition and arrangement of the tactics that must be implemented in any defensive center. Our goal was to identify these tactics from a global perspective, from the generation and registration of interesting activities in a protected IT asset to the analysis of data, alerting and final incident response. Although many studies focus on this later step, the one related to the incident response, we have tried to consider all mandatory tactics to achieve the SOC goals at the same importance level. For each identified tactic we have discussed and proposed sub–tactics that, in a particular arrangement, conform to the global tactic to achieve.

Finally, we identified some future research lines mainly related to the specification of sub–tactics and particular techniques in each of the identified tactics in our research. We consider this future work as a mandatory enhancement of this first approach to the process of detecting and neutralizing security threats.

# Chapter 6

# Key Requirements for the Detection and Sharing of Behavioral Indicators of Compromise

*Related to the detection of Advanced Threat Actors, one of the main problems we must face nowadays is the detection of their techniques. These techniques, in the form of behavioral indicators of compromise, are harder to modify for the actor, so being able to detect them provides organizations valuable intelligence for the detection. In this chapter we analyze the main problems of the detection and sharing of these indicators, identifying the key requirements to face these problems.*

This chapter was published as: Villalón-Huerta, A.; Ripoll-Ripoll, I.; Marco-Gisbert, H. Key Requirements for the Detection and Sharing of Behavioral Indicators of Compromise. Electronics. 2022, 11, 416. http://doi.org/10.3390/electronics11030416

## Contents

Cyber Threat Intelligence feeds focus on atomic and computed indicators of compromise. These indicators are the main source of tactical cyber intelligence most organizations benefit from. They are expressed in machine–readable formats and they are easily loaded into security devices in order to protect infrastructures. However, their usefulness is very limited, specially in terms of time of life. These indicators can be useful when dealing with non advanced actors, but they are easily avoided by advanced ones. To detect advanced actor's activities an analyst must deal with behavioral indicators of compromise, which represent tactics, techniques and procedures but that are not so common as the atomic and computed ones. In this paper we analyze why these indicators are not widely used and we identify key requirements for a successful behavioral IOC detection, specification and sharing. We follow the intelligence cycle as the arranged sequence of steps for a defensive team to work, thus providing a common reference for these teams to identify gaps in their capabilities.

## 6.1 Introduction

Indicators of Compromise (IOC) are a key piece in Cyber Threat Intelligence (CTI), as they enable and speed the detection of malicious activities in technological infrastructures. They allow to specify both the usage of technological capabilities, such as tools or artifacts, and the tactics, techniques and procedures (TTP) developed by threat actors. However, this last use case, the specification of TTP, is not extended among threat intelligence providers. These providers focus on the sharing of basic indicators, which provide an immediate result when loaded into security platforms but which present an important problem: their time of life. As they are easily modified by hostile actors, their usefulness is limited. In other words, most of the IOC shared today in threat intelligence sharing platforms are not the best ones, but the easiest to use ones.

We need cyber threat intelligence sharing to detect and respond to hostile actors activities. We usually get indicators that allow us to achieve this goal in two main fronts: compromised hosts, with indicators such as hashes, filenames or mutexes, and networks, with indicators such as IP addresses or domain names. In fact, it is usual to differentiate indicators based on where they are seen [11]: network and host–based ones. However, apart from the problem of false positives with these atomic and computed indicators of compromise [510], those simple IOC, as we have stated before, have a limited usefulness. For this reason, we must focus on the effective detection and sharing of behavioral IOC to face advanced threats, as these indicators are harder for a hostile actor to modify.

In this work we analyze this situation and we identify the key requirements for the effective detection and sharing of behavioral indicators of compromise; these kind of IOC represent the tactics, techniques and procedures of threat actors, and their value is much higher than the one of the basic indicators. By not exploiting and sharing them, defensive teams present an important gap in the detection of malicious activities.

The contributions of this paper are as follows:

- To analyze the problem of specification, detection and sharing of behavioral indicators of compromise.
- To extract the key features of cyber operations from advanced threat actors.
- To identify and structure the key requirements, from an intelligence perspective, for the detection and sharing of behavioral indicators of compromise.
- To identify current efforts to fulfill those requirements and the gap to achieve them.

The rest of the paper is organized as follows. The background, Section 6.2, provides concepts regarding indicators of compromise and the Intelligence Cycle. In Section 6.3 we assess the problem of the detection and sharing of indicators of compromise. Section 6.4 analyzes the different identified approaches to tactics, techniques and procedures specification and sharing. In Section 6.5 we identify the key requirements for this specification and sharing among actors and we discuss those requirements and the current status in Section 6.6, where future research lines are also identified. Finally, Section 6.7 highlights the main results of our work.

## 6.2 Background

### 6.2.1 Indicators of compromise

In CTI, an Indicator of Compromise is defined [259] as a piece of information that can be used to identify a potentially compromised system. This piece of information can range from a simple IP address to a complex set of tactics, techniques and procedures. In all cases, this information meets the definition of IOC: it can be used to identify a potentially compromised system.

Most researches [521] [441] [579] follow the classification of IOC stated by [141] and [284]. This classification defines the following three categories for IOC, based on their complexity and related to the granularity of data represented by them:

- Atomic. Atomic indicators are those which cannot be broken down into smaller parts and retain their meaning in the context of an intrusion. Examples of atomic indicators include IP addresses and domain names.
- Computed. Computed indicators are those which are derived from data involved in an incident. Examples of computed indicators include hash values and regular expressions.
- Behavioral. Behavioral indicators are collections of computed and atomic indicators, often subject to qualification by quantity and possibly combinatorial logic. An example of a complex behavioral indicator could be repeated social engineering attempts of a specific style via email against low–level employees to gain a foothold in the network, followed by unauthorised remote desktop connections to other computers on the network delivering specific malware [521]; a simpler example could be a document file creating an executable object. Such indicators are captured as Tactics, Techniques and Procedures, representing the *modus operandi* of the attacker [579].

While behavioral indicators of compromise are related to operational threat intelligence, atomic and computed ones are related to tactical threat intelligence. All those indicators are relevant to detect compromises, but tactical intelligence has a shorter time of life than operational intelligence, and it can also be more easily evaded, so in general terms it is less useful. In figure 6.1 the relationship between indicators of compromise and intelligence levels is shown.

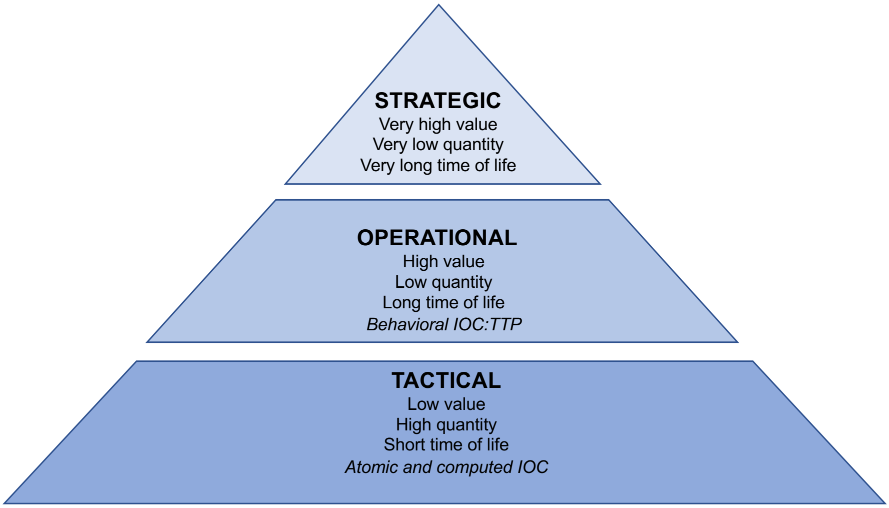

Figure 6.1: Indicators of Compromise and intelligence levels.

Although being less useful than behavioral ones, atomic and computed indicators of compromise are considered by many organizations the most valuable pieces of threat intelligence [90]. The main justification for this perception is related to the fact that the indicators representing tactical intelligence are usually expressed in machine–readable formats, so they can be easily loaded into security devices, providing an immediate result. On the other hand, operational or strategic intelligence feeds in most cases require a manual processing.

In today's interconnected world, it is not possible to deal with security in an isolated box; incidents are not unique, and organizations conform a security ecosystem with common threats, vulnerabilities, risks and capabilities. So in order to enhance one's own security it is mandatory to share cyber threat intelligence with other parties such as private companies, interest groups or law enforcement agencies. And of course, inside this scheme, IOC are a key piece: they are in fact the most shared type of threat intelligence. However, the major part of available shared data regards atomic and computed indicators [294] [631] [322] such as IP addresses, file hashes or domain names. Data related to higher level threat intelligence, the most useful one, is by far less shared, so it is less used. In fact, when dealing with indicators of compromise, it is usual to refer only to atomic and computed ones [671].

### 6.2.2 Intelligence cycle

NATO [454] defines intelligence as the product resulting from the directed collection and processing of information regarding the environment and the capabilities and intentions of actors, in order to identify threats and offer opportunities for exploitation by decision–makers; the same work also defines the intelligence cycle as the sequence of activities whereby information is obtained, assembled, converted into intelligence and made available to users. Although there are different versions of this cycle, we can summarize them in the following five steps:

- Direction. Determination of intelligence requirements, planning the collection effort, issuance of orders and requests to collection agencies and maintenance of a continuous check on the productivity of such agencies.
- Collection. The exploitation of sources by collection agencies and the delivery of the information obtained to the appropriate processing unit for use in the production of intelligence.
- Processing. The conversion of information into usable data suitable for analysis.
- Analysis. Tasks related to integration, evaluation or interpretation of information to turn it into intelligence: a contextualized, coherent whole.
- Dissemination. The timely conveyance of intelligence, in an appropriate form and by any suitable means, to those who need it.

A sixth step for the intelligence cycle [449] would be the evaluation and feedback: all steps, but specially the dissemination of the final product, feed a new iteration of the cycle, as shown in figure 6.2.

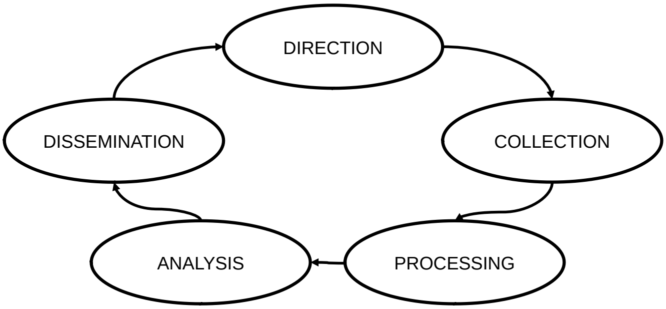

Figure 6.2: Intelligence Cycle.

Intelligence as a product is regarded as the final result of a set of actions, which are sequentially launched; this set of actions, the intelligence cycle, is a simple explanation of a complex intelligence process. It starts when someone, such as an authority or a government, has particular information needs in order to make the best decision about a subject. At this point the cycle starts, identifying the requirements and planning the acquisition of the information that will be later processed and analyzed, in order to generate intelligence.

Once planned, the next stage is to acquire information, and this acquisition can be performed through different intelligence collection disciplines [80] commonly referred as "the INTs": SIGINT (Signals Intelligence), OSINT (Open Sources Intelligence), MASINT (Measurements and Signature Intelligence), HUMINT (Human Intelligence) and GEOINT (Geospatial Intelligence). The essential elements of these INTs are not formally defined [139], neither are them consensuated between authors, but they define the families of sources the information can be gathered from: for example, a simple public website, a satellite, an intercepted artifact or a mole.

With the information gathered, processing and exploitation turn the information previously collected into a form suitable for the production of finished intelligence [520]; this stage includes tasks such as decryption, translation or data conversion and, as a part of the cycle, it is mandatory to the next one: analysis, in which the intelligence, the final product, is generated. This analysis must include the information gathered and processed no matter which collection discipline it comes from. In this sense, we can refer to all–source intelligence, defined [38] as the intelligence products, organizations and activities that incorporate all sources of information and intelligence, including open–source information, in the production of intelligence.

Finally, once the intelligence as a product has been generated, it is delivered to the customer , the entity which had the information needs stated before, in a suitable form for its use and by a variety of means. This product will be used to help the decision making process and, possibly, to start a new iteration of the intelligence cycle.

## 6.3 The issue

### 6.3.1 Threat specification

In CTI, many efforts have been made in order to characterize threats, campaigns and particular attacks by indicators of compromise, specially to establish a structured and standardized information sharing scheme between actors. Back in 2007 Internet Engineering Task Force (IETF) defined [165] IODEF (Incident Object Description Exchange Format), an XML data representation that provides a framework for sharing information commonly exchanged about computer security incidents. Although it is fairly static [100], during these years IODEF has been extended for different needs, such as the reporting of phishing events [104].

Private companies have also developed well–known standards to enable threat information sharing. Mandiant's OpenIOC is an extensible XML schema designed to describe the technical characteristics of evidences of compromise [388]. It provides indicators about files (such as full path, imports and exports, or compile time), hosts and networks (such as DNS or URI), processes (such as handles or paths), registry entries (such as names or text), services (such as name or DLL) and signatures (such as Snort or Yara), among others.

Structured Threat Information Expression (STIX) is a language and serialization format used to exchange cyber threat intelligence whose goal is to identify and represent all the elements of cyber threats in a flexible, automatable and human-readable way. Upon a standardized language, in XML format, the standard provides a common mechanism for addressing structured cyber threat information improving consistency, efficiency, interoperability, and overall situational awareness into a unified architecture that structures and links all those elements of a threat, from its lower level, observables or indicators, to the higher one, campaigns and actors [56]. STIX was defined in 2012, sponsored by US Department of Homeland Security, and in 2015 all the intellectual property and trademarks associated with STIX were licensed to OASIS, a nonprofit organization focused on the development and integration of open technological standards.

From its release 2.0, STIX integrates Cyber Observable eXpression (CybOX), a structured language for cyber observables also developed by MITRE. In STIX 2.1, the latest version at the time of this writing, the standard defines three types of core objects to represent cyber threat intelligence; one of them, SCO (STIX Cyber–observable Object), is used to characterize host–based and network–based information. SCO has been introduced in STIX 2; in previous versions cyber–observables could only exist as objects within an Observed Data object. SCO represents observed facts about a network or host that may be used and related to higher level intelligence to form a more complete understanding of the threat landscape. STIX 2.1 defines the Cyber–observable objects shown in table 6.1, each of them with its corresponding properties.

Although both of them are still available and extensions can be done, Ope-

Table 6.1: STIX 2.1 Cyber–observable objects.

| | | |
|---|---|---|
| Artifact | Autonomous System (AS) | Directory |
| Domain name | Email Address | Email Message |
| File | IPv4 Address | IPv6 Address |
| MAC Address | Mutex | Network Traffic |
| Process | Software | URL |
| User Account | Windows Registry Key | X.509 Certificate |

nIOC and IODEF are today considered legacy formats [507]. STIX has been widely adopted, being its particular applications explored in different works, from malware detection [409] to critical infrastructures and industrial control systems protection [3] [619]. It has been extended to detect more complex patterns [628], and it can be used to provide improvements to other formats, being able to embed not only IODEF extensions but also other formats such as OpenIOC or Yara rules [616] [628]. In this way, STIX can be considered the most accepted CTI standard among security community and the *de facto* one for describing threat intelligence data [550] [237]. The European Union Agency for Cybersecurity (ENISA) [556] has recommended European Union states to implement STIX as a globally accepted standard.

### 6.3.2 Real–world IOC

Despite all the efforts for the characterization of threats exposed in the previous section, the most shared indicators of compromise are still the simplest ones. STIX is a complex standard but it is mainly used to share atomic and computed indicators. The identification of the most used and shared types of indicators of compromise is a key question, as they will be the ones that hostile actors will try to evade in first place. IP addresses and file hashes are considered the most shared indicators of compromise in current literature [518]. Other authors progressively expand this list to domain names [294], URL [236] or malware signatures [631]. [617] also includes in this list file names, dynamic link libraries, registry keys, e–mail addresses, message objects and attachments or links inside a message.

We have processed all the information (available events) from some private MISP (Malware Information Sharing Platform) instances used in public and private Security Operations Centers. Hashes such as MD5, SHA1 or SHA256 represent the most used type of indicator (23,23%), followed by IP addresses (21,10%) and domains and hostnames (19,75%). All the other types of indicators analyzed represent the remaining 35,92% and their presence is far from these numbers (as an example, mutexes represent only 0,03% of global indicator types). These results are consistent with the hypothesis stated in [228] about the types of indicators generally available, and of course they are aligned with our own experience.

Hashes are mainly linked to implants, while IP addresses and domain names are linked to Command and Control (C2) or exfiltration servers. This means that, theoretically, only with these kind of indicators we could detect most activities on the persistence stage of an attack. In spite of the fact that most of the shared indicators among the community belong to these three classes, this security perception is not real. Being all of them easily changed by a hostile actor, the discovery of hashes, IP addresses and domain names causes little pain in the attacker, as the "Pyramid of Pain" [71] states. Any hostile actor who wants to evade detection will defeat, at least, these three types of IOC, as they are the most used ones. If an actor is able to cheat these indicators, it will be able to evade approximately 65% of the defender detection capabilities.

As main IOC types are almost useless when facing advanced threats, analysts have to look for alternatives to detect intrusions. Different approaches to provide other atomic and computed indicators of compromise have been developed, but they are not commonly shared among intelligence groups, so their usage is not as extended as it should be. In many cases this is mainly due to the lack of automatic tools to load intelligence and to get an immediate result. For example, to identify minor changes between objects, which of course produce different hashes, fuzzy hashes [470] [213] have been used. Through algorithms such as ssdeep or sdhash, this approach is able to detect similarities between files, helping the analyst to identify those changing objects. However, in any case this similarity hashing provides just another computed indicator that again can be easily evaded by an advanced actor.

To provide a more accurate detection capability, CTI must deal with the detection and sharing of behavioral indicators. This approach would allow analysts to detect the tactics and techniques of attackers, no matter which atomic or computed indicators they use in a particular campaign (this is, no matter which hash, which IP address or which domain name). However, CTI sharing is mainly focused on those simple indicators of compromise, that are easily evaded. In [25] the authors state that hashes, IP addresses and domain names are the easiest indicators to trace, to identify and to exploit quickly. As all of them can be easily expressed in machine–readable formats, for example a simple blacklist, many security devices can be configured to load them automatically, so they provide an immediate result. In fact, such types of indicators focus on immediacy [313], while upper ones, those related to goals or TTP, provide a richer analysis but with a longer time to process.

## 6.4 Approaches and limitations

The detection of behavioral indicators, in front of atomic or computed ones, has been largely studied in the malware field [297] [365]; the so called behavioral signatures are applied in dynamic malware analysis and rely on the malware behavior to identify patterns through multiple means, from the monitoring of system calls to temporal logic formulae [59]. These methods extend the classical static signature detection, complementing or superseding them. However, if we do not focus on such a specific piece of an attack, the malware, and we try to expand this detection approach to the tactics, techniques and procedures of an attacker, the work gets complicated.

STIX “Attack Pattern” SDO describes the tactics, techniques and procedures that the adversaries develop to compromise targets: this is, just what behavioral indicators represent. This SDO contains textual descriptions of the pattern along with references to external objects; it relates to other SDO, such as “Indicator”, but these ones are, again, pure atomic or computed indicators of compromise. So although STIX allows the specification of tactics, techniques and procedures, it does not provide a common vocabulary for describing TTP, making STIX not suitable for a machine–readable specification of behavioral indicators of compromise.

In [283] the authors present TTPDrill, a tool to extract threat actions from CTI unstructured text; it is the first approach to represent TTP in a structured form, in this case from published threat intelligence reports. However, although TTPDrill provides a novel model to identify threat actions in a machine–readable format from unstructured text, it relies on specific observables, not providing a general capability for the use of behavioral signatures.

Representing how an adversary works in an operation is not standardized among the CTI community, so this information has to be manually handled in most cases. As the relevant security information is usually consolidated in a SIEM (Security Information Event Management) platform, these technologies are the place where this information must be analyzed to detect indicators of compromise. Microsoft has developed Kusto Query Language, used in Azure both to monitor and to perform threat hunting [169], which is a non–standard language that can not be used *as is* outside the Microsoft ecosystem; in addition, providers such as Elastic have defined their own open rules and language (EQL, Event Query Language) to query Elastic SIEM. Both examples are proprietary ones, and their particular specifications can not be shared with other technologies.

An important effort towards the normalization of a language that allows analysts to query SIEM technologies in order to detect all kinds of IOC has been done with SIGMA rules. SIGMA[1] is a generic and open signature format to describe relevant log events in a straight forward manner. In other words, SIGMA is to SIEM events what Snort is to network traffic or Yara is to files. Although SIGMA goal is not to standardize a format to describe behavioral indicators of compromise, the language can be used to query SIEM events and provides a full coverage for all kind of indicators, from atomic to behavioral. For example, Turla Advanced Persistent Threat group executes different lateral movement techniques in compromised Windows system, identified as T1059, T1077, T1083 and T1135 by MITRE ATT&CK; ATT&CK, Adversarial Tactics, Techniques, and Common

[1] https://github.com/Neo23x0/sigma

Knowledge, is a globally accessible knowledge base of adversary tactics and techniques based on real–world observations. The following SIGMA rule allows to query a SIEM for these lateral movement techniques, looking for the execution of particular commands in Windows systems that are registered by sysmon and sent to the SIEM:

```
action: global
title: Turla Group Lateral Movement
id: c601f20d-570a-4cde-a7d6-e17f99cb8e7f
status: experimental
description: Detects automated lateral movement by Turla group
references:
    - https://securelist.com/the-epic-turla-operation/65545/
tags:
    - attack.g0010
    - attack.execution
    - attack.t1059
    - attack.lateral_movement
    - attack.t1077
    - attack.discovery
    - attack.t1083
    - attack.t1135
author: Markus Neis
date: 2017/11/07
logsource:
    category: process_creation
    product: windows
falsepositives:
   - Unknown
---
detection:
   selection:
      CommandLine:
         - 'net use \\%DomainController%\C$ "P@ssw0rd" *'
         - 'dir c:\\*.doc* /s'
         - 'dir %TEMP%\\*.exe'
   condition: selection
level: critical
---
detection:
   netCommand1:
      CommandLine: 'net view /DOMAIN'
   netCommand2:
      CommandLine: 'net session'
   netCommand3:
      CommandLine: 'net share'
   timeframe: 1m
   condition: netCommand1 | near netCommand2 and netCommand3
level: medium
```

This simple example does fit into the category of behavioral IOC, and of course the language allows to specify more complex rules. SIGMA has become the *de facto* standard to query SIEM events, but it does not provide full coverage for the specification of all behavioral procedures. This standard must be improved and complemented with post processing capabilities or equivalent over the stored data to be able to specify a full range of behavioral indicators of compromise. In addition, although SIGMA is supported by most SIEM technologies such as QRadar or Splunk, it is not the native query language in these technologies. This fact, in addition to the limitations of the language, forces analysts to maintain queries in different platform–dependent query languages if they want to explode to the maximum the SIEM capabilities.

User behavior and entity analytics, UEBA, sometimes referred simply as UBA, User Behavior Analytics (an elder concept superseded by UEBA), is also an effort to track behavioral indicators with own tools and with the integration into SIEM technologies; UEBA offers [101] profiling and anomaly detection based on a range of analytic approaches, usually using a combination of basic (e.g., rules that leverage signatures, pattern matching and simple statistics) and advanced analytics methods (e.g., supervised and unsupervised machine learning). UEBA strengths are related to the advanced methods but, although these could be efficient approaches in many cases, machine learning has not been proved to be a practical solution for the hunting of advanced actors. The problems in the anomaly detection field are clear and identified years ago [220], and in the case of the detection of advanced actors, many of them are linked to the stealth attacks these actors usually perform, as well as to the complexity to identify anomalies by establishing a user or entity behavior baseline. Although UEBA is a promising approach, it still lacks maturity while regarding advanced actors and, in any case, it does not provide a suitable standard to share behavioral indicators between different technologies or platforms.

## 6.5 Key requirements for behavioral IOC detection and sharing

By adopting the intelligence cycle as a working model we have identified key requirements for the detection and sharing of behavioral IOC. In this section we provide those requirements in each of the steps of the cycle. They all should be considered in the planning stage, to establish the direction of the process from the collection of data to the dissemination of intelligence in the form of behavioral IOC. Our final goal is to generate behavioral IOC that can be shared and used in an effective way to detect threat actors TTP. These requirements are focused on the detection and sharing of behavioral IOC: they must be considered together with other requirements for a Security Operations Center to perform threat hunting activities in an effective way. For example, we will not make emphasis on requirements as automatizing or false positive reduction, as they are not focused on behavioral IOC detection but on general detection capabilities.

During our research we have analyzed threat actors cyber operations, available both in frameworks like MITRE ATT&CK and in intelligence reports about particular threat actors such as [292] [428] [629] or campaigns such as [569]. Particularly, we have analyzed the activities of different advanced persistent threats from Russia (APT28, APT29, Turla...), China (APT1, APT17, Ke3chang...) and Iran (APT33, Clever Kitten...). This analysis allows the identification of key features of advanced offensive cyber operations, including not only their techniques but also characteristics related to their goals, targets or artifacts. The most relevant identified features are shown in table 6.2.

Table 6.2: Key features in offensive operations.

| **Feature** | **Description** |
|---|---|
| Multiple targets | Advanced threat actors target a wide spectrum of victims, including sectors such as military, government, technology, energy or even non–profit organizations |
| Broad range of techniques | Advanced threat actors achieve their goals through a broad range of techniques. These techniques are usually stealth, in order to go unnoticed, and one single threat actor can execute different techniques linked to the same tactic, even in a single operation against a particular target |
| Tailored tools and artifacts | Advanced threat actors can use multiple tools and artifacts in their operations. These tools and artifacts range from specifically developed malware to legit system tools, and in many cases the threat actor is aware of the deployed counter measures in the target and knows how to evade them |
| Potential indicators | Hostile activities leave traces in targeted systems and internal network traffic. In addition, the target perimeter security must be monitored in order to detect connections to Command and Control or exfiltration servers |
| Compromises spread over time | Once a target is compromised, this compromise spreads over time in most operations, thus giving the threat actor the ability to control its target for months or years |

The fact that a single threat actor targets multiple victims from different sectors is directly linked to the targeting of multiple infrastructures, protected by multiple security technologies. These technologies are provided by different vendors, and each of them uses its custom logging formats and data. In order to guarantee an accurate detection and response capability, this data must be normalized regardless of its technological data source through a common format that allows analysts the searching of hostile activities regardless the technology where they are performed. This data must be analyzed through a common platform–independent language that can be shared through different defensive teams and exploited no matter which technologies are monitored and which specific analysis platform is used in each case.

As advanced threat actors can achieve their tactics through different techniques, even in the same campaign, defensive teams must be able to search hostile behaviours regardless of the mechanisms used in each case. This means that the analysis tool, usually the SIEM [338] [339], must provide these teams the ability to specify all previously identified techniques as well as the new ones that are discovered during a particular analysis. This specification must be quick for the defensive team, in order to provide agility to the investigation of potential compromises.

The employment of tailored tools and artifacts, including legit tools provided by operating systems, is related to the stealth techniques executed by threat actors. Being covertness a must in hostile operations most movements will not generate any alert in security systems such as antivirus software or firewalls. This situation forces the defensive team to identify not only misuses, but especially anomalies, so it is mandatory to analyze, so to acquire and process, normal activities in systems and networks and, in most cases, to establish a baseline or a reference to define the normal behaviour of the users and infrastructures.

Hostile activities leave traces in different points of the targeted infrastructure: the compromised systems and the network traffic. In fact, as we have stated before, atomic and computed indicators of compromise are usually divided into host or network based. A particular sub set of network indicators are those related to domain names and IP addresses, which are usually seen on the network perimeter. This distribution of indicators forces the defensive team to analyze data from multiple sources, establishing relationships between them in a central repository where this data is received and stored. Although not particularly focused on the detection of behavioral IOC, works like [406] or [47] reflect the requirement to analyze, so to acquire, this multi source data. In fact, ATT&CK matrices available online from MITRE define the mandatory data sources for the detection of each technique, being most of them multi source.

Once a threat actor compromises its target, this compromise spreads over time; persistence periods range from months to years in many cases. For example, a threat actor such as APT1 can maintain access to victim networks for an average of 356 days, being four years and ten months the longest persistence period [398]. For this reason, to make a whole picture of the operation, the defensive team must be able to analyze, so to store, historical data in order to identify the initial entry point, the hostile activities performed and the internal systems that have been compromised.

To deal with the identified features of advanced offensive cyber operations, we have identified the key requirements for effective behavioral IOC detection and sharing. Following each of the steps of the intelligence cycle, in table 6.3 we summarize these requirements. Please note that they are applicable not only in Information Technology infrastructure, but also in Industrial Control Systems environments, where cyber threat intelligence is also a must [4] [648] and where all the problems we have identified in our work are also present.

Table 6.3: Key requirements for TTP detection.

| IC Stage | Key Requirements |
|---|---|
| Acquisition | Acquire data from multiple, relevant sources |
| | Acquire not only alerts, but regular events |
| Processing | Central data repository where relationships can be established |
| | Common format for stored data |
| | Long term retention |
| Analysis | Platform–agnostic implementation |
| | Full native coverage for all techniques |
| | Correlation of data from multiple sources |
| | Comparison of correlated data against a reference |
| Dissemination | Machine readable and exportable format |
| | Standard query language among providers |

### 6.5.1 Acquisition

For an effective TTP detection it is mandatory to acquire information from multiple data sources, those where main TTP can be identified. Taking as a reference the MITRE ATT&CK framework [596], where tactics and techniques are analyzed, we find the different data sources that enable the detection of each particular technique. Summarizing these data sources, we identify three main points to acquire data from:

- Endpoint, including not only user endpoints but also servers, where processes are created, files are opened and threat activities are performed at last; this data source includes global infrastructures for endpoints, just as Windows Active Directory.
- Network, including payload and net flow, where threat movements, both lateral and external, are performed.
- Perimeter, where input and output of data between the threat actor and its

target is performed, including network devices such as firewalls, Data Loss Prevention systems or Virtual Private Network servers.

In all cases, the mandatory information to acquire is that related to regular activities, not only the one related to alerts. Although we identify this key requirement, this global, regular data acquisition is not widely extended [508], thus impacting on the quality of later steps of the intelligence cycle and in the final intelligence product. A suspicious behavior is an event or a sequence of events, being in most cases each of these events not suspicious by itself. In other words, a behavioral indicator of compromise is not simply a set of atomic or computed ones. Thus, while atomic and computed indicators of compromise can be detected by rule–based systems such as Snort (for network indicators) or Yara (mainly for host indicators), many of the behavioral ones can not be identified this way. These systems provide specific misuse detection capabilities, so it would be hard for them, especially for Yara, to allow an analyst to identify tactics and techniques represented by behavioral indicators of compromise. This is an important point, as most tactics can not be detected laying only in the security alerts from a data source: as stated before, most threat actors will not generate security alerts in their regular activities. In fact, from a classical intrusion detection systems perspective, alerts could be considered "misuse detection", which has to be complemented with "anomaly detection" through the processing and analysis of the acquired regular events.

### 6.5.2 Processing

We have also identified particular requirements for the detection of behavioral IOC in the processing stage, once the analysts have acquired specified data. In first place, to achieve this detection, it is mandatory to have a centralized point where data can be collected; this point is usually a SIEM, where logs from multiple sources are stored in a common format. Related to the data acquisition from multiple sources requirement detailed before, in the processing stage it is a key requirement to receive and store this data, as it will be later correlated in the analysis stage.

Although this one would be also a requirement for the detection of all kinds of IOC, in the case of behavioral ones the requirement of having a centralized point with a common format for logs from multiple sources is especially relevant, as it allows the correlation between events from these different sources. As we will detail in next section, unlike the detection of atomic or computed indicators, techniques to detect behavioral IOC are usually based on the correlation of data from multiple sources, so we must consider this particular requirement as a key one.

In addition, as we will detail on the analysis stage, long–term retention is a must for a successful behavioral IOC detection. In fact, it is a must, from a forensic point of view, to detect all kind of IOC: when a threat intelligence feed is received, analysts must look backwards for its presence in the stored events.

However, when dealing with behavioral IOC, apart from this forensic approach, long–term retention is mandatory to identify stealthy behaviors. The detection of these stealthy techniques requires the analysis of events far in time, to compare them, and to establish relationships to identify the behavior of a threat actor. Without this long–term retention it may not be possible to identify techniques linked to advanced threat actors. Of course, to enable this kind of retention, and considering that an identified key requirement is to gather not only alerts, but also regular events from different sources, big data architectures that can handle all this information are a must, not only for the processing stage, but for all of the activities of the intelligence cycle [663] [421].

### 6.5.3 Analysis

The analysis stage is perhaps the most important one in the intelligence cycle, although it can not be accomplished without proper acquisition and processing activities. The requirements we have stated in previous steps of the intelligence cycle are mandatory for a successful analysis, thus enabling different capabilities for analysts to work, especially through the SIEM.

The first identified key requirement for the analysis is to be able to specify the behavioral IOC in a technology–agnostic way. This requirement implies that the IOC can be used regardless of the SIEM deployed in each case, but also that the same IOC specification regarding particular data sources can be used regardless of the technology of these data sources. For example, given a particular firewall (data source) and SIEM technologies, an IOC to detect a malicious behaviour on this firewall, by analyzing the data on this SIEM should be able to be loaded in any other SIEM and to provide the same results analyzing the data from other firewall technology in the industry. As we will see later in this section, it is a must to be able to load shared behavioral IOC from third parties.

This IOC specification capability has to provide native full–coverage for all identified techniques; this coverage can be achieved through a common query language, such as SIGMA, or through a common format for stored data and a suitable API to query this data. As we have detailed in section 6.4, the first option is the most extended among SIEM providers, being in this case the problem on each particular language and its incompatibility with the rest of approaches. Following this identified requirement, in the case of SIGMA it is mandatory to expand the language to provide full–coverage for all techniques, or to give SIGMA post processing capabilities to generate results in a format suitable to be managed by common programming languages.

Most techniques can not be identified by analyzing events from a single data source [479]; in fact, only about ten particular MITRE ATT&CK techniques (out of 185) can be detected using a single data source. So as we have stated before, the ability to establish relationships between events from different sources is a must. To achieve this requirement, SIEM technologies have to provide this capability, specially through the normalization to a common format of the information received from different data sources.

To identify most behavioral indicators it is also mandatory, in the analysis stage, to establish a relationship between events or alerts by comparing them against a specific reference; this relationship is usually a temporal one, but it can also be based on dependencies or simply on a comparison against a normality model. For example, a parent–child process match that can be considered suspicious.

SIEM environments are the main repository to store events from relevant sources from a security perspective; they collect not only misuse alerts, for example from systems such as intrusion detection systems or antivirus, but also "normal" events from all kinds of data sources, from endpoints to firewalls. They also provide the ability to establish relationships between events, so as we have stated before, SIEM should be the right place to identify all types of indicators of compromise.

### 6.5.4 Dissemination

At this point, successful detection of behavioral IOC could have been performed in a Security Operations Center. However, as we have stated before, threat intelligence sharing between defensive centers is a must: without a proper sharing, no single center can detect most hostile operations, especially those performed by advanced actors. For this reason, organizations collaborate to define defensive actions against complex attack vectors by sharing information about threats [492]. So now we will approach the last stage of the intelligence cycle, that related to the dissemination, to the sharing, of the final intelligence product.

As with atomic or computed indicators, behavioral ones must be usually shared by uploading them to a threat intelligence sharing platform such as MISP. To do so, behavioral IOC must be exportable as a single intelligence unit, commonly called "hunt". This implies that a single behavioral IOC generated in one platform, for example in a specific SIEM, can be exported from this SIEM in the form of intelligence unit.

In this dissemination stage, another key requirement for sharing behavioral IOC is that the intelligence unit has to be machine readable: this is, it has to be actionable when loaded into a SIEM. As with atomic and computed indicators of compromise, whose extended use mainly relies on the fact that they can be loaded into security devices and generate immediate results, as we have stated before [25], with behavioral IOC we identify the same requirement in order to be effectively shared.

To achieve this last requirement we also identify, as a previous mandatory requirement, a common accepted language for the specification and sharing of behavioral IOC among providers. Without such a capability, the effort that has to be done to translate TTP from natural language to specific, sometimes proprietary, technical standards as the ones referenced in this work, is not acceptable. With the current lack of a common language analysts have to iterate the specification among many platforms, with independent and non–compatible specification languages, thus multiplying their work not only for the initial specification but, which is most important, to its maintenance and upgrade.

Please note that in this section we have focused on the technological requirements for behavioral IOC sharing; other relevant aspects, such as legal or human matters, are out of the scope of our work, and in fact are common to all CTI sharing approaches [526] [649], not specifically to the dissemination of behavioral IOC.

### 6.5.5 A practical example

To provide a practical result for our proposal, we have analyzed a particular technique performed by threat actors in their operations and how the different requirements we have identified must be fulfilled. We have chosen the Command and Control (C2) tactic stated by MITRE ATT&CK, in particular the T1071.001 technique, related to command and control through web traffic protocols. While using this technique, adversaries may communicate using application layer protocols associated with web traffic to avoid detection and network filtering by blending in with existing traffic. A malicious HTTP/S hit is hidden inside the whole legitimate web traffic. Commands to the remote system, and often the results of those commands, will be embedded within the protocol traffic between the client and server, as MITRE ATT&CK framework states. For our example we have chosen a particular implementation of the technique, in which the hostile actor buys and uses domains whose name contains strings related to legit sites, trying in this way to go unnoticed. We must highlight that this technique implementation is behavioral, so it is not possible to detect it through the use of atomic or computed indicators of compromise.

The detection of this particular technique requires in first place to acquire, process and analyze information from a main data source: the web proxy or equivalent, where navigation logs are stored. These logs must be analyzed with a mechanism that is able to detect the use of legit strings inside malicious web names; a simple approach can be looking for pre identified legit strings related to existing companies, such as “google”, “microsoft” or “adobe”, into the domain names that have been navigated from the organization.

Once a suspicious navigation has been identified, a set of actions must be performed by the analysis team. These activities are shown in figure 6.3 in the form of playbook. In first place, the analysts will query intelligence sources to check if the domain is suspicious or not, for example by confirming that it is registered by the original company or by a third party entity. In this last case, as the domain will be considered suspicious, the analysts will look for more hits in the historical navigation records (for example, in the stored proxy logs with the available retention period) with the goal of identifying the time period the domain has been connected from the organization.

By analyzing data from these historical records, the analyst will identify the set

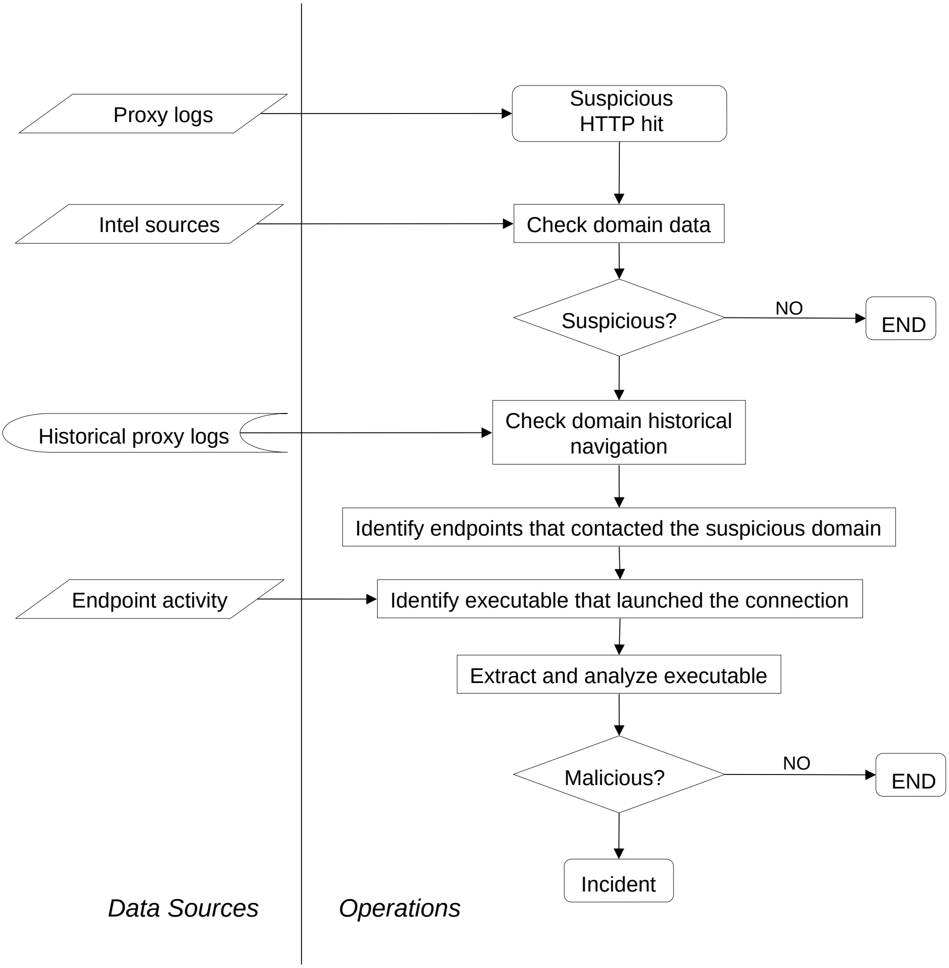
Proxy logs
Suspicious HTTP hit
Intel sources
Check domain data
Suspicious?
NO
END
Historical proxy logs
Check domain historical navigation
Identify endpoints that contacted the suspicious domain
Endpoint activity
Identify executable that launched the connection
Extract and analyze executable
Malicious?
NO
END
Incident
Data Sources
Operations

Figure 6.3: Playbook for the analysis of suspicious HTTP hits.

of endpoints that contacted the suspicious domain; for all of them, and by the analysis of endpoint activity logs, the executable that performed the suspicious connection will be identified, in order to extract and analyze it in a malware laboratory. If this last analysis confirms that the executable is malicious, an incident will be raised; otherwise, it will be considered a false positive and the investigation will be closed. If an incident is detected, it is mandatory to specify the detection of the technique in the form of behavioral IOC, to automate its detection not only in the organization, but also to share it with other interest groups.

With this simple example, we can confirm the mandatory requirements for the detection of a behavioral IOC. In first place, it is clear that data from multiple sources must be acquired to enable the detection of the technique: in this case, from the navigation logs and from the endpoint activity. None of these sources can provide by itself the full picture of the situation to the analyst, as they log different types of information that must be considered together for the whole detection scheme. In addition, as an anomalous domain web hit is not a security violation by itself, it does not generate an alert in any security mechanism. In this way, if we register only alerts, these anomalous hits will go unnoticed.

Regarding the Processing step, we face in first place the retention of the data; we must have a long term retention scheme to identify historical records that help us to draw the global picture of our investigation: the whole of endpoints contacting to the suspicious domain, the time this domain was first seen in the organization or the frequency of hits. In addition, when dealing with medium or large organizations, it is mandatory to centralize information from multiple sources in a single repository and with a common format. This requirement will enable analysts to establish relationships between events from these sources; otherwise, these relationships would have been manually established, which is not possible in most organizations.

The Analysis step is performed in the SIEM platform. This platform must provide mechanisms for the full analysis of all data from multiple sources, as it has to centralize all the information from these sources and provide the capability to exploit it. In this exploitation, correlation is mandatory for the analysts to establish relationships between different data sets, as it is the establishment of a common reference (usually a temporal one) for all the data. A platform–agnostic implementation is also a must for two main reasons: the first one, to compare data from different technologies across the organization. A single organization can have proxies, firewalls or EDR (Endpoint Detection and Response) from multiple vendors, each of them generating logs in its own proprietary format. The second one is perhaps the most important: to be able to automate and share, as we will see in next step, the specification of the technique among multiple organizations that will surely have different security technologies.

Finally, in the Dissemination step, to be able to share the specification of the technique with other security groups, and also to receive and exploit specifications from these security groups, the two key requirements we have identified must be fulfilled. The technique must be specified in a standard language among technology providers, particularly SIEM ones, as not all the organizations use the same SIEM. In addition, this specification must be exportable, to allow sharing and, most important, it must be machine readable to load it automatically in security tools to detect the identified technique.

We have provided a practical example for the detection of a particular technique and identified the fulfillment of all the key requirements we have proposed. Although we have chosen a simple technique used by advanced threat actors, our findings can be applied to all kind of techniques; if we face more complex techniques against large organizations, the fulfilling of all the requirements is also mandatory.

## 6.6 Discussion

The most used types of indicators in CTI sharing are atomic and computed ones: in particular, file hashes, IP addresses and network domains. However, as these indicators of compromise are easily changed by a threat actor, they can be evaded with little or no effort. While they are definitely not the most useful ones, they are the most used and shared, mainly because they can be expressed in machine–readable format: this feature allows these indicators to be automatically loaded into security devices, thus providing an immediate result for the customer. But this situation represents a problem for security analysts, as most of the shared intelligence is easily evaded by hostile actors, rendering it useless. This problem must be addressed with the detection and sharing of behavioral indicators of compromise. In this work we have identified the key requirements for both of these activities and discussed them to detect hostile techniques, providing a practical example.

All but one of the key requirements we have identified in our work are today fulfilled with most SIEM technologies. The main unfulfilled requirement is the definition of a standard query language among providers. There are some efforts to fill this gap, mainly aligned with MITRE ATT&CK and especially in SIEM technologies, being the most relevant the SIGMA language. However, until now there is not a commonly–accepted standard suitable for sharing behavioral signatures between analysts using different technologies. This lack motivates two problems: the inability to share, thus the difficulty to detect, and the effort to translate TTP from natural language to specific vendor–dependent formats, not compatible between them.

These two main problems can be only addressed by the specification of a commonly accepted standard that provides full behavioral detection capabilities to SIEM or to any technologies and products focused on the detection of advanced threat actors. The definition of such standard will enable the use of these indicators among analysts using different technologies, and as a direct consequence, this usage will enable information sharing between analysts. As this standard becomes accepted and used, and behavioral signatures can be shared and automatically loaded into security systems, the detection capabilities for security teams will increase significantly.

Although there are many efforts to characterize threats among the CTI community, including their behavior, little progress has been achieved to define a machine–readable format accepted as an standard. In this sense we have identified two main references to be considered: STIX and SIGMA. STIX provides capabilities to define TTP as SDO, but it does not allow to specify behavioral signatures; the specification of TTP is based on atomic and computed indicators of compromise, so if an attacker changes them, the particular TTP will not be detected. SIGMA language is the closest approach for such a standard, but it has to be improved to expand its capabilities. In addition, at the time of this writing it is not natively supported by many SIEM technologies, so SIGMA queries must be turned into specific SIEM signatures through a converter. This fact introduces two main problems: complexity in the management of SIGMA queries among multiple SIEM providers and dependency on the capability of those converters.

As we have stated before, from a technical point of view the only unfulfilled requirement to detect and share behavioral indicators of compromise is the definition of a standard query language among providers. However, being technically fulfilled does not mean that all defensive teams are able to perform this detection and sharing. The complexity of our work relies on the identification of a structured methodology suitable for the analysis of the lessons learned after an incident, as well as on this particular analysis. We adopt the Intelligence Cycle as we defend this identification of hostile operations is a counter intelligence activity, so it has to be structured and analyzed in this way. The highlight of requirements and their arrangement based on a consistent model provides a homogeneous framework to identify gaps on the detection capabilities provided by CERT teams. The analysis of the lessons learned to identify the mandatory requirements for the detection and sharing of behavioral indicators of compromise has its own complexity. This kind of analysis after an incident is not usually public, so we have partially based our research in our own experience in incident handling.

We identify as a future research line the measurement of the quality of indicators of compromise. Although behavioral indicators have a longer time of life and provide a more accurate detection capability, the analysis of the quality of IOC, in order to decay them, is an ongoing work. We defend that, in general terms, a behavioral IOC is more useful than an atomic or a computed one, but this generality has exceptions, and the quantification of this usefulness must be analyzed in all cases.

Another relevant research line is the use of behavioral indicators of compromise in emerging or changing applications and technologies in which security is a especially relevant issue. In this sense, we identify research fields such as smart contracts [478], smart cities [193], internet of things [380] [563] [355], 5G communications [601] or crypto currencies [437], as well as the combined application of these trends. In all of them, the application of all kinds of IOC, including behavioral, is a key issue for the detection of security compromises.

## 6.7 Conclusions

A relevant problem in cyber threat intelligence is the sharing of behavioral indicators of compromise, those that specify the tactics, techniques and procedures of hostile actors. The fact that atomic and computed indicators of compromise are the most shared and exploited ones reduces the detection capabilities for defensive teams: these indicators are easy to evade, which leaves a window of opportunity for threat actors.

In this work we have analyzed this problem and identified the key requirements for the detection and sharing of behavioral indicators of compromise. To structure our requirements we have followed the intelligence cycle, so we identify these requirements in each of the stages of the cycle. This identification is based on the analysis of threat actors and their campaigns, from where we extract the main features of these hostile operations. The detection of any type of IOC has to be performed through a SIEM, where the information gathered from different sources is normalized and centralized. Nowadays, most SIEM technologies provide the mandatory technical capabilities and fulfill all the identified requirements. However, the main barrier to the detection and sharing of behavioral indicators of compromise is the lack of a common machine–readable format to specify behavioral signatures and share them among different SIEM providers.

This lack must be addressed in short term through the definition and acceptance of such a standard, a common format to describe behavioral indicators of compromise. Until the security community does not define and accept this standard, these indicators will not be massively used, so they will not be massively shared, and the ability to detect the TTP of advanced actors will remain residual. In this sense, we identify SIGMA as the main current effort, but its acceptance among technology providers has to be increased.

Once this common standard has been designed and accepted among the community, the rest of the work will rely in each particular team defensive capabilities, from acquisition to dissemination. The identified key requirements are mandatory for an effective detection of behavioral IOC, but all of them can be fulfilled with current technologies. In this sense, our proposal, structured in the form of intelligence cycle, can be used to measure and compare the capabilities of different defensive teams.

Although the value of a behavioral indicator of compromise is usually higher than the value of a computed or atomic one, in some cases a simple IOC can provide a high quality detection. Regarding the measurement and quantification of the quality of an indicator of compromise, the type of IOC (atomic, computed or behavioral) is of course relevant, but many other parameters must be considered, such as the source reliability or the last time the IOC was seen. We identify this measurement as a relevant research line for future works.

# Chapter 7

# Discussion

## 7.1 Advanced Threat Actors in cyberspace

Cyberspace is a global domain within the information environment consisting of the interdependent network of information technology infrastructures and resident data, including the Internet, telecommunications networks, computer systems, and embedded processors and controllers [160]. The employment of cyberspace for all kind of activities in modern societies is questionless, and it increases exponentially, from electronic banking to video conference systems, remote working capabilities or general entertainment (streaming, online gaming, etc.). In this way, global dependence on cyberspace for common activities also increases exponentially for particulars, organizations and even whole countries and societies. An example of this situation was shown in 2020, when the global COVID–19 pandemic increased our dependence on online systems for many of our daily activities [464] [373] [203].

However, cyberspace is not only used for legitimate purposes. All kind of threat actors also exploit the benefits of working in cyberspace. In fact, during COVID–19 pandemic, where our dependence on cyberspace was especially relevant, hostile activities increased dramatically [665] [457], mostly cyber crime related ones [97] [324]. It is important to highlight that, from a threat actor's perspective, cyberspace provides enormous benefits, such as:

**Accessibility** Access to cyberspace capabilities is easy in front of access to ground, air or naval ones. This fact boosts asymmetric conflicts.

**Plausible deniability** Cyberspace operations are hard to link to a real physical threat, such as a nation–state. For this reason, while classical operations can have direct consequences for the hostile agents, such as diplomatic conflicts or even personal losses, cyberspace operations are easy to deny for an actor.

**Geographical offshoring** In front of classical conflicts, cyberspace operations are not easily linked to a specific geographical context, so in some cases it is difficult to face them, as defenders do not have a localized enemy who can be responded.

All of these benefits for a threat actor, in terms of money, information, or deniability, increase the probability of being the victim of hostile activities through cyberspace. On the other hand, from a target's perspective, dependence on cyberspace increases day by day. As exposed before, today it is difficult to find activities that are not performed through technological resources, so when these resources are compromised, a whole organization or even a society can be damaged. For this reason, the associated impact of an hostile activity through cyberspace is also increasing day by day.

With high probability and high impact, the risk of being targeted by hostile activities through cyberspace is also high. The number of hostile operations has increased dramatically during last years, as well as their sophistication and the range of potential targets: from small companies to whole countries. For most organizations, being targeted by a nation–state threat actor or by a cyber crime gang is a real fact. Hostile activities through cyberspace represent a threat not only for private or public organizations, but for the whole society.

Not all hostile activities through cyberspace are relevant for this work, focused on advanced threat actors. By this concept, I refer to threat actors with high capabilities (technical, economical, etc.) that perform hostile activities through the cyberspace. Although most of these hostile activities are performed by nation–state actors, during the last years different cyber crime gangs, not sponsored by governments, have evolved their tactics, techniques and tools to reach, in some cases, high capabilities. For this reason, nowadays it is possible to find two main types of advanced threat actors: nation–state and non–state ones.

Related to nation–state threat actors it is mandatory to refer to cyberspace operations, a term which is usually linked to the military [450] [625], defined as the employment of cyberspace capabilities where the primary purpose is to achieve objectives in or through cyberspace. Although the most common activities are related to cyber espionage, these cyberspace operations include all kind of offensive information operations, ranging from electronic warfare to influence campaigns, as well as destructive actions (cyberspace attack). The number and impact of these last ones is particularly increasing during last years, with cyber physical systems being connected to Internet [144] [181] in critical infrastructures.

In addition to purely military cyberspace or nation–state operations, when dealing with advanced threat actors, non–state activities must be considered, particularly cyber crime related ones, where the hostile agent has an economic goal. Not only governments use the cyberspace as a medium to achieve their goals, but also criminal gangs do, and some of them have deployed advanced capabilities that are a serious threat for all organizations. Human Operated Ransomware groups are an example of advanced threats linked to criminal gangs, whose techniques and tools are in some cases similar to those of nation–state actors.

## 7.2 Categorization of Advanced Threat Actors

The approaches to threat actors's categorization are based on different attributes of the actor. These attributes, such as funding or motivation, are always a common starting point for their categorization. In this way, different classification schemes for the categorization of threat actors have been defined in literature. In [321] Dimitrios Kavallieros et al. classify threat actors into three groups: internal to the organization, external to the organization and mixed; the authors also provide a taxonomy of seven categories of cyber crime actors, from coders to criminals. Alternative classifications are provided in [539], [389] or [198], where Youngsup Shit et al. propose an automated classification scheme for threat actors based on the MITRE ATT&CK framework.

However, the assessment of the threat actors' attributes is in some cases impossible. For this reason, most accepted industry categorizations rely on the attributes whose assessment is more reliable. For example [7] and [395] rely on attributes such as motivation and capability to establish a classification for threat actors. Capability can be specified in features such as expertise, knowledge, time or equipment [531]. In real offensive cyberspace operations, these terms are reflected in the form of observables, from low–level ones to tactics, techniques and procedures [105]. In this work, I defend that capability is linked to the characterization of the threat actor, not to its categorization, so we must discuss the motivation of advanced threat actors in order to establish a categorization for them.

For this reason, in this work the approach for categorization for threat actors is based on their motivation: on what the threat actor wants to achieve through its hostile activities. Following this approach, we can find actors such as nation–states performing espionage or cyber war actions, hacktivists, cyber terrorists, etc. Of course, not all of these actors can be considered advanced ones. As I have stated in this work, it is possible to find two main families of advanced threat actors: mostly nation–state sponsored ones, as well as specific non–state actors performing advanced cyber crime operations with an economic goal. The rest of types of hostile threat actors in cyberspace, such as hacktivists or even cyber terrorists, can not be considered advanced ones. Their tactics and techniques, their tools and their general capabilities are not advanced. However, this fact does not mean that the impact of their hostile activities can be high.

When referring to advanced threat actors it is common to focus on Advanced Persistent Threats and, more specifically, on Advanced Persistent Threats performing cyber espionage operations. These state–sponsored actors and their operations have become one of the biggest threats for all types of organizations. However, nowadays we can identify as advanced threat actors both state–sponsored ones (usually referred as strategically motivated, or S–APT) and criminal gangs (usually related to operationally motivated APT, or O–APT), without entering into the potential relationships that both of them can maintain in certain countries. In addition, state–sponsored threat actors perform not only cyber espionage or even cyber attack actions, but a whole range of Information Operations through cyberspace: cyberspace operations.

Related to nation–state actors, most of their activities are related to cyber espionage, as previously stated. However, these actors have been involved in cyber attack campaigns, influence operations and even non–strategically motivated actions, such as direct economical benefit [334] [251]. Although these ones are not as common as cyberspace operations, their impact can be enormous for an organization not only in economical terms, but also in reputation. On the other hand, criminal gangs, threat actors linked to cyber crime activities, have improved their technical capabilities by the employment of powerful tools and tactics, techniques and procedures that in some cases are difficult to differentiate from the ones used by state–sponsored actors. The goal of these actors is purely economic, targeting all types of organizations. Their modus operandi, although less stealth, is very similar to that of nation–states actors.

Regardless of their motivation, it is important to highlight that the impact of a hostile activity performed by an advanced threat actor is usually high. Particularly, when the targeted infrastructure is a critical one, those which provide essential services to the society: in these cases, a whole country can be seriously damaged. Advanced threat actors have not only the capabilities, but also the intent to target a victim, so their success rate is high. For this reason, it is important to delve into their general knowledge and particularly into their characterization. And it is even more important to be able to detect and to neutralize their activities in an efficient manner.

In this section the motivations of different threat actors are analyzed and discussed, in order to establish which of them can be considered advanced ones. Two main families of these advanced threat actors are defined: nation–state and non–state actors, delving into their operations in the cyberspace. Finally, I analyze the impact of these operations in critical infrastructures, as a key target for all kinds of threat actors, including advanced ones.

### 7.2.1 Motivation of Advanced Threat Actors

The motivation of threat actors has been largely addressed in different works [256] [523] and it provides a clear classification for threat actors. A commonly accepted model for threat actors' categorization is a fivevold classification based on the motivational factors of these actors [357] [223]. Although different classifications for threat actors have been presented, including more complex approaches [539], I consider the fivevold model is simple and valid enough for an initial classification. This model defines cyber activism, cyber crime, cyber espionage, cyber terrorism and cyber warfare as the main motivations for threat actors and states a commonly accepted categorization for them [607]:

- Cyber activism mainly consists of vandalism and hacktivism, although the difference between these activities is not clear in some cases. It is related to different forms of activism where the cyberspace is the base or allows the enhancement of online actions [643].

- Cyber crime is defined [72] as criminal acts committed using electronic communications networks and information systems or against such networks and systems.
- Cyber espionage is [414] about gaining access to a computer systems and retrieving information from them. It is defined as Computer Network Exploitation in military doctrine.
- Cyber terrorism is defined [488] as the premeditated attack or threat thereof by non–state actors with the intent to use cyberspace to cause real–world consequences in order to induce fear or coerce civilian, government, or non–government targets in pursuit of social or ideological objectives. Real–world consequences include physical, psycho social, political, economic, ecological, or otherwise that occur outside of cyberspace.
- Cyber warfare has no universally accepted definition [527], although it is a widely used term to describe the operations of state–actors in cyberspace in order to gain superiority over an adversary.

Nowadays, advanced threat actors are mainly linked to nation–states, but also to some criminal groups that have enough capabilities to perform complex operations. State–actors are not only performing cyberspace exploitation, cyber espionage, but a whole range of Information Operations in the cyberspace. From an offensive point of view, these operations include cyberspace exploitation, cyberspace attack, psychological operations on the cyberspace, or electronic attack. Other IO core capabilities have a defensive role, as shown in figure 7.1.

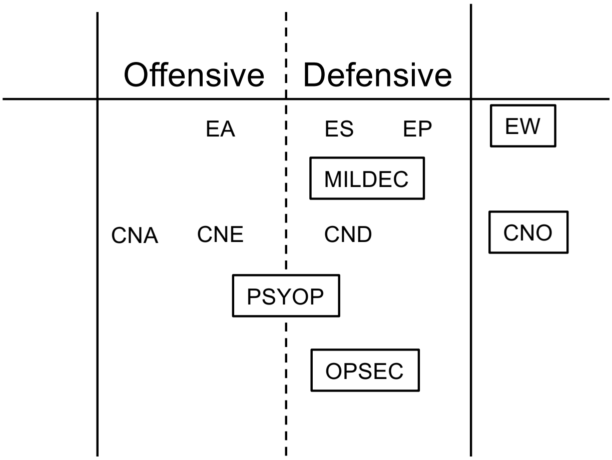

Figure 7.1: Defensive and offensive Information Operations.

As I have previously stated, it is mandatory to highlight that, apart from nation state actors, cyber crime actors have dramatically increased their hostile capabilities during the last years. Although their goals are related to economic gain, in some cases their techniques are so complex and stealth that it is hard to differentiate them from nation state actors. This situation involves that when dealing with advanced hostile cyberspace operations, nowadays we face not only those regarding Information Operations performed by nation–state actors, but also those related to cyber crime.

Cyber terrorism and the capabilities of cyber terrorist groups have been largely addressed in many works from late nineties [514] [215] to nowadays [528] [156] [20]. At the moment of this writing, the use of Internet by terrorist groups is limited to actions regarding recruitment, radicalisation, fundraising or propaganda [207] [45] [630]. No terrorist action with a relevant impact has been performed through cyberspace [551] [270] [358]. However, cyber terrorism must be considered as a growing threat, and cyber terrorists could be considered advanced threat actors in short term. As terrorist groups increase their "know how" about cyberspace operations, a terrorist action on cyberspace is more likely to occur, especially targeting critical infrastructures. In this sense, these infrastructures are not only on risk by cyber warfare offensive operations, but probably also by cyber terrorism ones in a short or middle term.

Finally, cyber activists can not be considered advanced threat actors in general terms. Cyberspace as an arena for political activism and social movements is increasing its importance [542]. However, the modus operandi of cyber activists is usually not advanced. In most cases they do not have a defined strategy and clear goals, although some advanced strategies have been adopted by these actors [247]. Their tactics and techniques are not well structured and their activities are not stealth. In addition, they do not employ advanced tools, malware or exploits [382] [131]. However, please note that these elements do not imply that their hostile activities can have an enormous impact on their targets.

Despite the fact that advanced threat actors are a real threat, it is important to note that just as in every other instance of human conflict, they are composed of human beings and have properties in common with all other types of adversary [378]:

- They have goals and objectives.
- They have resource limitations.
- They engage in mission planning, practice, development and testing.
- They translate their behavior into Computer Network Operations.

Summarizing, nowadays we face two main types of advanced threat actors (based on their motivation): state–sponsored and non-state actors. The first ones are engaged in a whole range of offensive cyberspace operations, although in some cases they also have a operational economic goal. Non-state actors are engaged in economic crime through operations such as Human Operated Ransomware (HOR). In figure 7.2 this summary is shown. This data is consistent with the analysis of motivations of Advanced Persistent Threats performed by different groups and researchers [14]; for example, ThaiCERT periodically publishes [610] a report regarding Advanced Persistent Threats where elements such as attribution, motivation or operations are described. Analyzing this report, it is possible to confirm that most of these groups are motivated by information theft and espionage, a much reduced set of the is motivated by financial gain/crime and, finally, less than fifteen groups are devoted to sabotage and destruction. None of them is performing hacktivism or terrorist activities.

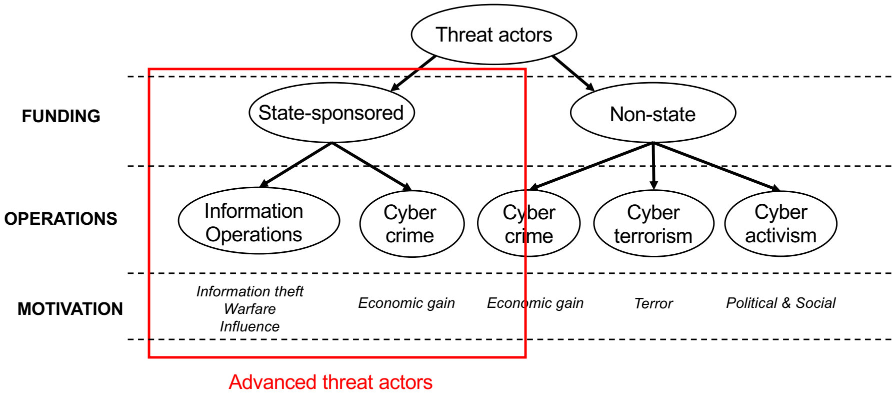

Figure 7.2: Threat actors' categorization and motivation.

## 7.2.2 Nation–state actors

Nation–state actors represent a serious threat to the technologies and to the information of most organizations. If an organization handles relevant information (this is, information which a high value from any point of view, from economic to diplomatic or military), it is a clear target for nation–state actors. This includes high–technology companies, international non–governmental organizations (NGO) and, especially, all kind of critical infrastructures (see section 7.2.4), from public administration to banking or utilities. Comparing with other threat actors, nation–state ones have a fixed interest in their targets, as well as the resources (both human and material) and the time to execute their operations.

In section 1.4.6 the concept of Advanced Persistent Threat, or APT, was introduced. Usually, APT groups are linked to states, and they are engaged in different offensive information operations, as stated before. Cyber espionage activities are the most common ones among APT groups, including the groups performing other types of offensive information operations, such as APT28, a Russian APT which performs not only cyberspace exploitation, but also disruptive operations, especially against critical infrastructures as noted in section 7.2.4, and even psychological operations [543] [369] [560]. However, it is important to highlight that there are cases of nation–state actors performing economic operations to finance their states, such as North Korean Lazarus [471] [664]. Although this is not a common trend, some nations need to bear the expenses for their cyber espionage operations through economic crime: they have the knowledge, the resources and the intent.

As previously stated, cyber espionage activities are the core ones for nation–state groups. Most unveiled operations are related to cyberspace exploitation, and most of the identified groups perform only exploitation activities [610]. The vast majority of published reports regarding APT groups are focused on cyber espionage operations [398] [292] [192] [632]. As their own name states, these operations are focused on intelligence gathering. When performed through cyberspace, these activities have less risk for the threat actor than information gathering through other means, and in many cases their results are better. For this reason, nation–state actors not only engage in such activities, but also work in programs to overtly leverage commercial products' design, making systems exploitable by inserting vulnerabilities and overt intelligence collection capabilities. The SIGINT Enabling Project [53] [65] [606] is an example of such programs.

Regarding cyberspace attack, few groups are known to be engaged in destructive operations [125] [209] [244] [428]. Among them, we can find Sandworm Team and APT28, linked to Russian GRU [309] [245] [383], APT37, APT38 and Lazarus, linked to North Korean Reconnaissance General Bureau [512] [471] [273] and Gamaredon, linked to Russian FSB [161] [108]. As we can guess, Russian offensive capabilities in cyberspace are strong, far away from the known capabilities of other actors, with the exception of United States [224]. I must highlight that destructive operations are commonly linked to cyber war. This association is wrong, as in a potential cyber war we would see both attack, exploitation and of course defence operations. Until now, no cyber war has ever been formally declared. However, it is common to refer to cyber warfare, a different concept than cyber war: while cyber warfare is an activity, cyber war is a state [527]. As I have stated before, cyber warfare has no universally accepted definition. However, in this context it can be considered [243] as an extension of policy by actions taken in cyberspace by state actors (or by non–state actors with significant state direction or support) that constitute a serious threat to another state's security, or an action of the same nature taken in response to a serious threat to a state's security (actual or perceived). Attending to this definition, cyber warfare is a term closer to a whole range of information operations in cyberspace than to purely destructive actions.

Finally, an interesting discussion is the employment of non–state actors in cyberspace operations performed by nation–state actors. As in classical warfare, these non–state actors are usually know as proxies [423]. Although the use of proxies has operational and legal drawbacks [572], there are different benefits to use them in cyberspace operations [162], such as cost savings, plausible deniability, skills and specialization and punitive power [206]. The use of proxies is obviously not publicly known in most cases; however, it is apparently a rising trend for state–actors [555].

### 7.2.3 Non–state actors

The main rise of non–state actors, and the development and improvement of their tools, tactics and techniques, is in most cases linked to financial motivations. Advanced groups engaging in criminal activities for their own financial benefit are referred as Operationally Motivated APT (O–APT) [14], in front of Strategically Motivated ones, or S–APT. Economic cyber crime is a growing threat that all kinds of organizations must face nowadays [455], and these actors are continuously improving their technical capabilities in order to increase their revenues [463]. In this sense, as I have stated before, Human Operated Ransomware tools and techniques are the closest ones to these linked to nation–state actors, and Human Operated Ransomware (HOR) groups are considered advanced actors in many cases [234] [467]. Although end users were the primary target for ransomware gangs, in the last years organizations have become their main victims. These target organizations are chosen in advance, and the goal of the threat actor is to cause maximum disruption to get a big ransom payment [467].

HOR criminal groups, such as HIVE [117], CONTI [622] or DARKSIDE [445] target all kinds of organizations, from academic to government. However, critical infrastructures are becoming key victims, as detailed in section 7.2.4: in these cases, the impact of a ransomware operation is very high for a whole society, so the requested ransom and the probability to get paid increases. However, HOR is not the only type of threat actors that, under certain circumstances, can be considered advanced. From 2013 to 2015 cyber criminals targeted financial institution through the Carbanak remote backdoor [416], initially designed for cyber espionage. This campaign exposed different advanced techniques, linked to nation–state actors, used in cyber crime operations [242].

In addition to the activities of criminal gangs, the cyber crime marketplace is evolving over time. Cyber crime as a Service (CaaS) operations involve distributed denial of service attacks, credit card fraud, compromised systems selling and many other types of cyber crime [645] [287]. In the particular case of Ransomware as a Service (RaaS), criminal groups rent from packers and algorithms to social engineering capabilities [323]. However, advanced tools do not make advanced actors. Different factors, such as the lack of advanced tactics and techniques in the end user, and the absence of trusted marketplaces, for obvious reasons, make Ransomware as a Service a modest threat [402] in comparison with the activities of ransomware groups directly attacking their targets.

Finally, it is important to highlight that not all threat actors linked to cyber crime operations can be considered advanced ones. There are a lot of cyber crime gangs performing not advanced operations, but still being successful. Electronic mail frauds [473] [87], cyber stalking [486] [320] and even different financial crimes such as banking phishing [12] [319] do not require advanced tools or techniques to be executed. In these cases, the modus operandi of the threat actor is based on mass–targeted operations, not on directed ones. In this way, with basic tools and techniques, targeting millions of people, the global revenue is high: low investment and high benefits.

### 7.2.4 Critical infrastructures

When facing hostile actors' activities, a key asset to protect are critical infrastructures: infrastructures so vital that their incapacitation or destruction would have a debilitating impact on defense or economic security [419] [420]. As its own definition states, the security of these infrastructures influence the life of particulars and the normal operation of any sector critical for the operation of whole societies [50]. Although these sectors are defined by national laws, they commonly include energy, water, telecommunications, banking or government facilities and services, among others [48]. As an example, in figure 7.3 the twelve strategic sectors defined in [170] by the Spanish Government are shown. Due to their relevance for the society, critical infrastructures are a key target for different groups of threat actors, particularly for advanced ones: nation–states (cyber espionage and cyberspace attack) and cyber crime [362] [61] [66]. And not only from a cyberspace operations' perspective, but also from a pure physical one [37] [505].

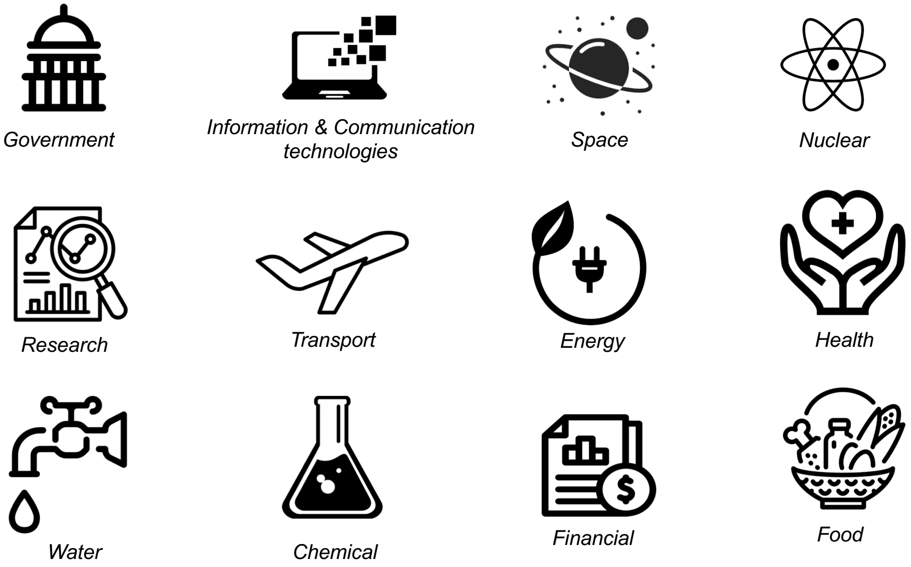

Figure 7.3: Spanish strategic sectors.

Related to nation–state threat actors engaged in cyberspace operations, critical infrastructures are a target both for cyberspace attack and for cyberspace exploitation actions. From a cyber espionage perspective, critical infrastructures' operators handle sensitive information of interest for foreign countries. This information can be targeted for cyber espionage with multiple purposes, from military and diplomatic to pure industrial espionage. In addition, as stated before, all cyber attack operations, or at least the advanced ones, require a previous cyberspace exploitation, this is, a mandatory information acquisition action to plan and execute a tailored attack.

On the other hand, from a cyberspace attack perspective, critical infrastructures are especially vulnerable to sabotage operations. Particularly, those that are based on Industrial Control Systems (ICS), the so called cyber–physical systems [132] [676]. They are a key target for threat actors due to their vulnerability level and to the impact an hostile action can have. Disruptive attacks to critical infrastructures are analyzed in [401] from a military approach, where strategy, operational and tactical considerations are considered. In fact, different disruptive threats, such as APT28 (Russian GRU), are linked to military intelligence services or units, not to civilian ones. Focusing on ICS, in [296] Eduardo Izycki and Eduardo Wallier Vianna detail seven cyberspace attack operations targeting critical industrial infrastructures. All of them are attributed to nation–state actors such as Iran, Russia, Israel or US.

Related to cyber crime gangs, money is a key motive behind operations on critical infrastructures [587]. Human Operated Ransomware groups have hit different critical infrastructures, mainly in the health sector during COVID-19 pandemic [328] [491]. As in every sector or industry, ransomware also affects critical infrastructures. However, in this case, with individuals and whole societies relying on the proper operation of these infrastructures, the impact of a ransomware operation is enormous. Criminal gangs are aware of this fact and target critical infrastructures, from particular industrial ones [32] [150] to country–global technologies, such as in the Costa Rica 2022 incidents [107] [33].

## 7.3 Characterization of Advanced Threat Actors

As previously stated in section 1.4.7, the characterization of threat actors is the extraction and analysis of threat actors' features, in order to identify their interests, goals or operations, among others. Although this characterization can be performed through all the intelligence gathering disciplines, SIGINT and TECHINT are the most relevant ones in most cases, as the characterization usually starts by direct observables that are turned into indicators of compromise. However, to discuss the whole characterization of threat actors, we must consider both direct observables and non–observable elements, such as goals, strategy and even attribution. As these ones are not directly seen in an operation, they must be inferred from an intelligence analysis, apart from the purely technical aspects of the operation. This analysis, outside of the scope of this work, will deduce, with an associated probability, why a threat actor is conducting a hostile operation against a particular target. The identification of goals, strategies and attribution provides valuable information to establish tailored security countermeasures to face specific threat actors.

In table 7.1 the main families of features regarding threat actors are summarized.

We must differentiate between observable features (those that can be directly seen on an operation) and non–observables ones (those that are not directly seen, so they must be inferred or acquired by external intelligence). Low–level observables are linked to tactical intelligence and TTP are linked to operational intelligence; both of them can be expressed in the form of indicators of compromise. On the other hand, non–observables are linked to strategical intelligence. All of them are relevant for the correct characterization of a threat actor, although strategical intelligence is rarely actionable.

Table 7.1: Threat actors' features.

| | |
|---|---|
| Non–observables | Attribution |
| | Goals and strategy |
| Observables | TTP |
| | Low–level indicators |

In this section, the characterization of threat actors through their observable and non–observable features is discussed. This approach starts with low–level observables and ends with the attribution, one of the main relevant problems that threat intelligence analysts face nowadays. All of the discussed features are important to the whole characterization of a threat actor, from its arsenal to its interests. However, I defend that the characterization of advanced threat actors must be mainly approached by the analysis of their tactics and techniques. They are the most valuable observables in the context of a cyberspace operation. This value is linked to the fact that lower level observables, such as atomic indicators of compromise, or even tools or artifacts, are easily modified by an actor, so their value is limited. On the other hand, features such as goals and strategies, or even attribution, are not direct observables in a hostile operation and in most cases they must be inferred from the operational and tactical levels, where observables are usually found. For this reason, this work has delved into the tactics and techniques of advanced threat actors to improve their characterization.

While tactics represent what a threat actor executes during an operation, techniques represent how tactics are performed. Most common tactics and their arrangement are consensuated among the cyber security industry and academia. In this context, it is mandatory to refer to MITRE ATT&CK as the main framework for the identification of threat actors' tactics and techniques, as this work has detailed. In addition, proposals such as the Cyber Kill Chain®, also discussed in previous sections, establish a common arrangement for tactics in hostile operations.

However, to be effective, tactics and techniques must be represented in a machine–readable format that can be loaded into security devices and automatically provide accurate results. I consider this is one of the biggest challenges we must face today. As I have discussed in chapter 6, different formats and languages have been developed in order to allow this specification, as well as the sharing of tactics and techniques in the form of actionable intelligence. However, the lack of a common standard is a current problem, as most of these formats are vendor–dependent. Without such a common standard, actionable intelligence is based nowadays mostly in atomic and computed indicators of compromise, easy to consume but with a very short time of life.

MITRE ATT&CK represents an enormous effort to characterize threat actors and their activities. However, this framework must be improved over time. In this sense, this work provides the dissection, in the form of taxonomy, of the techniques linked to different tactics. This dissection delves into the particularities of each tactic and allows analysts to establish a suitable classification for the associated techniques, thus improving not only previous characterization approaches but, which is most important, to identify and deploy appropriate countermeasures to face hostile activities. Particularly, three key tactics for advanced operations through cyberspace have been analyzed: delivery, persistence and impact. Both delivery and persistence are mandatory tactics in advanced cyberspace operations. If tactics are arranged in the form of kill–chain, these tactics are always performed in the early stages of the operation, so their knowledge and dissection are critical for an accurate detection capability.

Although it is not mandatory in all operations, Impact is a especially relevant tactic in which this work delves into. This tactic is linked to destructive and manipulative activities in Cyberspace Attack operations. Most cyberspace operations are related to cyber espionage, and few of them include impact techniques; in fact, in the context of cyber espionage, destruction or manipulation techniques are linked to the covering of the actors' tracks. However, the relevance of Computer Network Attack, or Cyberspace Attack, is dramatically growing during the last years. Particularly, with the rise of cyber physical systems, elements connected to Internet controlling physical processes, from nuclear plants to hospitals, the impact of these operations is very high, shutting down all kind of organizations and even whole countries.

### 7.3.1 Low level observables

Threat actors' tools and artifacts represent the lowest level of observables in an offensive cyber operation. Their associated indicators of compromise are usually atomic and computed ones, such as IP addresses, file hashes, mutexes or domain names. They are linked to tactical intelligence, and their exploitation allows the immediate detection of an operation. However, their time of life is short, as their related indicators are easy to modify for an advanced threat actor, as discussed in chapter 6. For example, a hash value associated with an implant can be changed at many steps of the generation of this particular implant, from source code to the executable itself (preprocessor, compiler, assembler, linker). Techniques such as domain fronting [208], in which the attacker abuses a trusted third party domain to hide the real identity of the command and control server, can easily hide the real command and control domain name or IP address. For this reason, although low level observables are usually a mandatory first line of analysis and defense, an important goal for all analysts is to rely not only on them, but also in the detection of the tactics and techniques of threat actors. The characterization of a threat can not be performed on a pure technical basis, this is, with focus on elements such as command and control infrastructures, malware analysis or communications' forensics. When possible, it is mandatory to identify, with a given probability, higher level elements such as goals or attribution.

In an offensive operation, advanced hostile actors employ different types of tools. Their arsenal can range from commercial or open–source applications to capabilities developed by the own threat actor, and even legitimate system utilities. This arsenal comprises all of the mandatory capabilities to complete the mission, from the reconnaissance step to the final actions on the target. The execution of pieces of the arsenal on the targeted infrastructure leaves traces on this infrastructure, and most of these traces represent the lowest level indicators of the compromise. Actions such as process execution, network connections or user access leaves traces that must be monitored in order not only to detect hostile operations, as detailed in section 7.4, but also to characterize threat actors, as the first line of observables.

When referring to offensive tools, a common term among literature is offensive cyber capabilities (OCC) [585] [584] or cyber weapons. In [522] Thomas Rid and Peter McBurney define cyber weapon as computer code that is used, or designed to be used, with the aim of threatening or causing physical, functional or mental harm to structures, systems or living beings. From this definition, cyber espionage tools and capabilities can not be considered cyber weapons. Most cyber weapons are designed to attack critical infrastructures [480], being Stuxnet the most well–known example of cyber weapon until now [146] [96]. Other examples include malware such as Petya [133] [16], WannaCry [133] or Sunburst [167]. Please note that all of these examples are destructive capabilities used in cyberspace attack operations, not in cyber espionage ones. Although cyberspace exploitation is an offensive discipline, the tools used in these operations are not considered cyber weapons. However, this work considers not only destructive capabilities, but the whole range of tools and artifacts used in both cyberspace attack and cyberspace exploitation operations.

Related to the vulnerabilities that advanced threat actors exploit, it is important to highlight that they rarely rely on zero–days and typically use public known vulnerabilities for their operations [177]. This is an interesting discussion, as in this context "advanced" does not mean to use technically advanced techniques, tools or exploits, but only that the actor has knowledge enough to decide if tools or exploits have to be "advanced", because the target is well–protected (hard target), or not, if the target is not hardened (soft target).

Not all hostile operations rely on malware to be successful. As stated in chapter 3, malwareless operations are those that do not use malicious code, but legitimate system utilities. These common utilities are abused by the threat actor in order to achieve its goals in the targeted infrastructure, such as lateral movement or data exfiltration. Although they do not leave artifacts in the environment, the usage of these legitimate tools leaves traces as well, traces that must be acquired and analyzed in order to detect the malicious activities. In addition, malwareless operations examples include not only the use of legitimate tools such as psexec to perform specific movements, but also the abuse of leaked credentials that grant access to a direct corporate server, where the threat actor is able to acquire the required information from its victim. In these cases, where the threat actor abuses legitimate credentials to remotely access webmails, remote desktops or virtual private networks of the targeted infrastructure, the leaved traces must be also acquired and analyzed to detect this malicious activity.

### 7.3.2 Tactics, techniques and procedures

This work has delved into the tactics and techniques of advanced threat actors. These tactics and techniques are a key element for the characterization of threat actors. While other observable elements, such as artifacts or tools, are easy to modify by the hostile actor, their tactics and techniques are harder to change, so they provide a more valuable intelligence for their characterization and detection. This intelligence, identified as operational, has a longer time of life than tactical intelligence, based on simple observables that are usually changed in every operation.

Tactics represent what a threat actor is doing, at the highest level of description, to accomplish a certain mission. In literature, they have been structured in frameworks such as MITRE ATT&CK and in different kill–chain models such as the Cyber Kill Chain®. Techniques specify how a tactic is implemented. From an intelligence point of view, their value is very high for the characterization of a threat actor, as well as for its detection. Finally, procedures are particular implementations of a given technique, linked to specific threat actors of even operators. Being so particular, procedures are not useful for a global detection of an offensive cyberspace operation, as in general terms they do not provide relevant information that is not provided by their superior techniques, so they are out of the scope of this work.

Tactics and techniques, operational intelligence, describe the modus operandi of a threat actor and they are a key element for its characterization, as they are not easily modified. However, we face a relevant problem when dealing with tactics and techniques: they are difficult to specify in a machine–readable format. For this reason, their sharing is not as common as the one of low level indicators, a fact that directly impacts on the detection of hostile operations. In section 7.4 I will delve into this problem and into the different approaches proposed in this work to address it.

Being tactics and techniques the key element for characterization, in this work the main techniques for well–known tactics linked to advanced threat actors have been analyzed. MITRE ATT&CK is the main industry framework for the structuring of tactics and techniques [262]. However, it has a relevant problem in its structure: the framework provides a plain alignment for all the techniques linked to a particular tactic. This fact impacts in the modeling of the activities of advanced threat actors, hindering both their detection and the implementation of countermeasures to face their activities. For this reason, this work has proposed a dissection and a taxonomy for techniques linked to different tactics, such as Delivery, Persistence or Impact (this one regarding destructive operations). Such taxonomies provide not only a valuable knowledge about advanced threat actors, but also an accurate characterization for them. The definition of taxonomies for techniques linked to other relevant tactics, such as Exfiltration or Command and Control, is an ongoing work.

In the case of Delivery, in chapter 2 an accurate taxonomy for delivery techniques, which allows the detection of novel techniques and the identification of appropriate countermeasures, is provided. Delivery is a key stage in offensive cyberspace operations, as it is the first moment in which a threat actor interacts with its victim in a hostile way. My proposal allows the accurate classification of techniques, overcoming the identified problems in current approaches and allowing both the discovery of new techniques and the detection of gaps in deployed countermeasures. It follows a logical structure that can be easy to expand and adapt, and it can be directly used in commonly accepted industry standards, such as MITRE ATT&CK.

After the Delivery, one of the main goals of an advanced threat actor is to maintain persistence over time. In chapter 3 a taxonomy for persistence techniques is proposed. This approach allows the detection of novel techniques, as well as the identification of appropriate countermeasures. A platform–agnostic model that structures persistence techniques through persistence points is provided. Persistence points are a novel concept introduced in this work as the core of the proposed taxonomy. The identification of these points is an useful tool for analysts, helping them to identify, classify and detect compromises and significantly reducing the amount of effort needed for these tasks. As with the Delivery tactic, my work follows a logical structure easy to expand and adapt, and which is aligned with MITRE ATT&CK as the main framework for threat actors' TTP identification.

Related to the Impact tactic, in chapter 4 destructive and control operations are analyzed and a global approach for these operations, that can be used to map real–world activities, is proposed. I identify the main Impact tactics (this is, degradation, disruption, destruction and manipulation) and, for each of them, a suitable structure for their associated techniques is proposed. In this way, a novel doctrine for these operations is defined, with two main goals: the modeling of advanced threat actors and, especially, the identification of proper countermeasures to face offensive cyberspace operations. This approach provides the basis for the identification of techniques in destructive and control operations. As with previous tactics, it is aligned with MITRE ATT&CK to ease its usage in a commonly accepted framework; in fact, it improves MITRE ATT&CK, given that this framework does not identify accurate tactics and techniques for destructive and control operations, grouping all of its techniques under a global umbrella tactic identified as "Impact".

The proposal for the structuring of offensive cyber operations' tactics and techniques, and the improvement of current approaches, significantly reduces the amount of effort needed to identify, analyze, and neutralize hostile activities from advanced threat actors. No previous global approaches to an accurate structure for techniques, aligned in all cases with industry standards such as MITRE ATT&CK, had been done until this moment. In this sense, it is mandatory to continue defining those structures, in particular for the rest of tactics that advanced threat actors develop in their operations.

### 7.3.3 Goals and strategy

When facing an operation, an advanced threat actor has a set of specific goals and a defined strategy or strategies to accomplish them. They represent what the hostile actor wants to perform on the target. It is important to highlight once again that advanced threats are not casual actors, and that they have clear and defined motivations for their operations, from strategic (geopolitical, diplomatic, military, etc.) to operational (especially economical).

The threat actor's goal represents a condition that the threat actor would like to achieve [681] [677], although it can not be always achieved. Its strategy defines, at high–level, how this goal can be achieved. A goal can be achieved through different parallel or sequential strategies, and both elements are non–observables: they can not be directly observed on the target infrastructure, so they must be inferred or acquired through external intelligence. In section 7.4.1 the role of this intelligence in the detection of threat actors' operations is discussed.

From an intelligence perspective, goals and strategy may not be shared with the threat actor's staff. The threat actor works at a tactical or operational level, while these features are at strategical one. This fact is especially relevant, and it is based on a need to know approach: few people know the whole picture of an operation, so if some information is leaked, the threat actor, or its sponsor, is not fully damaged. In cyberspace exploitation operations, and following the intelligence cycle exposed in section 1.4.3, from particular information needs as a starting point, a threat actor will plan a particular operation. This planning will include specific goals and defined strategies to accomplish these goals. The threat actor's operators will deploy particular techniques to acquire the mandatory information from the target, and once acquired this information will be sent to potentially different teams to process, analyze and disseminate the intelligence. Unfortunately, there is no equivalent to the intelligence cycle regarding other offensive information operations, such as cyberspace attack or psychological operations, so we must consider this point as an interesting research line.

### 7.3.4 Attribution

A relevant element for the characterization of advanced threat actors is attribution. Attribution is defined [662] [501] as the process to determine the identity or location of an attacker or an attacker's intermediary. The attribution process tries to answer a key question regarding cyberspace operations: the identity of the threat actor [521]. Attribution is particularly relevant for the characterization of a threat actor, not only to identify a culprit, but especially to determine the threat actor's interests. If we return to the Detection Maturity Level, exposed in 1.4.8, attribution is useful for the identification of goals and strategies of a threat actor, thus contextualizing targeted operations. And most important, it is useful for the prioritization of countermeasures [590], to formulate actions against the threat actor [453] [280] and even to create deterrence [422].

Attribution is one of the counterintelligence challenges in cyberspace operations [280] [590]. It is a complex task that faces many problems, both legal [94] [554] and technological [524]. From a legal point of view it is considered the most important practical obstacle to applying the law of *jus ad bellum* [576], this is, the criteria to be consulted before engaging in offensive actions, in order to determine whether they are permissible, so the response is a just one. The technical challenges are discussed in [272], [485] and specially in [590].

Attribution can be performed at three levels [366]: machines, human operators and the ultimately responsible party, a specific adversary. This last one is particularly important, as it implies legal, political, diplomatic or even military issues to a much greater degree. In [469] Timea Pahi and Florian Skopik propose the Cyber Attribution Model (CAM), a novel model to address the technical and socio–political sides of attribution. This model is based on the Diamond Model for Intrusion Analysis, exposed in section 1.4.8. It combines the cyber threat actor profiling with the cyber attack investigation indicators, in both cases from a technical and socio–political point of view. With this combination, the CAM proposes an approach to the attribution of cyberspace operations, including the analysis of false–flag ones.

To guide analysts towards a timely and accurate attribution, the US Office of the Director of National Intelligence (ODNI) identifies [453] five key indicators:

- **Tradecraft**. The behaviour of the actor, considered the most important indicator: habits are more difficult to change than technical tools.
- **Infrastructure**. The communication structures used during the attack, from delivery to command and control.
- **Malware**. Malicious software designed to enable unauthorized functions on a compromised system.
- **Intent**. The actor's commitment to carry out certain actions based on the context.

- **Indicators for external sources**. External information used by the Intelligence Community to get information or hypotheses about the perpetrator.

In all of the characterization elements, but especially in the attribution of hostile cyberspace operations, it is important to refer to probabilities. In section 1.4.3 Words of Estimative Probability were introduced, as an approach for the homogenization of terms. As this work has stated, in threat intelligence, as in intelligence, it is very difficult to have a complete certainty about a particular fact. From atomic indicators and artifacts, to strategies and goals, and of course tactics, techniques and procedures, we must consider elements with a given probability. In fact, it is especially relevant when dealing with observables: an IP address, a file hash or even a particular technique we are observing during an investigation is a fact, but we must always consider that an advanced threat may want us to watch these observables, so their usage must be carefully exploited.

Finally, it is important to highlight that although it is not a common tactic [590], nation–state actors can be engaged in false flag and no–flag cyberspace operations. A false flag operation is a diversionary or propaganda tactic of deceiving an adversary into thinking that an operation was carried out by another party, while a no–flag one is an operation conducted while being undeclared and when the operatives are either scantily marked or entirely unmarked [279] [484]. In both of them, the threat actor deceives or misguides attribution attempts, including the actor's origin, identity, movement, and exploitation [580]. To achieve this, the threat actor can distort from atomic indicators of compromise to tactics, techniques and procedures, and even goals and objectives. These operations can be used to support influence campaigns [84], such as TV5 Monde cyber attack, performed by APT28 facing as IS "Cyber Caliphate" [84] [152] [261], to launch or reinforce parallel conflicts, to generate noise as a diversionary tactic [466] or to test offensive capabilities [95].

## 7.4 Detection of Advanced Threat Actors

The detection of advanced threat actors' operations is not an easy task. In front of actors with low capabilities, whose operations are usually not well structured, and they can be detected and neutralized with easier countermeasures, and even influenced with more or less basic deterrence mechanisms [658], advanced threat actors must be faced with accurate technical and non–technical detection and response capabilities. However, the final goal for a defensive team is not the characterization of threats, but the detection and response to these threats. For this reason, characterization efforts and improvements are focused on this final goal, and they must provide better detection and response capabilities. The modeling of tactics and techniques, the identification of goals and strategies, and even attribution hypotheses, can be interesting discussions, but they have no real value if this modeling is not put in practice through detection capabilities in Security Operations Centers.

Nation–state actors usually perform stealth offensive cyberspace operations [122], both related to exploitation [669] [654] and attack [116] [214] actions. The operations of non–state actors on targeted infrastructures are less stealth than these of nation–state actors, although their arsenals and capabilities are growing in features to keep them unnoticed [460]. Due to this stealthiness, most countermeasures and security tools are not fully useful to face nation–state actors. Chapter 6 has delved into the identification and sharing of behavioral indicators of compromise. These indicators specify the tactics, techniques and procedures of advanced threat actors, and they are a key element for the detection of their hostile operations. As discussed, the detection of behavioral indicators is mandatory for an accurate detection capability, mainly because lower–level indicators, such as atomic or computed ones, are easy to change for an advanced threat actor, thus rendering detection schemes useless. As behavioral indicators are harder to modify for the actor, they have a longer time of life, so their usefulness is higher. However, one of the main conclusions of this work is that although these indicators are the most useful, they are the less shared ones. This is an interesting point: as there is not commonly–accepted standards to specify behavioral indicators of compromise, they are not actionable intelligence, so their sharing is limited. For this reason, I consider the inability to specify and share behavioral IOC in a common platform–independent format, one of the main problems for the detection of hostile activities, especially when dealing with sophisticated cyberspace operations.

This section, in addition to discuss the requirements for the detection of advanced threat actors and the role of intelligence and intelligence sharing, delves into the detection process. This process, as stated in chapter 5, is one of the main contributions of this work, and will allow defensive teams the identification and definition of mandatory tactics and techniques for an accurate detection. I have provided the mandatory tactics to accomplish this detection goal with a novel kill–chain model. This model defines and arranges the tactics that must be implemented in defensive centers. In addition, for each identified tactic different sub–tactics have been discussed and proposed. Further work should complete these sub tactics, as well as identify particular techniques to achieve each tactic or sub tactic. This model enables defensive teams, such as a SOC or a CSIRT, to accomplish their goals: this is, to detect and respond to threats. Although most works focus on the pure incident response approaches, I have considered all mandatory tactics to achieve the SOC goals at the same importance level. I defend that this approach is the MITRE ATT&CK equivalent for defensive teams, as previously discussed, so MITRE has been contacted in order to collaborate in the establishment of such a commonly–accepted defensive kill-chain.

### 7.4.1 The role of intelligence

The detection and the later analysis of an offensive cyberspace operation can be performed through all of the intelligence gathering disciplines that have been exposed in section 1.4.3, from HUMINT to MASINT. All of the information acquisition disciplines are relevant to identify features of an operation, from the strategical to the tactical ones. In this way, all of them are helpful to characterize the operation, and through this characterization all of them allow the detection of hostile activities. Especially, through operational and tactical intelligence specified as indicators of compromise (both low level ones, atomic and computed, and behavioral ones, tactics and techniques). This process is summarized in figure 7.4.

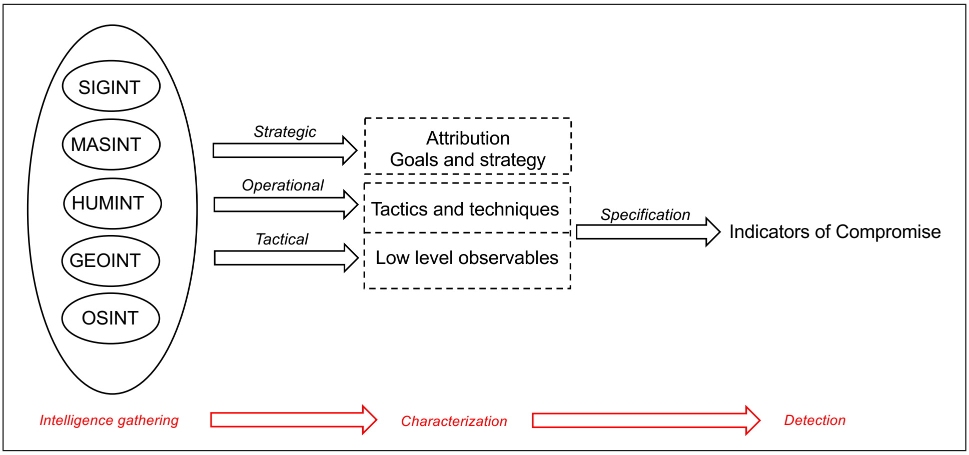

Figure 7.4: The role of information gathering disciplines in threat detection.

However, not all of these disciplines have the same weight on the detection equation. The main intelligence gathering discipline in cyber intelligence is SIGINT, recognized as the primary driver for operations within the cyberspace operating environment [212] [446]. In fact, many of the services or units historically focused on SIGINT activities are nowadays tasked with cyber operations, such as US NSA [372] [341], UK GCHQ [22] or IL IDF Unit 8200 [151]. Cyberspace has become the main way to communicate, and interception and gathering of network signals muddies the traditional notion of SIGINT [519]. Most detection approaches are based nowadays on SIGINT capabilities: this is, on the detection of anomalous activities in one's own infrastructure, through the monitoring of systems and networks. Usually, SIGINT provides tactics, techniques and procedures of implants communicating laterally and externally (C2 and exfiltration), as well as the relevant atomic indicators regarding these communications.

MASINT, specifically TECHINT, also plays a key role in the cyberspace domain. In the kinetic sphere, TECHINT refers to the collection and analysis of adversary's equipment and materiel; in cyberspace, media and software, particularly malware [200], are the equivalent to this equipment and materiel. TECHINT provides relevant information not only in the tactical level, but also in the operational and strategical, from the most technical indicators of compromise to aspects such as an adversary's budget or interest in its target.

HUMINT remains fundamental for understanding threats' capabilities and intentions [230] in cyberspace, not being replaced by any of the other acquisition disciplines. While non–human sources provide vast volumes of intelligence, HUMINT provides excellent (not vast, but excellent) information about adversaries. Different services have shown interest in deploying cover HUMINT capabilities, targeting units in hostile services or telecommunications industries (GCHQ Human Operations Team, HOT, could be an example [186]). In addition, overt capabilities among interest groups to get effective information sharing regarding cyber capabilities, interests or activities of potential adversaries is also a key element for HUMINT approaches [91]. An example of an overt cyber intelligence sharing effort could be the European Government CERT (EGC) group [290].

Although OSINT is a big player in cyber intelligence, its usefulness for the detection of hostile operations is limited. However, nowadays open source is usually a synonym for Internet source and OSINT is seen as monitoring, analysis and research of information coming from the Internet [351]. A global monitoring schema must include open source monitoring for the tracking of adversarial capabilities. This monitoring will be more useful for the prevention than for the detection of hostile activities: an eminent role is [475] the analysis of social opinion and sentiment. OSINT provides useful information about general trends that could be relevant to intelligence analysis, but as in classical intelligence, from an analytic perspective, one of the main problems to face in OSINT is the reliability of the source where information is gathered from [589] [236].

Finally, GEOINT regarding cyber is clear in military operations, but its role in cyber intelligence is usually not as relevant as the one of SIGINT or MASINT. GEOINT is mostly exploited in cyber Situational Awareness solutions [336] [680]. [39] states that cyberspace can be viewed as three layers (physical, logical, and social) made up of five components (geographic, physical network, logical network, cyber persona, and persona). The lowest of these three layers, the physical one, includes the geographic component, referring to the physical location of elements of the network and denoting a physical aspect tied to the rest of components. Being commonly accepted that information cannot exist without a physical infrastructure to support it, and that cyberspace has been created as a domain by this infrastructure and has a relevant geospatial component [605], there have been some efforts to “visualize” cyber using intelligence fusion and GEOINT techniques, trying to connect the “bits and the bytes” with the “bricks and mortar” [493]. To ensure this connection it is mandatory to geolocate network activity, tracking actions in both network time and space [212] towards cyber–physical spatialization in order to detect hostile operations.

As stated before, all information gathering disciplines are relevant for the detection and analysis of hostile activities. Although SIGINT, MASINT and HUMINT are the most exploited ones, all of them can provide strategical, operational and tactical intelligence. For this reason, an accurate security approach must consider all of these disciplines, not only for the pure detection but for the whole analysis and modeling of the threat actors' activities and interests. All of them provide the mandatory intelligence for the characterization of threat actors and their activities, so all of them can enable the detection of hostile activities in our infrastructures, as figure 7.4 summarizes.

### 7.4.2 Indicators of Compromise

Being SIGINT the main information gathering discipline for the detection of hostile activities, most of the current approaches to this detection rely on the specification and sharing of atomic and computed indicators of compromise. These indicators have a limited value and time of life, as they are easily modified by threat actors. As stated in chapter 6, for an effective detection capability it is mandatory to work at the operational intelligence level, this is, the one regarding tactics, techniques and procedures: behavioral indicators of compromise. For this reason I have delved into these tactics and techniques for advanced threat actors.

Nowadays, intelligence sharing, from strategical to tactical, is a must for threat detection, as in most cases we face global threats and there is a consensus that no intelligence actor can successfully act alone [312]. Collaboration between organizations is a key point to prevent, detect and neutralize threats. As an example, it is possible to refer to formalized information sharing and collaboration groups, such as FIRST or TF-CSIRT [337], US ISAC [394] or UK WARP [495]. Intelligence must be shared among a community, a group of trusted stakeholders who work together to address shared threats or vulnerabilities [667], usually with common interests; the formalized groups referenced below are examples of communities. Inside each type of community, elements such as the trust model or the sharing intelligence policy define how intelligence is shared.

To be considered valid threat intelligence, information shared must meet three requirements [164]: it must be relevant, actionable and valuable. As previously stated, nowadays most shared intelligence is in the form of low level data [476], especially atomic indicators [550]: this is, a very tactical approach that focuses on elements such as malicious IP addresses, DNS domains or URL. Operational, and even strategical, intelligence is much less shared, although it is more valuable than tactical one.

To share intelligence, nowadays it is mandatory to establish exchange mechanisms over a technological platform that can be deployed in many forms, for example as centralized or peer to peer exchange approaches. [550] states that there is no common definition of threat intelligence sharing platforms, being most of them focused on the exchange of tactical intelligence in STIX format. In fact, what we usually call threat intelligence sharing platforms, such as MISP, are focused on this kind of tactical and even operational intelligence, but they are not usually suitable for strategic intelligence sharing.

In this work the key requirements for the detection and sharing of behavioral indicators of compromise have been identified, as stated in chapter 6. One of the main problems I have identified during this research is the focus on atomic and computed indicators of compromise for the detection. These types of indicators are easily avoided by advanced threat actors, so to improve the success rate of detection capabilities we must deal with behavioral indicators of compromise. These indicators represent the tactics and techniques developed by threat actors, and although their value is much higher than the one of low level indicators, they are not widely used. For an accurate detection capability, the key requirements for defensive centers, such as SOC or CERT, have been identified. They are shown in table 7.2.

Table 7.2: Key requirements for TTP detection.

| **IC Stage** | **Key Requirements** |
|---|---|
| Acquisition | Acquire data from multiple, relevant sources<br>Acquire not only alerts, but regular events |
| Processing | Central data repository where relationships can be established<br>Common format for stored data<br>Long term retention |
| Analysis | Platform–agnostic implementation<br>Full native coverage for all techniques<br>Correlation of data from multiple sources<br>Comparison of correlated data against a reference |
| Dissemination | Machine readable and exportable format<br>Standard query language among providers |

These requirements provide the main directions for an accurate detection of advanced threat actors. However, they obviously do not guarantee success. In fact, with current technologies and capabilities, the only unfulfilled requirement for detecting and sharing behavioral indicators of compromise is the definition of a standard query language among providers. The absence of such a standard gives a competitive advantage to threat actors, as the inability to share knowledge about their tactics and techniques hardens their detection.

A relevant research line I have identified in this work is the quantification of the value of an indicator of compromise over time. This quality measurement is mandatory to determine a quantitative value for a given indicator at a specific time. Data quality is a especially relevant challenge in CTI sharing and exploitation [575] [411] [10] [335]. However, although most researchers agree on this fact, few work has been done to identify common metrics for a quality assessment of feeds. Quality assessment of threat intelligence data from a quantitative point of view, through specific metrics, is a complex problem [82]. In addition, it is also important to decay indicators and rotate them in security technologies, removing useless ones for efficiency reasons [288]. As the number of indicators, no matter their type, grows every day, security and network devices such as firewalls, intrusion detection systems or even proxies are slowed down by a such an amount of data to be processed [289].

### 7.4.3 The detection process

In addition to the identification of key requirements for the detection and sharing of behavioral indicators of compromise, in this work a common kill–chain model has been proposed. In this model, the mandatory actions to detect and enable the neutralization of threats are structured and arranged. Different kill–chain models have been developed from the threat actors' perspective; however, the detection process of an hostile operation must also have defined tactics and techniques in an arranged model and, until now, no such model had been defined. This lack causes not only problems such as unstructured approaches and conceptual errors but, what is most important, inefficiency in the detection and response to threats, as defensive tactics are not well identified. In this work, the SOC Critical Path (SCP) has been proposed. It is exposed in chapter 5 and shown in figure 7.5. SCP is defined as an arrangement of tactics that a SOC must perform to achieve effective detection. SCP is a technology–independent kill–chain model that provides an arrangement of mandatory steps, in the form of tactics, to be executed by Computer Network Defense teams to detect hostile cyber operations.

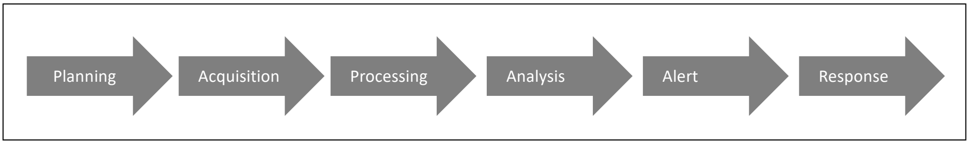

Figure 7.5: SOC Critical Path

Following the SOC Critical Path, to perform an accurate detection of advanced threat actors' activities it is mandatory to plan the activities to be executed on each step of the SCP. In this planning all the later steps of the cycle must be considered, as well as the identified requirements (see chapter 6) for each of the common ones. Please note that when facing advanced threat actors, it is especially relevant to focus on the detection of tactics, techniques and procedures, although atomic and computed indicators are always a must.

In the acquisition step, it is mandatory to identify the relevant data sources whose information will be useful for an accurate detection. In [175] the authors propose three families for data sources: network, payload and endpoint. As a data source, network acquisition provides the upper–level picture of what is signaling on the monitored infrastructure, from network flows [547] to protocol dissection [103]. This initial data source must be complemented with payloads, as they take one more step: payloads allow analysts to inspect the contents of traffic [644], thus allowing the detection and neutralization of threats before they arrive to the endpoint, which is the last data source. Endpoints can provide the most detailed picture of what is executed and stored on the monitored infrastructure. Tools such as sysmon [390] or Google Rapid Response [472] can acquire most of the relevant information to perform an accurate detection. In all of these data sources it is especially important to acquire the relevant information, and this relevant information includes the regular activities on all data sources. As advanced threat actors' activities are stealth in most cases, their detection is not possible relying only on alerts generated by security devices.

Once the relevant information has been acquired, Processing is identified as the next step of the SOC Critical Path. In this case, an accurate format for stored data and, of course, storage capabilities, are key elements of the processing tactic [537]. When dealing with advanced threat actors, we face stealth operations that extend over time. For this reason, to achieve their detection, it is mandatory to handle acquired information with a long–time retention period (in some cases, for months). In addition, the stealthiness of these operations involves a mandatory analysis of data acquired from different data sources, so data normalization is a must [34] [558] to establish relationships between information acquired from multiple data sources.

Although it is obviously a mandatory tactic, processing is not usually the core of the cyberspace defense operations, being in some cases even out of the scope of the SOC security team. Analysis is the next key tactic in the SOC Critical Path for an accurate detection and response. In this case, apart from the key requirements that have identified in chapter 6, it is especially important to highlight the need for automated analysis. The amount of data acquired and stored in previous steps is hard to handle for a human being. For this reason, automated misuse and anomaly detection approaches are put in place in all defensive schemes. Misuse intrusion detection systems are a mandatory countermeasure, but due to the nature of their stealth operations, they are not fully effective when dealing with advanced threat actors. Anomaly detection approaches are needed to identify compromises, in a process usually called threat hunting [67] [298]. In this sense, knowledge–based approaches to anomaly detection [204] and, specially, machine learning ones [258] are found. Machine learning approaches [258] have been identified as one of the most effective strategies to face cyberspace threats [564], as they can detect not only known threats and techniques, but also novel ones.

Once the analysis detects a potential incident, it is mandatory to raise an alert to be handled. This handling involves the expert examination of the alert and its associated data, in order to confirm if the alert is a false or true positive. One of the goals of a defensive team is to reduce the false positive rate [691] [434] [536], as each of these false positives introduces a needless cost and, most important, increases alert fatigue. The alert fatigue problem is caused by a large number of security alerts that finally causes security operators to become desensitized to these alerts [55]. As this is a relevant problem in most Security Operations Centers, it has been addressed through different approaches, from isolation forests [29] [30] [15] to data provenance analysis [263] or automatic context data aggregation to the raised alert [93].

An accurate response to the raised incident is the final step of the threat actors' detection process. The goal of this final tactic is the neutralization of the threat, which is in fact the goal of the whole detection process. As stated in chapter 5, although the pure incident response cycle is a well–known process, not all responses involve a whole incident response approach. This response will depend on multiple additional factors such as contractual and economic agreements, technical capabilities or service level agreements, and it is out of the scope of this work.

# Chapter 8

# Conclusions

The threat from advanced actors is a real fact. Being targeted by one of these actors is easy. Every organization with valuable information, every critical operator for basic services and even every single citizen is a potential target. We face two types of advanced threat actors: those linked to nation–states and those linked to criminal gangs. Both of them have the budget, the intent, the time and the capability to perform hostile activities. This is a growing trend that is expected to increase for years: cyberspace provides to threat actors enormous benefits, in front of classical operations.

Nation–state actors engaged in cyberspace operations are very active in cyber espionage and cyberspace attack campaigns. Among others, they target organizations with military, political, technological, environmental, diplomatic or scientific interest, as well as the infrastructures that support them. In addition, they are able to perform psychological operations, particularly influence ones, through cyberspace, targeting particular citizens. Definitely, nation–state actors benefit from cyberspace capabilities to perform all types of information operations through cyberspace, turning information into power: our data and infrastructures are a valuable asset.

In addition, in our daily basis we have to deal not only with these nation–state groups, but also with cyber crime actors that also target all kind of organizations, in this case with a pure economic goal: this is, to earn money. Their hostile operations are usually not as advanced as the ones linked to nation–state actors, but their complexity is growing day by day, as their revenue is very high. In fact, in the early stages of an incident handling, in many cases it is difficult to differentiate between nation–state actors and non–state ones: although their goals differ, most of their tactics and techniques are becoming similar day by day. Related to these actors, it is mandatory to highlight Human Operated Ransomware groups such as CONTI or HIVE. Their modus operandi is similar to that of cyber espionage operations: they target an organization, infiltrate it, move laterally and take control of the infrastructure. Once the target is fully or almost fully compromised, they exfiltrate the relevant information and, finally, they encrypt all data. At this point, they ask their target for a ransom for not to publish the exfiltrated data and to get the decryption key to recover the compromised information.

In this work, the categorization of advanced threat actors has been discussed, and I have delved into their characterization through the detailed analysis of their tactics and techniques. In addition, approaches to improve the detection of hostile activities through their analysis and through a novel kill–chain model for defensive tactics have been proposed. As stated in section 1, the main contributions of this work are as follows:

- To analyze and dissect key tactics for hostile operations performed by advanced threat actors, providing taxonomies for them. This contribution significantly improves not only the knowledge about advanced threat actors, but also the identification of their actions, thus providing accurate detection capabilities.
- To improve the main framework for the analysis of tactics and techniques of advanced threat actors, MITRE ATT&CK, easing not only the detection of these tactics and techniques but also the identification of hostile capabilities of the actors.
- To identify the key requirements for an accurate detection of hostile activities regarding advanced threat actors. The stealthiness of these activities makes them harder to detect, so approaches for an improved detection scheme following common intelligence models are proposed.
- To define the mandatory tactics to detect and respond to security incidents, providing a novel arrangement for them in the form of kill–chain model. This common set of tactics allows defensive centers the planning and continuous improvement of their work.

With the dissection of particular tactics, and with the identification of suitable taxonomies for their techniques, this work not only helps to improve the characterization of advanced threat actors, but also to establish accurate detection capabilities, thus improving global security for all kind of organizations. In addition, the proposed approaches are fully compatible with MITRE ATT&CK, so they can be directly used in this framework to provide an in–depth view of particular tactics. Of course, future research lines might include the equivalent dissection and taxonomy proposals for the rest of MITRE ATT&CK tactics.

This work on the detection of advanced threat actors definitely helps to improve the defensive capabilities of security teams. The novel kill–chain model that has been proposed provides a homogeneous detection process for all security centers, no matter their budget, sector or constituency are. This common process improves not only technical capabilities, but it also allows the detection of monitoring gaps and the measurement of the detection controls. This measurement is a must for the continuous improvement of quality of the detection scheme. Together with the identification of key requirements for the detection of TTP, I consider that these proposals help defensive centers to provide a globally improved detection approach, standardized, repeatable and measurable.

From a global point of view, this work helps analysts to categorize and characterize advanced threat actors, particularly through the in–depth analysis of their tactics and techniques. And most important, it helps analysts to detect hostile operations, as a first step to their neutralization. Improving the knowledge about the tactics and techniques of these actors is a step forward towards a better detection capability. In addition, with the requirements for an improved detection scheme, and the definition of the arranged tactics for detection and response, I consider this work as a relevant piece in Computer Network Defense. However, it is an ongoing work that must be completed with a detailed structure for other mandatory threat tactics, such as exploitation or exfiltration. Together with prevention mechanisms, out the scope of this thesis, we must achieve a global better protection not only for our information, but also for our lives.

As stated before, future research lines on the characterization of advanced threat actors include the dissection and taxonomy proposals for the tactics not covered in this work, mainly those identified in MITRE ATT&CK. Related to the detection of their activities, I identify the measurement of the quality of all kind of indicators of compromise as a key research line, as well as the definition and classification of techniques associated to the main defensive tactics I have proposed.

# Appendix A

# Acronyms

| | |
|---|---|
| APT | Advanced Persistent Threat |
| ATT&CK | Adversarial Tactics, Techniques, and Common Knowledge |
| CAPEC | Common Attack Pattern Enumeration and Classification |
| CERT | Computer Emergency Response Team |
| CMO | Civil–Military Operations |
| CNA | Computer Network Attack |
| CND | Computer Network Defense |
| CNE | Computer Network Exploitation |
| CNO | Computer Network Operations |
| CSIRT | Computer Security Incident Response Team |
| COI | Capability, Opportunity and Intent |
| COMINT | Communications Intelligence |
| COMCAM | Combat Camera |
| CI | Counterintelligence |
| CCI | Cyber Counterintelligence |
| CYBINT | Cyber Intelligence |
| CKC | Cyber Kill Chain |
| CTI | Cyber Threat Intelligence |
| CO | Cyberspace Operations |
| CVE | Common Vulnerabilities and Exposures |
| CVRF | Common Vulnerability Reporting Framework |

| | |
|---|---|
| CWE | Common Weaknesses Enumeration |
| CybOX | Cyber Observable eXpression |
| DMIA | Diamond Model for Intrusion Analysis |
| DML | Detection Maturity Level |
| EA | Electronic Attack |
| ELINT | Electronic Intelligence |
| ENISA | European Union Agency for Cybersecurity |
| EP | Electronic Protection |
| ES | Electronic Warfare Support |
| EW | Electronic Warfare |
| FISINT | Foreign Instrumentation Signals Intelligence |
| GEOINT | Geospatial Intelligence |
| HOR | Human Operated Ransomware |
| HUMINT | Human Intelligence |
| IA | Information Assurance |
| IC | Intelligence Cycle |
| ICS | Industrial Control Systems |
| IDS | Intrusion Detection System |
| IETF | Internet Engineering Task Force |
| IMINT | Imagery Intelligence |
| INSA | Intelligence and National Security Alliance |
| IOC | Indicator of Compromise |
| IODEF | Incident Object Description Exchange Format |
| IO | Information Operations |
| IT | Information Technologies |
| MASINT | Measures and Signatures Intelligence |
| MILDEC | Military Deception |
| MISO | Military Information Support Operations |
| MISP | Malware Information Sharing Platform |
| NATO | North Atlantic Treaty Organization |
| NIST | National Institute of Standards and Technology |
| NSA | National Security Agency |
| OCC | Offensive Cyber Capabilities |

| | |
|---|---|
| OSINT | Open Source Intelligence |
| OPSEC | Operations Security |
| OSVDB | Open Source Vulnerability DataBase |
| PA | Public Affairs |
| PLC | Programmable Logic Controller |
| PSYOP | Psychological Operations |
| SCO | STIX Cyber–observable Object |
| SCP | SOC Critical Path |
| SDO | STIX Domain Object |
| SOC | Security Operations Center |
| SIEM | Security Information Event Management |
| SIGINT | Signals Intelligence |
| STIX | Structured Threat Information Expression |
| TAXII | Trusted Automated Exchange of Intelligence Information |
| TECHINT | Technical Intelligence |
| TI | Threat Intelligence |
| TTP | Tactics, Techniques and Procedures |
| UEBA | User Behavior and Entity Analytics |
| WEP | Words of Estimative Probability |

# Bibliography

[1] ISO/IEC JTC 1/SC 27. Iso/iec 27035-1:2016. information technology — security techniques — information security incident management — part 1: Principles of incident management. Technical report, International Organization for Standardization, November 2016.

[2] Nurul Hidayah Ab Rahman and Kim-Kwang Raymond Choo. A survey of information security incident handling in the cloud. *computers & security*, 49:45–69, 2015.

[3] Shingo Abe, Yukako Uchida, Mitsutaka Hori, Yuichiro Hiraoka, and Shinichi Horata. Cyber threat information sharing system for industrial control system (ics). In *2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, pages 374–379. IEEE, 2018.

[4] Shingo Abe, Yukako Uchida, Mitsutaka Hori, Yuichiro Hiraoka, and Shinichi Horata. Cyber threat information sharing system for industrial control system (ics). In *2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, pages 374–379. IEEE, 2018.

[5] Marwan Abi-Antoun, Daniel Wang, and Peter Torr. Checking threat modeling data flow diagrams for implementation conformance and security. In *Proceedings of the twenty-second IEEE/ACM international conference on Automated software engineering*, pages 393–396, 2007.

[6] Mehmud Abliz. Internet denial of service attacks and defense mechanisms. *University of Pittsburgh, Department of Computer Science, Technical Report*, pages 1–50, 2011.

[7] Mohamed Abomhara and Geir M Køien. Cyber security and the internet of things: vulnerabilities, threats, intruders and attacks. *Journal of Cyber Security and Mobility*, pages 65–88, 2015.

[8] Natascha Abrek. Attack taxonomies and ontologies. In *Seminar Future Internet SS2014, Network Architectures and Services*, 2015.

[9] Md Sahrom Abu, Siti Rahayu Selamat, Aswami Ariffin, and Robiah Yusof. Cyber threat intelligence – issue and challenges. *Indonesian Journal of Electrical Engineering and Computer Science*, 10(1):371–379, 2018.

[10] Md Sahrom Abu, Siti Rahayu Selamat, Aswami Ariffin, and Robiah Yusof. Cyber threat intelligence–issue and challenges. *Indonesian Journal of Electrical Engineering and Computer Science*, 10(1):371–379, 2018.

[11] Sahrom Abu, Siti Rahayu Selamat, Robiah Yusof, and Aswami Ariffin. An enhancement of cyber threat intelligence framework. *Journal of Advanced Research in Dynamical and Control Systems*, 10:96–104, 2018.

[12] Maher Aburrous, M Alamgir Hossain, Keshav Dahal, and Fadi Thabtah. Experimental case studies for investigating e-banking phishing techniques and attack strategies. *Cognitive Computation*, 2(3):242–253, 2010.

[13] Muhammad I Adeka. *Cryptography and Computer Communications Security. Extending the Human Security Perimeter through a Web of Trust.* PhD thesis, University of Bradford, Bradford, UK, 2015.

[14] Atif Ahmad, Jeb Webb, Kevin C Desouza, and James Boorman. Strategically-motivated advanced persistent threat: Definition, process, tactics and a disinformation model of counterattack. *Computers & Security*, 86:402–418, 2019.

[15] Tariq Ahmed, Aayush Shah, Morarjee Kolla, and Ramadevi Yellasiri. Reduction of alert fatigue using extended isolation forest. In *2021 International Conference on Forensics, Analytics, Big Data, Security (FABS)*, volume 1, pages 1–5. IEEE, 2021.

[16] Jagmeet Singh Aidan, Harsh Kumar Verma, and Lalit Kumar Awasthi. Comprehensive survey on petya ransomware attack. In *2017 International Conference on Next Generation Computing and Information Systems (ICNGCIS)*, pages 122–125. IEEE, 2017.

[17] Maxat Akbanov, Vassilios G Vassilakis, and Michael D Logothetis. Wannacry ransomware: Analysis of infection, persistence, recovery prevention and propagation mechanisms. *Journal of Telecommunications and Information Technology*, pages 113–124, January 2019.

[18] Sm Al-Amin, Shipra Rani Sharkar, M Shamim Kaiser, and Milon Biswas. Towards a blockchain-based supply chain management for e-agro business system. In *Proceedings of International Conference on Trends in Computational and Cognitive Engineering*, pages 329–339, Online, 2021. Springer.

[19] Rawan Al-Shaer, Jonathan M Spring, and Eliana Christou. Learning the associations of mitre att&ck adversarial techniques. In *2020 IEEE Conference on Communications and Network Security (CNS)*, pages 1–9. IEEE, 2020.

[20] Marwan Albahar. Cyber attacks and terrorism: A twenty-first century conundrum. *Science and engineering ethics*, 25(4):993–1006, 2019.

[21] Hussain Aldawood and Geoffrey Skinner. An advanced taxonomy for social engineering attacks. *International Journal of Computer Applications*, 177(30):1–11, 2020.

[22] Richard J Aldrich. From sigint to cyber: a hundred years of britain's biggest intelligence agency. *Intelligence and National Security*, 36(6):910–917, 2021.

[23] Mohammed H Almeshekah. *Using deception to enhance security: A Taxonomy, Model, and Novel Uses.* PhD thesis, Purdue University, 2015.

[24] Mohammed H Almeshekah and Eugene H Spafford. Planning and integrating deception into computer security defenses. In *Proceedings of the 2014 New Security Paradigms Workshop*, pages 127–138. ACM, 2014.

[25] Hamad Almohannadi, Irfan Awan, Jassim Al Hamar, Andrea Cullen, Jules Pagan Disso, and Lorna Armitage. Cyber threat intelligence from honeypot data using elasticsearch. In *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*, pages 900–906. IEEE, 2018.

[26] Muhammad Turki Alshurideh, Barween Al Kurdi, Haitham M Alzoubi, Taher M Ghazal, Raed A Said, Ahmad Qasim AlHamad, Samer Hamadneh, Nizar Sahawneh, and Amer Hani Al-kassem. Fuzzy assisted human resource management for supply chain management issues. *Annals of Operations Research*, pages 1–19, 2022.

[27] Izzat Alsmadi. Cyber intelligence analysis. In *The NICE Cyber Security Framework*, pages 91–134. Springer, 2019.

[28] Ali Amanlou, Amir Abolfazl Suratgar, Jafar Tavoosi, Ardashir Mohammadzadeh, and Amir Mosavi. Single-image reflection removal using deep learning: A systematic review. *IEEE Access*, 2022.

[29] Muhamad Erza Aminanto, Lei Zhu, Tao Ban, Ryoichi Isawa, Takeshi Takahashi, and Daisuke Inoue. Automated threat-alert screening for battling alert fatigue with temporal isolation forest. In *2019 17th International Conference on Privacy, Security and Trust (PST)*, pages 1–3. IEEE, 2019.

[30] Muhamad Erza Aminanto, Lei Zhu, Tao Ban, Ryoichi Isawa, Takeshi Takahashi, and Daisuke Inoue. Combating threat-alert fatigue with online anomaly detection using isolation forest. In *International Conference on Neural Information Processing*, pages 756–765. Springer, 2019.

[31] Oxford Analytica. Kaseya ransomware attack underlines supply chain risks. Technical Report oxan-es, Oxford Analytica, Oxford, UK, July 2021.

[32] Oxford Analytica. Us pipeline hack to make ransomware risks a priority. *Emerald Expert Briefings*, May 2021.

[33] Oxford Analytica. Costa rica's new president faces immediate challenges. *Emerald Expert Briefings*, May 2022.

[34] Igor Anastasov and Danco Davcev. Siem implementation for global and distributed environments. In *2014 World Congress on Computer Applications and Information Systems (WCCAIS)*, pages 1–6. IEEE, 2014.

[35] Roberto Andrade, Jenny Torres, and Susana Cadena. Cognitive security for incident management process. In *International Conference on Information Technology & Systems*, pages 612–621, Cham, Switzerland, 2019. Springer.

[36] Uchenna Daniel Ani, Hongmei He, and Ashutosh Tiwari. Human factor security: evaluating the cybersecurity capacity of the industrial workforce. *Journal of Systems and Information Technology*, 2019.

[37] Claudia Aradau. Security that matters: Critical infrastructure and objects of protection. *Security dialogue*, 41(5):491–514, 2010.

[38] US Army. Field manual 2-0 intelligence, May 2004.

[39] US Army. Tradoc pamphlet 525-7-8. cyberspace operations concept capability plan 2016-2028. Technical report, United States Army, February 2010.

[40] John Arquilla and Mark Guzdial. The solarwinds hack, and a grand challenge for cs education. *Communications of the ACM*, 64(4):6–7, 2021.

[41] C Arslan and M Yanık. A new discipline of intelligence: Social media. *Military and Security Studies 2015*, page 69, 2015.

[42] Srinivas Arukonda and Samta Sinha. The innocent perpetrators: reflectors and reflection attacks. *Advances in Computer Science: an International Journal*, 4(1):94–98, 2015.

[43] Michael J Assante and Robert M Lee. The industrial control system cyber kill chain. Technical report, SANS Institute, October 2015.

[44] LV Astakhova and IA Medvedev. An information tool for increasing the resistance of employees of an organization to social engineering attacks. *Scientific and Technical Information Processing*, 48(1):15–20, 2021.

[45] Imran Awan. Cyber-extremism: Isis and the power of social media. *Society*, 54(2):138–149, 2017.

[46] Robert Axelrod and Rumen Iliev. Timing of cyber conflict. *Proceedings of the National Academy of Sciences*, 111(4):1298–1303, 2014.

[47] Gbadebo Ayoade, Swarup Chandra, Latifur Khan, Kevin Hamlen, and Bhavani Thuraisingham. Automated threat report classification over multi-source data. In *2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC)*, pages 236–245. IEEE, 2018.

[48] Abdul Haseeb Khan Babar and Yousaf Ali. Framework construction for augmentation of resilience in critical infrastructure: Developing countries a case in point. *Technology in Society*, 68:101809, 2022.

[49] Raj Badhwar. Commentary on insider threat. In *The CISO's Next Frontier*, pages 345–351. Springer, Berlin/Heidelberg, Germany, 2021.

[50] Ebrahim Bagheri and Ali A Ghorbani. The state of the art in critical infrastructure protection: a framework for convergence. *International Journal of Critical Infrastructures*, 4(3):215, 2008.

[51] Pooneh Nikkhah Bahrami, Ali Dehghantanha, Tooska Dargahi, Reza M Parizi, Kim-Kwang Raymond Choo, and Hamid HS Javadi. Cyber kill chain-based taxonomy of advanced persistent threat actors: analogy of tactics, techniques, and procedures. *Journal of Information Processing Systems*, 15(4):865–889, 2019.

[52] Şerif Bahtiyar, Mehmet Barış Yaman, and Can Yılmaz Altıniğne. A multi-dimensional machine learning approach to predict advanced malware. *Computer networks*, 160:118–129, 2019.

[53] James Ball, Julian Borger, Glenn Greenwald, et al. Revealed: how us and uk spy agencies defeat internet privacy and security. *Know Your Neighborhood*, 2013.

[54] George Bamford, John Felker, and Troy Mattern. Operational levels of cyber intelligence. *Cyber Intelligence Task Force, Intelligence and National Security Alliance*, September 2013.

[55] Tao Ban, Ndichu Samuel, Takeshi Takahashi, and Daisuke Inoue. Combat security alert fatigue with ai-assisted techniques. In *Cyber Security Experimentation and Test Workshop*, pages 9–16, 2021.

[56] Sean Barnum. Standardizing cyber threat intelligence information with the structured threat information expression (stix). *Mitre Corporation*, 11:1–22, 2012.

[57] Peter Barry and Patrick Crowley. *Modern Embedded Computing. Designing Connected, Pervasive, Media-Rich Systems.* Elsevier, 2012.

[58] Wilson Bautista. *Practical cyber intelligence: how action-based intelligence can be an effective response to incidents.* Packt Publishing Ltd, 2018.

[59] Philippe Beaucamps, Isabelle Gnaedig, and Jean-Yves Marion. Behavior abstraction in malware analysis. In *International Conference on Runtime Verification*, pages 168–182. Springer, 2010.

[60] Kristian Beckers, Leanid Krautsevich, and Artsiom Yautsiukhin. Analysis of social engineering threats with attack graphs. In *Data privacy management, autonomous spontaneous security, and security assurance*, pages 216–232. Springer, 2014.

[61] Zsolt Bederna, Zoltan Rajnai, and Tamas Szadeczky. Attacks against energy, water and other critical infrastructure in the eu. In *2020 IEEE 3rd International Conference and Workshop in Óbuda on Electrical and Power Engineering (CANDO-EPE)*, pages 000125–000130. IEEE, 2020.

[62] Richard Bejtlich. What is apt and what does it want? https://taosecurity.blogspot.com/2010/01/what-is-apt-and-what-does-it-want.html, January 2010. Accessed: 20190313.

[63] J. Bowyer Bell and Barton Whaley. *Cheating and deception.* Library of Congress, 1991.

[64] Mercy Bere, Fungai Bhunu-Shava, Attlee Gamundani, and Isaac Nhamu. How advanced persistent threats exploit humans. *International Journal of Computer Science Issues (IJCSI)*, 12(6):170, 2015.

[65] Daniel J Bernstein, Tanja Lange, and Ruben Niederhagen. Dual ec: A standardized back door. In *The New Codebreakers*, pages 256–281. Springer, 2016.

[66] Humairaa Bhaiyat and Siphesihle Sithungu. Cyberwarfare and its effects on critical infrastructure. In *International Conference on Cyber Warfare and Security*, volume 17, pages 536–543, 2022.

[67] Akashdeep Bhardwaj and Sam Goundar. A framework for effective threat hunting. *Network Security*, 2019(6):15–19, 2019.

[68] Ganesh D Bhatt. Knowledge management in organizations: examining the interaction between technologies, techniques, and people. *Journal of knowledge management*, 2001.

[69] Sandeep Bhatt, Pratyusa K Manadhata, and Loai Zomlot. The operational role of security information and event management systems. *IEEE security & Privacy*, 12(5):35–41, 2014.

[70] Wesam Bhaya and Mehdi Ebady Manaa. Review clustering mechanisms of distributed denial of service attacks. *Journal of Computer Science*, 10(10):2037, 2014.

[71] David Bianco. The pyramid of pain. *Enterprise Detection & Response*, 2013.

[72] Tatjana Bilevičienė and Eglė Bilevičiūtė. Dynamics of crimes against the security of electronic data and information systems, and its influence on the development of electronic business in lithuania. *Jurisprudencija*, 18(2):689–702, 2011.

[73] Matt Bishop and Carrie Gates. Defining the insider threat. In *Proceedings of the 4th annual workshop on Cyber security and information intelligence research: developing strategies to meet the cyber security and information intelligence challenges ahead*, pages 1–3, Oak Ridge, TN, USA, 2008.

[74] Paul Black, Iqbal Gondal, and Robert Layton. A survey of similarities in banking malware behaviours. *Computers & Security*, 77:756–772, 2018.

[75] Dan Blum. Institute resilience through detection, response, and recovery. In *Rational Cybersecurity for Business*, pages 259–295. Apress, Berkeley, California, USA, 2020.

[76] D Bodeau, C McCollum, and D Fox. Cyber threat modeling: Survey, assessment, and representative framework. *HSSEDI, The Mitre Corporation*, 2018.

[77] Matteo E Bonfanti. Cyber intelligence: In pursuit of a better understanding for an emerging practice. *Cyber, Intelligence, and Security*, 2(1):105–121, 2018.

[78] E Lincoln Bonner. Defending our satellites: the need for electronic warfare education and training. Technical report, Air Force Research Institute Maxwell AFB United States, 2015.

[79] Todd Booth and Karl Andersson. Network security of internet services: eliminate ddos reflection amplification attacks. *Journal of Internet Services and Information Security (JISIS)*, 5(3):58–79, 2015.

[80] Anne-Claire Boury-Brisset, Anissa Frini, and Réjean Lebrun. All-source information management and integration for improved collective intelligence production. Technical report, Defence Research and Development Canada Valcartier (Quebec), 2011.

[81] Jean-Ian Boutin. Gamaredon group grows its game. Technical report, ESET, June 2020.

[82] Xander Bouwman, Harm Griffioen, Jelle Egbers, Christian Doerr, Bram Klievink, and Michel van Eeten. A different cup of TI? the added value of commercial threat intelligence. In *29th USENIX Security Symposium (USENIX Security 20)*, pages 433–450, 2020.

[83] Brian M Bowen, Shlomo Hershkop, Angelos D Keromytis, and Salvatore J Stolfo. Baiting inside attackers using decoy documents. In *International Conference on Security and Privacy in Communication Systems*, pages 51–70, Turin, Italy, 3-5 June 2009. Springer.

[84] Pascal Brangetto and Matthijs A Veenendaal. Influence cyber operations: The use of cyberattacks in support of influence operations. In *2016 8th International Conference on Cyber Conflict (CyCon)*, pages 113–126. IEEE, 2016.

[85] Rory Breuk and Albert Spruyt. Integrating dma attacks in exploitation frameworks. *University of Amsterdam, Tech. Rep*, pages 2011–2012, 2012.

[86] Calvin Brierley, Jamie Pont, Budi Arief, David J Barnes, and Julio C Hernandez-Castro. Persistence in linux-based iot malware. In *25th Nordic Conference on Secure IT Systems (Nordsec)*, 2020.

[87] Richard G Brody, Sara Kern, and Kehinde Ogunade. An insider's look at the rise of nigerian 419 scams. *Journal of Financial Crime*, 2020.

[88] Siri Bromander, Audun Jøsang, and Martin Eian. Semantic cyberthreat modelling. In *STIDS*, pages 74–78, 2016.

[89] Judith M Brown, Steven Greenspan, and Robert Biddle. Incident response teams in it operations centers: the t-tocs model of team functionality. *Cognition, Technology & Work*, 18(4):695–716, 2016.

[90] Rebekah Brown and Robert M Lee. The evolution of cyber threat intelligence (cti): 2019 sans cti survey. *SANS Institute. February*, 2019.

[91] Sarah Brown, Joep Gommers, and Oscar Serrano. From cyber security information sharing to threat management. In *Proceedings of the 2nd ACM workshop on information sharing and collaborative security*, pages 43–49, 2015.

[92] Blake D Bryant and Hossein Saiedian. A novel kill-chain framework for remote security log analysis with siem software. *Computers & Security*, 67:198–210, 2017.

[93] Blake D Bryant and Hossein Saiedian. Improving siem alert metadata aggregation with a novel kill-chain based classification model. *Computers & Security*, 94:101817, 2020.

[94] Russell Buchan. Cyber attacks: unlawful uses of force or prohibited interventions? *Journal of Conflict and Security Law*, 17(2):212–227, 2012.

[95] Ben Buchanan and Michael Sulmeyer. Russia and cyber operations: Challenges and opportunities for the next us administration. *Carnegie Endowment for International Peace*, 3, 2016.

[96] Scott Steele Buchanan. *Cyber-Attacks to Industrial Control Systems since Stuxnet: A Systematic Review.* PhD thesis, Capitol Technology University, April 2022.

[97] David Buil-Gil, Fernando Miró-Llinares, Asier Moneva, Steven Kemp, and Nacho Díaz-Castaño. Cybercrime and shifts in opportunities during covid-19: a preliminary analysis in the uk. *European Societies*, 23(sup1):S47–S59, 2021.

[98] LA Bukowski and J Feliks. Fuzzy logic expert system for supply chain resilience modelling and simulation. *Journal of Polish Safety and Reliability Association*, 6:31–38, 2015.

[99] Jan-Willem Bullee, Lorena Montoya, Marianne Junger, and Pieter Hartel. Spear phishing in organisations explained. *Information & Computer Security*, (25):593–613, 2017.

[100] Eric W Burger, Michael D Goodman, Panos Kampanakis, and Kevin A Zhu. Taxonomy model for cyber threat intelligence information exchange technologies. In *Proceedings of the 2014 ACM Workshop on Information Sharing & Collaborative Security*, pages 51–60, 2014.

[101] T Bussa, Avivah Litan, and T Phillips. Market guide for user and entity behavior analytics. Technical report, Gartner, 2016.

[102] Andrei Bytes and Jianying Zhou. Post-exploitation and persistence techniques against programmable logic controller. In *International Conference on Applied Cryptography and Network Security*, pages 255–273. Springer, 2020.

[103] Olivier Cabana, Amr M Youssef, Mourad Debbabi, Bernard Lebel, Marthe Kassouf, Ribal Atallah, and Basile L Agba. Threat intelligence generation using network telescope data for industrial control systems. *IEEE Transactions on Information Forensics and Security*, 16:3355–3370, 2021.

[104] Patrick Cain and David Jevans. Extensions to the iodef-document class for reporting phishing. *RFC 5901*, 2010.

[105] Sergio Caltagirone, Andrew Pendergast, and Christopher Betz. The diamond model of intrusion analysis. Technical report, Center For Cyber Intelligence Analysis and Threat Research Hanover, 2013.

[106] Joan Calvet, Jessy Campos, and Thomas Dupuy. Visiting the bear den. Technical report, ESET, June 2016.

[107] Jeimy J Cano. El ransomware: una estrategia de desestabilización geopolítica. el caso de costa rica. *Global strategy reports*, (15):1, 2022.

[108] Terry L Capps. *The Evolution of Russian Cyber Tactics, Techniques, and Procedures.* PhD thesis, Utica University, 2022.

[109] Robert Cardillo. Geospatial intelligence (geoint) basic doctrine. Technical report, National System for Geospatial Intelligence, April 2018.

[110] Rodney Carlisle. *Encyclopedia of intelligence and counterintelligence.* Routledge, 2015.

[111] James E Cartwright and W James. Joint terminology for cyberspace operations. *Joint Chiefs of Staff (JCS) Memorandum*, 2010.

[112] Defense Use Case. Analysis of the cyber attack on the ukrainian power grid. *Electricity Information Sharing and Analysis Center (E-ISAC)*, 388, March 2016.

[113] Timothy Casey. Threat agent library helps identify information security risks. Technical report, Intel Corporation, 2007.

[114] Timothy Casey. Understanding cyber threat motivations to improve defense. Technical report, Intel Corporation, 2015.

[115] M Cayford, C Van Gulijk, and PHAJM van Gelder. All swept up: An initial classification of nsa surveillance technology. *Safety and Reliability: Methodology and Applications*, pages 643–650, 2014.

[116] Lorena Cazorla, Cristina Alcaraz, and Javier Lopez. Cyber stealth attacks in critical information infrastructures. *IEEE Systems Journal*, 12(2):1778–1792, 2016.

[117] CCN-CERT. Análisis del ransomware hive o hiveleaks. Technical report, Centro Criptológico Nacional, December 2021.

[118] Ahmet Cetinkaya, Hideaki Ishii, and Tomohisa Hayakawa. An overview on denial-of-service attacks in control systems: Attack models and security analyses. *Entropy*, 21(2):210, 2019.

[119] S Sibi Chakkaravarthy, D Sangeetha, and V Vaidehi. A survey on malware analysis and mitigation techniques. *Computer Science Review*, 32:1–23, 2019.

[120] Junaid Ahsenali Chaudhry and Robert G Rittenhouse. Phishing: classification and countermeasures. In *2015 7th International Conference on Multimedia, Computer Graphics and Broadcasting (MulGraB)*, pages 28–31. IEEE, 25–28 November 2015.

[121] Checkpoint. Naikon APT: cyber espionage reloaded. Technical report, Chekpoint, May 2020.

[122] Jim Chen and Alan Dinerman. On cyber dominance in modern warfare. In *European Conference on Cyber Warfare and Security*, page 52. Academic Conferences International Limited, 2016.

[123] Ping Chen, Lieven Desmet, and Christophe Huygens. A study on advanced persistent threats. In *IFIP International Conference on Communications and Multimedia Security*, pages 63–72, Aveiro, Portugal, 25-26 September 2014. Springer.

[124] Shuo Chen, Jun Xu, Emre Can Sezer, Prachi Gauriar, and Ravishankar K Iyer. Non-control-data attacks are realistic threats. In *USENIX Security Symposium*, volume 5, 2005.

[125] Thomas M Chen and Saeed Abu-Nimeh. Lessons from stuxnet. *Computer*, 44(4):91–93, 2011.

[126] Yi-Hsien Chen, Yen-Da Lin, Chung-Kuan Chen, Chin-Laung Lei, and Chun-Ying Huang. Construct macos cyber range for red/blue teams. In *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security*, pages 934–936, 2020.

[127] Kaouthar Chetioui, Birom Bah, Abderrahim Ouali Alami, and Ayoub Bahnasse. Overview of social engineering attacks on social networks. *Procedia Computer Science*, 198:656–661, 2022.

[128] Meghna Chhabra, B B Gupta, and Dr.Ammar Almomani. A novel solution to handle DDOS attack in MANET. *Journal of Information Security*, 04:165–179, 01 2013.

[129] Ramkumar Chinchani, Anusha Iyer, Hung Q Ngo, and Shambhu Upadhyaya. Towards a theory of insider threat assessment. In *2005 International Conference on Dependable Systems and Networks (DSN'05)*, pages 108–117, Yokohama, Japan, 28 June–1 July 2005. IEEE.

[130] Kim-Kwang Raymond Choo. The cyber threat landscape: Challenges and future research directions. *Computers & security*, 30(8):719–731, 2011.

[131] Thomas Chopitea. Threat modelling of hacktivist groups. organization, chain of command, and attack methods. Technical report, Chalmers University of Technology, August 2012.

[132] Michał Choraś, Rafał Kozik, Adam Flizikowski, Witold Hołubowicz, and Rafał Renk. Cyber threats impacting critical infrastructures. In *Managing the Complexity of Critical Infrastructures*, pages 139–161. Springer, Cham, 2016.

[133] Kristoffer Kjærgaard Christensen and Tobias Liebetrau. A new role for 'the public'? exploring cyber security controversies in the case of wannacry. *Intelligence and National Security*, 34(3):395–408, 2019.

[134] Paul Cichonski, Thomas Millar, Tim Grance, and Karen Scarfone. Computer security incident handling guide. Technical Report SP 800-61, National Institute of Standards and Technology, August 2012.

[135] Paul Cichonski, Tom Millar, Tim Grance, Karen Scarfone, et al. Computer security incident handling guide. *NIST Special Publication*, 800(61):1–147, 2012.

[136] CISA. Advanced persistent threat compromise of government agencies, critical infrastructure, and private sector organizations. Technical report, Cybersecurity and Infrastructure Security Agency, December 2020.

[137] Robert M. Clark. Perspectives on intelligence collection. *Journal of U.S. Intelligence Studies*, 20(2):47–53, 2013.

[138] Robert M. Clark and Peter C. Oleson. Intelligence in public literature. *Studies in Intelligence*, 60(1):81–96, March 2016.

[139] Robert M. Clark and Peter C. Oleson. Cyber intelligence. *Journal of U.S. Intelligence Studies*, 24(3):11–23, 2018.

[140] Richard Alan Clarke and Robert K Knake. *The fifth domain: Defending our country, our companies, and ourselves in the age of cyber threats*. Penguin Press, July 2019.

[141] Mike Cloppert. Security intelligence: Attacking the cyber kill chain. *SANS Computer Forensics*, 26, 2009.

[142] Daniel R. Coats. National intelligence strategy of the united states of america 2019. *Office of the Director of National Intelligence, Washington, DC*, 2019.

[143] Fred Cohen, Dave Lambert, Charles Preston, Nina Berry, Corbin Stewart, and Eric Thomas. A framework for deception. *National Security Issues in Science, Law, and Technology*, 2001.

[144] Silvia Colabianchi, Francesco Costantino, Giulio Di Gravio, Fabio Nonino, and Riccardo Patriarca. Discussing resilience in the context of cyber physical systems. *Computers & Industrial Engineering*, 160:107534, 2021.

[145] Zachary A Collier and Joseph Sarkis. The zero trust supply chain: Managing supply chain risk in the absence of trust. *International Journal of Production Research*, 59(11):3430–3445, 2021.

[146] Sean Collins and Stephen McCombie. Stuxnet: the emergence of a new cyber weapon and its implications. *Journal of Policing, Intelligence and Counter Terrorism*, 7(1):80–91, 2012.

[147] Carl Colwill. Human factors in information security: The insider threat–who can you trust these days? *Information security technical report*, 14(4):186–196, 2009.

[148] Mauro Conti, Tooska Dargahi, and Ali Dehghantanha. Cyber threat intelligence: challenges and opportunities. In *Cyber Threat Intelligence*, pages 1–6. Springer, 2018.

[149] Allan Cook, Helge Janicke, Richard Smith, and Leandros Maglaras. The industrial control system cyber defence triage process. *Computers & Security*, 70:467–481, 2017.

[150] Shaen Corbet and John W Goodell. The reputational contagion effects of ransomware attacks. *Finance Research Letters*, page 102715, 2022.

[151] Sean Cordey. The israeli unit 8200–an osint-based study: Trend analysis. Technical report, ETH Zurich, 2019.

[152] Gordon Corera. How france's tv5 was almost destroyed by 'russian hackers'. *BBC News*, October 2016.

[153] Aneta Coufalíková, Ivo Klaban, and Tomáš Šlajs. Complex strategy against supply chain attacks. In *2021 International Conference on Military Technologies (ICMT)*, pages 1–5, Brno, Czech Republic, 8-11 June 2021. IEEE.

[154] National Research Council, Mapping Science Committee, et al. *Priorities for GEOINT research at the National geospatial-intelligence agency.* National Academies Press, 2006.

[155] Jerry M Couretas. Cyber influence operations. In *An Introduction to Cyber Analysis and Targeting*, pages 57–89. Springer, 2022.

[156] Christopher Cox. Cyber capabilities and intent of terrorist forces. *Information Security Journal: A Global Perspective*, 24(1-3):31–38, 2015.

[157] Emanuele Cozzi, Mariano Graziano, Yanick Fratantonio, and Davide Balzarotti. Understanding linux malware. In *2018 IEEE symposium on security and privacy (SP)*, pages 161–175. IEEE, 2018.

[158] Matthew Crosston and Frank Valli. An intelligence civil war: Humint vs. techint. *Cyber, Intelligence, and Security*, 1(1):67–82, January 2017.

[159] Richard M Crowell. War in the information age: a primer for cyberspace operations in 21st century warfare. Technical report, Naval War College, 686 Cushing Road, Newport, RI, January 2010.

[160] Glenn Alexander Crowther. The cyber domain. *The cyber defense review*, 2(3):63–78, 2017.

[161] Conor Cunningham. A russian federation information warfare primer. *The Henry M. Jackson School of International Studies*, 2020.

[162] Christian Czosseck and Katharina Ziolkowski. State actors and their proxies in cyberspace. *Peacetime regime for state activities in cyberspace*, 1:1–3, 2013.

[163] Assaf Dahan. Operation cobalt kitty: A large-scale APT in Asia carried out by the oceanlotus group. Technical report, Cyber Reason, 2017.

[164] Henry Dalziel. *How to define and build an effective cyber threat intelligence capability.* Syngress, 2014.

[165] Roman Danyliw, Jan Meijer, Yuri Demchenko, et al. The incident object description exchange format. *IETF Request For Comments*, 5070, 2007.

[166] Tooska Dargahi, Ali Dehghantanha, Pooneh Nikkhah Bahrami, Mauro Conti, Giuseppe Bianchi, and Loris Benedetto. A cyber-kill-chain based taxonomy of crypto-ransomware features. *Journal of Computer Virology and Hacking Techniques*, 15(4):277–305, 2019.

[167] Pratim Datta. Hannibal at the gates: Cyberwarfare & the solarwinds sunburst hack. *Journal of Information Technology Teaching Cases*, page 2043886921993126, 2021.

[168] Mark de Bruijne, Michel van Eeten, Carlos Hernández Gañán, and Wolter Pieters. Towards a new cyber threat actor typology. Technical report, Delft University of Technology, 2017.

[169] Peter De Tender, David Rendon, and Samuel Erskine. *Pro Azure Governance and Security.A Comprehensive Guide to Azure Policy, Blueprints, Security Center, and Sentinel.* Springer, 2019.

[170] Jefatura del Estado. Spanish law 8/2011 (LPIC), April 2011.

[171] Konstantinos Demertzis, Nikos Tziritas, Panayiotis Kikiras, Salvador Llopis Sanchez, and Lazaros Iliadis. The next generation cognitive security operations center: adaptive analytic lambda architecture for efficient defense against adversarial attacks. *Big Data and Cognitive Computing*, 3(1):6, 2019.

[172] Dorothy E Denning. Assessing the computer network operations threat of foreign countries. Technical report, Naval Postgraduate School, Monterey, CA, August 2007.

[173] Dorothy E Denning. Stuxnet: What has changed? *Future Internet*, 4(3):672–687, 2012.

[174] Vasily Desnitsky, Igor Kotenko, and Danil Zakoldaev. Evaluation of resource exhaustion attacks against wireless mobile devices. *Electronics*, 8(5):500, 2019.

[175] Jeremy D'Hoinne and Lawrence Orans. Five styles of advanced threat defense. Technical report, Gartner, Inc., August 2013.

[176] Roberto Di Pietro, Simone Raponi, Maurantonio Caprolu, and Stefano Cresci. Critical infrastructure. In *New Dimensions of Information Warfare*, pages 157–196. Springer, 2021.

[177] Giorgio Di Tizio, Michele Armellini, and Fabio Massacci. Software updates strategies: a quantitative evaluation against advanced persistent threats. *arXiv preprint arXiv:2205.07759*, 2022.

[178] Alina Díaz-Curbelo, Rafael Alejandro Espin Andrade, and Ángel Manuel Gento Municio. The role of fuzzy logic to dealing with epistemic uncertainty in supply chain risk assessment: review standpoints. *International Journal of Fuzzy Systems*, 22(8):2769–2791, 2020.

[179] Hermann Dornhackl, Konstantin Kadletz, Robert Luh, and Paul Tavolato. Malicious behavior patterns. In *2014 IEEE 8th international symposium on service oriented system engineering*, pages 384–389, Oxford, UK, 7-11 April 2014. IEEE.

[180] Christos Douligeris and Aikaterini Mitrokotsa. Ddos attacks and defense mechanisms: a classification. In *Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology (IEEE Cat. No. 03EX795)*, pages 190–193. IEEE, 2003.

[181] Wenli Duo, MengChu Zhou, and Abdullah Abusorrah. A survey of cyber attacks on cyber physical systems: Recent advances and challenges. *IEEE/CAA Journal of Automatica Sinica*, 9(5):784–800, 2022.

[182] Thomas Dupuy and Matthieu Faou. Gelsemium. Technical report, ESET, June 2021.

[183] Petrus Duvenage, Victor Jaquire, and Sebastian von Solms. A selective literature review on cyber counterintelligence. In *ECCWS 2018 17th European Conference on Cyber Warfare and Security V2*, pages 137–145, 2018.

[184] Petrus Duvenage and Sebastian von Solms. Cyber counterintelligence: Back to the future. *Journal of Information Warfare*, 13(4):42–56, 2014.

[185] Petrus Duvenage and Sebastian von Solms. Putting counterintelligence in cyber counterintelligence: back to the future. In *13th European Conference on Cyber Warfare and Security ECCWS 2014*, page 70, 2014.

[186] Petrus Duvenage and Sebastian von Solms. Putting counterintelligence in cyber counterintelligence: back to the future. In *13th European Conference on Cyber Warfare and Security ECCWS-2014 The University of Piraeus Piraeus, Greece*, page 70, 2014.

[187] Petrus Duvenage, Sebastian von Solms, and Manuel Corregedor. The cyber counterintelligence process: A conceptual overview and theoretical proposition. In *Proceedings of the 14th European Conference on Cyber Warfare and Security 2015*, pages 42–51, 2015.

[188] Anita D'Amico and Kirsten Whitley. The real work of computer network defense analysts. In *VizSEC 2007*, pages 19–37. Springer, Berlin, Heidelberg, 2008.

[189] Peter Eder-Neuhauser, Tanja Zseby, Joachim Fabini, and Gernot Vormayr. Cyber attack models for smart grid environments. *Sustainable Energy, Grids and Networks*, 12:10–29, 2017.

[190] John Ehrman. What are we talking about when we talk about counterintelligence? *Studies in Intelligence*, 53(2):9, 2009.

[191] Samar H El-sherif, Rabab F Abdel-kader, and Rawya Y Rizk. Two-factor authentication scheme using one time password in cloud computing. In *International Conference on Advanced Intelligent Systems and Informatics*, pages 425–434. Springer, August 2018.

[192] Jamie M Ellis. Chinese cyber espionage: a complementary method to aid pla modernization. Technical report, Naval Postgraduate School Monterey CA, 2015.

[193] Adel S Elmaghraby and Michael M Losavio. Cyber security challenges in smart cities: Safety, security and privacy. *Journal of advanced research*, 5(4):491–497, 2014.

[194] ENISA. Threat landscape for supply chain attacks. Technical report, European Union Agency for Cybersecurity (ENISA), Athens, Greece, July 2021.

[195] Jared Ettinger. Cyber intelligence tradecraft report. the state of cyber intelligence practices in the united states. Technical report, Carnegie–Mellon University. Software Engineering Institute, 2019.

[196] Charles D Faint. Exploitation intelligence (exint) a new intelligence discipline? *American Intelligence Journal*, 29(1):65–69, 2011.

[197] Kevin Fairbanks and Simson Garfinkel. Column: Factors affecting data decay. *Journal of Digital Forensics, Security and Law*, 7(2):1, 2012.

[198] Courtney Falk and Tatiana Ringenberg. Classifying cyber threat actors using an ontological approach, 2022.

[199] Wenjun Fan, Kevin Lwakatare, and Rong Rong. Social engineering: Ie based model of human weakness for attack and defense investigations. *International Journal of Computer Network & Information Security*, 9(1):1–11, 2017.

[200] R Fanelli. On the role of malware analysis for technical intelligence in active cyber defense. *Journal of Information Warfare*, 14(2):69–81, 2015.

[201] Barbara Fast, Michael Johnson, and Dick Schaeffer. Cyber intelligence. setting the landscape for an emerging discipline. *Cyber Intelligence Task Force, Intelligence and National Security Alliance*, September 2011.

[202] FBI/NSA. Russian GRU 85th GTsSS deploys previously undisclosed drovorub malware. Technical report, National Security Agency, Federal Bureau of Investigation, August 2020.

[203] Jana Feldkamp. The rise of tiktok: The evolution of a social media platform during covid-19. In *Digital responses to covid-19*, pages 73–85. Springer, 2021.

[204] Gilberto Fernandes, Joel JPC Rodrigues, Luiz Fernando Carvalho, Jalal F Al-Muhtadi, and Mario Lemes Proença. A comprehensive survey on network anomaly detection. *Telecommunication Systems*, 70(3):447–489, 2019.

[205] Ed Ferrara, Christopher McCClean, Rick Holland, and Thayer Frechette. Security operations center (SOC) staffing. Technical report, Forrester Research, Inc., Cambridge, MA, USA, August 2013.

[206] Samantha V Feuer, Amy Zegart, and Andrew Grotto. *From the Shadows to the Front Page: State Use of Proxies for Cyber Operations*. PhD thesis, Freeman Spogli Institute for International Studies, Stanford University, May 2020.

[207] David P Fidler. Cyberspace, terrorism and international law. *Journal of Conflict and Security Law*, 21(3):475–493, 2016.

[208] David Fifield, Chang Lan, Rod Hynes, Percy Wegmann, and Vern Paxson. Blocking-resistant communication through domain fronting. *Proceedings on Privacy Enhancing Technologies*, 2015(2):46–64, 2015.

[209] FireEye. Insights into Iranian Cyber Espionage: APT33 Targets Aerospace and Energy Sectors and has Ties to Destructive Malware. https://www.fireeye.com/blog/threat-research/2017/09/apt33-insights-into-iranian-cyber-espionage.html, September 2017. Accessed: 20190303.

[210] FireEye. APT37 (REAPER). the overlooked north korean actor. Technical report, FireEye, Inc., Milpitas, CA, USA, 2018.

[211] Cik Feresa Mohd Foozy, Rabiah Ahmad, and Mohd Faizal Abdollah. Phishing detection taxonomy for mobile device. *International Journal of Computer Science Issues (IJCSI)*, 10(1):338–344, 2013.

[212] George Franz, Galen Kane, and Jeff Fair. Reshaping intelligence operations in the cyberspace domain. *The Cyber Defense Review*, 4(1):33–40, 2019.

[213] David French and William Casey. Fuzzy hashing techniques in applied malware analysis. *Results of SEI Line-Funded Exploratory New Starts Projects*, page 2, 2012.

[214] Paul J Frontera and Erick J Rodríguez-Seda. Network attacks on cyber–physical systems project-based learning activity. *IEEE Transactions on Education*, 64(2):110–116, 2020.

[215] Stephen M Furnell and Matthew J Warren. Computer hacking and cyber terrorism: The real threats in the new millennium? *Computers & Security*, 18(1):28–34, 1999.

[216] Maryam Gallab, Hafida Bouloiz, Youssef Lamrani Alaoui, and Mohamed Tkiouat. Risk assessment of maintenance activities using fuzzy logic. *Procedia computer science*, 148:226–235, 2019.

[217] Keertika Gangwar, Subhranshu Mohanty, and AK Mohapatra. Analysis and detection of ransomware through its delivery methods. In *International Conference on Recent Developments in Science, Engineering and Technology*, pages 353–362, Gurgaon, India, 13 October 2017. Springer.

[218] Hongbo Gao, Qingbao Li, Yu Zhu, Wei Wang, and Li Zhou. Research on the working mechanism of bootkit. In *2012 8th International Conference on Information Science and Digital Content Technology (ICIDT2012)*, volume 3, pages 476–479. IEEE, 2012.

[219] Faisal A Garba, Sahalu B Junaidu, I Ahmad, and MS Tekanyi. Proposed framework for effective detection and prediction of advanced persistent threats based on the cyber kill chain. *Scientific and Practical Cyber Security Journal*, 3(3), September 2018.

[220] Carrie Gates and Carol Taylor. Challenging the anomaly detection paradigm: a provocative discussion. In *Proceedings of the 2006 workshop on New security paradigms*, pages 21–29, 2006.

[221] Leonardo Gavaudan, Swann Legras, and Vèronique Ventos. Cyber range automation, a bedrock for ai applications. *Proceedings of the 28th C&ESAR*, page 165, 2021.

[222] Zeljko Gavric and Dejan Simic. Overview of dos attacks on wireless sensor networks and experimental results for simulation of interference attacks. *Ingeniería e Investigación*, 38(1):130–138, 2018.

[223] Kenneth Geers. Coder, hacker, soldier, spy. In *Cyber Security: Analytics, Technology and Automation*, pages 73–87. Springer, Cham, Switzerland, 2015.

[224] Kenneth Geers, Darien Kindlund, Ned Moran, and Rob Rachwald. World war c: Understanding nation-state motives behind today's advanced cyber attacks. Technical report, Fireeye, Inc., 2014.

[225] Anna Georgiadou, Spiros Mouzakitis, and Dimitris Askounis. Detecting insider threat via a cyber-security culture framework. *Journal of Computer Information Systems*, pages 1–11, 2021.

[226] Moti Geva, Amir Herzberg, and Yehoshua Gev. Bandwidth distributed denial of service: Attacks and defenses. *IEEE Security & Privacy*, 12(1):54–61, 2014.

[227] Ibrahim Ghafir and Vaclav Prenosil. Proposed approach for targeted attacks detection. In H. Sulaiman, M. Othman, Y. Rahim, and N. Pee, editors, *Advanced Computer and Communication Engineering Technology*, pages 73–80. Springer, Cham, Switzerland, 2016.

[228] Yumna Ghazi, Zahid Anwar, Rafia Mumtaz, Shahzad Saleem, and Ali Tahir. A supervised machine learning based approach for automatically extracting high-level threat intelligence from unstructured sources. In *2018 International Conference on Frontiers of Information Technology (FIT)*, pages 129–134. IEEE, 2018.

[229] Daniel Gibert, Carles Mateu, and Jordi Planes. The rise of machine learning for detection and classification of malware: Research developments, trends and challenges. *Journal of Network and Computer Applications*, 153:102526, 2020.

[230] David V Gioe. 'the more things change': Humint in the cyber age. In *The Palgrave handbook of security, risk and intelligence*, pages 213–227. Springer, 2017.

[231] Zane Gittins and Michael Soltys. Malware persistence mechanisms. *24th International Conference on Knowledge–Based and Intelligent Information & Engineering Systems*, 176:88–97, 2020.

[232] VO Gnatyuk. Analysis of «incident» definitions and its interpretation in cyberspace. *Ukrainian Scientific Journal of Information Security*, 19(3), 2013.

[233] Roy Godson. *Dirty tricks or trump cards: US covert action and counterintelligence.* Transaction Publishers, 2001.

[234] Chris Goettl. Prioritising risk for better efficiency and collaboration. *Computer Fraud & Security*, 2021(4):13–16, 2021.

[235] Jan Goldman. *The Central Intelligence Agency: An Encyclopedia of Covert Ops, Intelligence Gathering, and Spies [2 volumes]: An Encyclopedia of Covert Ops, Intelligence Gathering, and Spies.* ABC-CLIO, 2015.

[236] Seonghyeon Gong, Jaeik Cho, and Changhoon Lee. A reliability comparison method for osint validity analysis. *IEEE Transactions on Industrial Informatics*, 14(12):5428–5435, 2018.

[237] Seonghyeon Gong and Changhoon Lee. Cyber threat intelligence framework for incident response in an energy cloud platform. *Electronics*, 10(3):239, 2021.

[238] William E Gortney. Department of defense dictionary of military and associated terms. Technical report, US Department of Defense, 2012.

[239] Spanish Government. Ley 11/2002, de 6 de mayo, reguladora del centro nacional de inteligencia, May 2002.

[240] Fanni Zsuzsanna Gozon, Daniel Vaczi, and Edit Toth-Laufer. Fuzzy-based human factor centered cybersecurity risk assessment. In *2021 IEEE 19th International Symposium on Intelligent Systems and Informatics (SISY)*, pages 83–88, Subotica, Serbia, 16-18 September 2021. IEEE.

[241] David L Grange. Asymmetric warfare: old method, new concern. In *National Strategy Forum Review*, volume 9, 2000.

[242] Great. Carbanak APT. the great bank robbery. Technical report, Kaspersky Lab, February 2015.

[243] James A Green. *Cyber Warfare*. Taylor & Francis, 2015.

[244] Andy Greenberg. The untold story of notpetya, the most devastating cyberattack in history. *Wired, August*, 22, 2018.

[245] Andy Greenberg. *Sandworm: A new era of cyberwar and the hunt for the Kremlin's most dangerous hackers*. Anchor, 2019.

[246] Bernhard Grill. *Bootkits revisited; detecting, analysing and mitigating bootkit threats*. PhD thesis, Technischen Universitat Wien, June 2016.

[247] Samuel Grooby, Tooska Dargahi, and Ali Dehghantanha. Protecting iot and ics platforms against advanced persistent threat actors: Analysis of apt1, silent chollima and molerats. In *Handbook of big data and IoT security*, pages 225–255. Springer, 2019.

[248] Kanika Grover, Alvin Lim, and Qing Yang. Jamming and anti-jamming techniques in wireless networks: A survey. *International Journal of Ad Hoc and Ubiquitous Computing*, 17(4):197–215, 2014.

[249] Thaddeus T. Grugq. Counterintelligence for cyber defence. intelligence analysis enables better defences against threat actors. Technical report, Medium, September 2017.

[250] Fuad Guliyev. National intelligence estimate. the outlook for intelligence collection. *Journal of Azerbaijani Studies*, 2010.

[251] Attila Gulyás. Lazarus. the north korean hacker group. *STRATEGIES XXI: The Complex and Dynamic Nature of the Security Environment*, pages 75–83, December 2021.

[252] Mordechai Guri, Matan Monitz, and Yuval Elovici. Usbee: Air-gap covert-channel via electromagnetic emission from usb. In *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, pages 264–268. IEEE, 2016.

[253] Adam Hahn, Roshan K Thomas, Ivan Lozano, and Alvaro Cardenas. A multi-layered and kill-chain based security analysis framework for cyber-physical systems. *International Journal of Critical Infrastructure Protection*, 11:39–50, 2015.

[254] Basel Halak. Cist: A threat modelling approach for hardware supply chain security. In *Hardware Supply Chain Security*, pages 3–65. Springer, Berlin/Heidelberg, Germany, 2021.

[255] Basel Halak. *Hardware Supply Chain Security: Threat Modelling, Emerging Attacks and Countermeasures.* Springer Nature, Berlin/Heidelberg, Germany, 2021.

[256] Sara LN Hald and Jens M Pedersen. An updated taxonomy for characterizing hackers according to their threat properties. In *2012 14th International Conference on Advanced Communication Technology (ICACT)*, pages 81–86. IEEE, 2012.

[257] Xiao Han, Nizar Kheir, and Davide Balzarotti. Deception techniques in computer security: A research perspective. *ACM Computing Surveys (CSUR)*, 51(4):80, 2018.

[258] Anand Handa, Ashu Sharma, and Sandeep K Shukla. Machine learning in cybersecurity: A review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(4):e1306, 2019.

[259] Chris Harrington. Sharing indicators of compromise: An overview of standards and formats. *EMC Critical Incident Response Center*, 2013.

[260] Jacob Harrison, Navid Asadizanjani, and Mark Tehranipoor. On malicious implants in pcbs throughout the supply chain. *Integration*, 79:12–22, 2021.

[261] Heather A Harrison Dinniss. The threat of cyber terrorism and what international law should (try to) do about it. *Geo. J. Int'l Aff.*, 19:43, 2018.

[262] Wajih Ul Hassan, Adam Bates, and Daniel Marino. Tactical provenance analysis for endpoint detection and response systems. In *2020 IEEE Symposium on Security and Privacy (SP)*, pages 1172–1189. IEEE, 2020.

[263] Wajih Ul Hassan, Shengjian Guo, Ding Li, Zhengzhang Chen, Kangkook Jee, Zhichun Li, and Adam Bates. Nodoze: Combatting threat alert fatigue with automated provenance triage. In *network and distributed systems security symposium*, 2019.

[264] Amin Hassanzadeh and Robin Burkett. Samiit: Spiral attack model in iiot mapping security alerts to attack life cycle phases. In *5th International Symposium for ICS & SCADA Cyber Security Research 2018 5*, pages 11–20, 2018.

[265] Sarah Hawley, Ben Read, Cristiana Brafman-Kittner, Nalani Fraser, Andrew Thompson, Yuri Rozhansky, and Sanaz Yashar. Apt39: An iranian cyber espionage group focused on personal information. Technical report, Mandiant, January 2019.

[266] Kaoru Hayashi and Vicky Ray. Bisonal malware used in attacks against russia and south korea. Technical report, PaloAlto Networks, July 2018.

[267] Ryan Heartfield and George Loukas. A taxonomy of attacks and a survey of defence mechanisms for semantic social engineering attacks. *ACM Computing Surveys (CSUR)*, 48(3):1–39, 2015.

[268] Ryan Heartfield, George Loukas, Sanja Budimir, Anatolij Bezemskij, Johnny RJ Fontaine, Avgoustinos Filippoupolitis, and Etienne Roesch. A taxonomy of cyber-physical threats and impact in the smart home. *Computers & Security*, 78:398–428, 2018.

[269] Udo Helmbrecht. ENISA overview of cybersecurity and related terminology. Technical report, European Union Agency For Network and Information Security, September 2017.

[270] Adam Henschke. Terrorism and the internet of things: Cyber-terrorism as an emergent threat. In *Counter-Terrorism, Ethics and Technology*, pages 71–87. Springer, Cham, 2021.

[271] Michael Herman. *Intelligence power in peace and war.* Cambridge University Press, 1996.

[272] Hinne Hettema. Rationality constraints in cyber defense: incident handling, attribution and cyber threat intelligence. *Computers & Security*, 109:102396, 2021.

[273] Seoung-Pyo Hong, Chae-Ho Lim, and Hoon Jae Lee. Apt attack response system through am-hids. In *2021 23rd International Conference on Advanced Communication Technology (ICACT)*, pages 271–274. IEEE, 2021.

[274] Aicia Hope. North korean lazarus hacking group leverages supply chain attacks to distribute malware for cyber espionage. *CPO Magazine*, November 2021.

[275] Jaromir Horejsi, Daniel Lunghi, Cedric Pernet, and Fujisawa Kazuki. Earth akhlut: exploring the tools, tactics and procedures of an advanced threat actor operating a large infrastructure. In *VB 2020 localhost*, 30 September - 2 October 2020.

[276] Reyhaneh HosseiniNejad, Hamed HaddadPajouh, Ali Dehghantanha, and Reza M Parizi. A cyber kill chain based analysis of remote access trojans. In *Handbook of big data and iot security*, pages 273–299. Springer, New York, USA, 2019.

[277] Zuzana Hromcová. At commands, TOR–based communications: meet Attor, a fantasy creature and also a spy platform. Technical report, ESET, October 2019.

[278] Jiankun Hu, Hemanshu R Pota, and Song Guo. Taxonomy of attacks for agent-based smart grids. *IEEE Transactions on Parallel and Distributed Systems*, 25(7):1886–1895, 2013.

[279] Geraint Hughes. *The military's role in counterterrorism: examples and implications for liberal democracies.* Strategic Studies Institute, May 2011.

[280] Jeffrey Hunker, Robert Hutchinson, and Jonathan Margulies. Attribution of cyber attacks on process control systems. In *International Conference on Critical Infrastructure Protection*, pages 87–99. Springer, 2008.

[281] Jeffrey Hunker and Christian W Probst. Insiders and insider threats-an overview of definitions and mitigation techniques. *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, 2(1):4–27, 2011.

[282] Matthew M Hurley. For and from cyberspace: Conceptualizing cyber intelligence, surveillance, and reconnaissance. *Air and Space Power Journal*, 26(6):12–33, 2012.

[283] Ghaith Husari, Ehab Al-Shaer, Mohiuddin Ahmed, Bill Chu, and Xi Niu. Ttpdrill: Automatic and accurate extraction of threat actions from unstructured text of cti sources. In *Proceedings of the 33rd Annual Computer Security Applications Conference*, pages 103–115, 2017.

[284] Eric M Hutchins, Michael J Cloppert, and Rohan M Amin. Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. *Leading Issues in Information Warfare & Security Research*, 1(1):80, 2011.

[285] William Hutchinson and Matthew J Warren. The use of deception in systems. In *1st International Conference on Systems Thinking in Management*, 2000.

[286] Jun-ho Hwang and Tae-jin Lee. Study of static analysis and ensemble-based linux malware classification. *Journal of the Korea Institute of Information Security & Cryptology*, 29(6):1327–1337, 2019.

[287] Thomas S Hyslip. Cybercrime-as-a-service operations. *The Palgrave Handbook of International Cybercrime and Cyberdeviance*, pages 815–846, 2020.

[288] Andras Iklody, Gerard Wagener, Alexandre Dulaunoy, Sami Mokaddem, and Cynthia Wagner. Decaying indicators of compromise. *arXiv preprint arXiv:1803.11052*, 2018.

[289] Andras Iklody, Gerard Wagener, Alexandre Dulaunoy, Sami Mokaddem, and Cynthia Wagner. Decaying indicators of compromise. Technical report, CIRC.lu, March 2018.

[290] Luukas K Ilves, Timothy J Evans, Frank J Cilluffo, and Alec A Nadeau. European union and nato global cybersecurity challenges. *Prism*, 6(2):126–141, 2016.

[291] Hakan Inac and Ercan Oztemel. An assessment framework for the transformation of mobility 4.0 in smart cities. *Systems*, 10(1):1, 2021.

[292] Threat Intelligence. Apt28: A window into russia's cyber espionage operations. Technical report, FireEye, 2014.

[293] Dragoş Ionică, Nirvana Popescu, Decebal Popescu, and Florin Pop. Cyber defence capabilities in complex networks. In *Internet of Everything*, pages 217–231. Springer, 2018.

[294] Zafar Iqbal and Zahid Anwar. Ontology generation of advanced persistent threats and their automated analysis. *NUST Journal of Engineering Sciences*, 9(2):68–75, 2016.

[295] Koteswara Ivaturi and Lech Janczewski. A taxonomy for social engineering attacks. In *International Conference on Information Resources Management*, pages 1–12, Shenzhen, China, 26-27 June 2011. Centre for Information Technology, Organizations, and People.

[296] Eduardo Izycki and Eduardo Wallier Vianna. Critical infrastructure: A battlefield for cyber warfare? In *ICCWS 2021 16th International Conference on Cyber Warfare and Security*, page 454. Academic Conferences Limited, 2021.

[297] Grégoire Jacob, Hervé Debar, and Eric Filiol. Behavioral detection of malware: from a survey towards an established taxonomy. *Journal in computer Virology*, 4(3):251–266, 2008.

[298] Zahra Jadidi and Yi Lu. A threat hunting framework for industrial control systems. *IEEE Access*, 9:164118–164130, 2021.

[299] Sunakshi Jaitly, Harshit Malhotra, and Bharat Bhushan. Security vulnerabilities and countermeasures against jamming attacks in wireless sensor networks: A survey. In *2017 International Conference on Computer, Communications and Electronics (Comptelix)*, pages 559–564. IEEE, 2017.

[300] Sushil Jajodia, VS Subrahmanian, Vipin Swarup, and Cliff Wang. *Cyber deception*. Springer, 2016.

[301] Markus Jakobsson. Modeling and preventing phishing attacks. In *Financial Cryptography*, volume 5. Citeseer, 2005.

[302] Victor Jaquire and Sebastiaan von Solms. Towards a cyber counterintelligence maturity model. In *Proceedings of the 12th International Conference on Cyber Warfare and Security, AR Bryant & RF Mills (eds), Wright State University, Air Force Institute of Technology, Dayton, OH, US*, pages 432–40, 2017.

[303] Victor John Jaquire. *A framework for a cyber counterintelligence maturity model*. PhD thesis, University of Johannesburg (South Africa), 2018.

[304] Chris Johnson, Lee Badger, David Waltermire, Lulie Snyder, and Clem Skrorupka. NIST SP 800-150. Guide to Cyber Threat Information Sharing, October 2016.

[305] David EA Johnson and Newton Howard. Network intelligence: An emerging discipline. In *2012 European Intelligence and Security Informatics Conference*, pages 287–288. IEEE, 2012.

[306] Leighton R. Johnson. *Computer incident response and forensics team management: Conducting a successful incident response.* Syngress, 2013.

[307] Loch K Johnson. *National security intelligence.* John Wiley & Sons, 2017.

[308] Joint Chiefs of Staff. *Joint Publication 1-02. Department of Defense Dictionary of Military and Associated Terms.* Department of Defense, November 2010.

[309] Oscar Jonsson. The next front: the western balkans. *Hacks, Leaks and Disruptions: Russian Cyber Strategies*, pages 85–91, 2018.

[310] Jestin Joy, Anita John, and James Joy. Rootkit detection mechanism: a survey. In *International Conference on Parallel Distributed Computing Technologies and Applications*, pages 366–374. Springer, 2011.

[311] Khurum Nazir Junejo and Jonathan Goh. Behaviour-based attack detection and classification in cyber physical systems using machine learning. In *Proceedings of the 2nd ACM International Workshop on Cyber-Physical System Security*, pages 34–43, 2016.

[312] Jori Pascal Kalkman and Lotte Wieskamp. Cyber intelligence networks: A typology. *The International Journal of Intelligence, Security, and Public Affairs*, 21(1):4–24, 2019.

[313] Yusuke Kambara, Yoshinori Katayama, Takanori Oikawa, Kazuyoshi Furukawa, Satoru Torii, and Tetsuya Izu. Developing the analysis tool of cyber-attacks by using cti and attributes of organization. In *Workshops of the International Conference on Advanced Information Networking and Applications*, pages 673–682. Springer, 2019.

[314] Ibrahim Kamel and Hussam Juma. A lightweight data integrity scheme for sensor networks. *Sensors*, 11(4):4118–4136, 2011.

[315] Vitally Kamluk and Alexander Gostev. Adwind – a cross–platform RAT. Technical report, Kaspersky, February 2016.

[316] Ulises Leon Kandiko. Cyber intelligence: Reinventing the wheel. *Triarius. Prevention and Security Bulletin on Terrorism and the new threats*, 2:27, 2018.

[317] Sean Kanuck. Content as infrastructure. *The Cyber Defense Review*, 7(1):181–192, 2022.

[318] Da-Yu Kao and Shou-Ching Hsiao. The dynamic analysis of wannacry ransomware. In *2018 20th International Conference on Advanced Communication Technology (ICACT)*, pages 159–166. IEEE, 2018.

[319] G Jaspher Willsie Kathrine, Paradise Mercy Praise, A Amrutha Rose, and Eligious C Kalaivani. Variants of phishing attacks and their detection techniques. In *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, pages 255–259. IEEE, 2019.

[320] Puneet Kaur, Amandeep Dhir, Anushree Tandon, Ebtesam A Alzeiby, and Abeer Ahmed Abohassan. A systematic literature review on cyberstalking. an analysis of past achievements and future promises. *Technological Forecasting and Social Change*, 163:120426, 2021.

[321] Dimitrios Kavallieros, Georgios Germanos, and Nicholas Kolokotronis. Profiles of cyber-attackers and attacks. In *Cyber-Security Threats, Actors, and Dynamic Mitigation*, pages 1–26. CRC Press, 2021.

[322] Yuta Kazato, Yoshihide Nakagawa, and Yuichi Nakatani. Improving maliciousness estimation of indicator of compromise using graph convolutional networks. In *2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC)*, pages 1–7. IEEE, 2020.

[323] Noël Keijzer. The new generation of ransomware: an in depth study of ransomware-as-a-service. Master's thesis, University of Twente, June 2020.

[324] Steven Kemp, David Buil-Gil, Asier Moneva, Fernando Miró-Llinares, and Nacho Díaz-Castaño. Empty streets, busy internet: A time-series analysis of cybercrime and fraud trends during covid-19. *Journal of Contemporary Criminal Justice*, 37(4):480–501, 2021.

[325] Geers Kenneth. Cyberspace and the changing nature of warfare. *SC Media*, 2008.

[326] Sherman Kent. Words of estimative probability. Technical report, Central Intelligence Agency, 1964.

[327] Muhammad Salman Khan, Sana Siddiqui, and Ken Ferens. A cognitive and concurrent cyber kill chain model. In *Computer and Network Security Essentials*, pages 585–602. Springer, New York, USA, 2018.

[328] Raghad Khweiled, Mahmoud Jazzar, and Derar Eleyan. Cybercrimes during covid-19 pandemic. *International Journal of Information Engineering & Electronic Business*, 13(2), 2021.

[329] Hyeob Kim, HyukJun Kwon, and Kyung Kyu Kim. Modified cyber kill chain model for multimedia service environments. *Multimedia Tools and Applications*, 78(3):3153–3170, 2019.

[330] Ivan Kirillov, Desiree Beck, Penny Chase, and Robert Martin. Malware attribute enumeration and characterization. Technical report, 2011.

[331] Richard Kissel, Andrew Regenscheid, Matthew Scholl, and Kevin Stine. *Guidelines for media sanitization. NIST SP 800–88.* US Department of Commerce, National Institute of Standards and Technology, December 2014.

[332] Dennis Kiwia, Ali Dehghantanha, Kim-Kwang Raymond Choo, and Jim Slaughter. A cyber kill chain based taxonomy of banking trojans for evolutionary computational intelligence. *Journal of computational science*, 27:394–409, 2018.

[333] A Kiyuna and L Conyers. *Cyberwarfare Sourcebook*. Lulu. com, 2015.

[334] Bruce Klingner. North korean cyberattacks: A dangerous and evolving threat. Technical Report 247, The Heritage Foundation, September 2021.

[335] Faris Bugra Kokulu, Ananta Soneji, Tiffany Bao, Yan Shoshitaishvili, Ziming Zhao, Adam Doupé, and Gail-Joon Ahn. Matched and mismatched socs: A qualitative study on security operations center issues. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, pages 1955–1970, 2019.

[336] Andreas Kornmaier and Fabrice Jaouën. Beyond technical data-a more comprehensive situational awareness fed by available intelligence information. In *2014 6th International Conference On Cyber Conflict (CyCon 2014)*, pages 139–154. IEEE, 2014.

[337] Klaus-Peter Kossakowski. Computer security incident response team (csirt) services framework. Technical report, FIRST, November 2019.

[338] Igor Kotenko and Andrey Chechulin. Attack modeling and security evaluation in siem systems. *International Transactions on Systems Science and Applications*, 8:129–147, 2012.

[339] Igor Kotenko and Andrey Chechulin. Common framework for attack modeling and security evaluation in siem systems. In *2012 IEEE International Conference on Green Computing and Communications*, pages 94–101. IEEE, 2012.

[340] Simon Kramer and Julian C Bradfield. A general definition of malware. *Journal in computer virology*, 6(2):105–114, 2010.

[341] David S Kris. The nsa's new sigint annex. *Journal of National Security Law & Policy*, 2021.

[342] N Krithika. A study on wha (watering hole attack)–the most dangerous threat to the organisation. *International Journal of Innovations in Scientific and Engineering Research (IJISER)*, 4(8):196–198, 2017.

[343] Katharina Krombholz, Heidelinde Hobel, Markus Huber, and Edgar Weippl. Advanced social engineering attacks. *Journal of Information Security and applications*, 22:113–122, 2015.

[344] Siva Kumar and Ramesh Anbanandam. An integrated delphi–fuzzy logic approach for measuring supply chain resilience: an illustrative case from manufacturing industry. *Measuring Business Excellence*, (23):350–375, 2019.

[345] Sushil Kumar et al. An emerging threat fileless malware: a survey and research challenges. *Cybersecurity*, 3(1):1–12, 2020.

[346] Sivaraju Kuraku and Dinesh Kalla. Emotet malware—a banking credentials stealer. *Iosr J. Comput. Eng*, 22:31–41, 2020.

[347] Bum Jun Kwon, Virinchi Srinivas, Amol Deshpande, and Tudor Dumitraş. Catching worms, trojan horses and pups: Unsupervised detection of silent delivery campaigns. *arXiv preprint arXiv:1611.02787*, 2016.

[348] David Lacey, Paul Salmon, and Patrick Glancy. Taking the bait: a systems analysis of phishing attacks. *Procedia Manufacturing*, 3:1109–1116, 2015.

[349] Anthony Cheuk Tung Lai, Ken Wai Kin Wong, Johnny Tsz Wun Wong, Austin Tsz Wai Lau, Alan Po Lun Ho, Shuai Wang, and Jogesh K. Muppala. Backdoor investigation and incident response: From zero to profit. In *12th EAI International Conference on Digital Forensics and Cyber Crime*, 12 2021.

[350] Marc Laliberte. A twist on the cyber kill chain: Defending against a javascript malware attack. *Dark Reading*, September 2017.

[351] Dmytro Lande and Ellina Shnurko-Tabakova. Osint as a part of cyber defense system. *Theoretical and Applied Cybersecurity*, 1(1), 2019.

[352] Ralph Langner. Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Security & Privacy*, 9(3):49–51, 2011.

[353] Austin LaSota. The present and potential future of mac hardware implants. 2019.

[354] Mohd Nasrulddin Abd Latif, Nurul Ashykin Abd Aziz, Nik Syuhailah Nik Hussin, and Zuraimi Abdul Aziz. Cyber security in supply chain management: a systematic review. *LogForum*, 17(1), 2021.

[355] Shahid Latif, Zeba Idrees, Zil e Huma, and Jawad Ahmad. Blockchain technology for the industrial internet of things: A comprehensive survey on security challenges, architectures, applications, and future research directions. *Transactions on Emerging Telecommunications Technologies*, 32(11):e4337, 2021.

[356] GyungMin Lee, ShinWoo Shim, ByoungMo Cho, TaeKyu Kim, and Kyounggon Kim. Fileless cyberattacks: Analysis and classification. *ETRI Journal*, 2020.

[357] Martti Lehto. The ways, means and ends in cyber security strategies. In *Proceedings of the 12th European conference on information warfare and security*, pages 182–190, 2013.

[358] Martti Lehto. Cyber-attacks against critical infrastructure. In *Cyber Security*, pages 3–42. Springer, 2022.

[359] Yassine Lemmou, Jean-Louis Lanet, and El Mamoun Souidi. A behavioural in-depth analysis of ransomware infection. *IET Information Security*, 15(1):38–58, 2021.

[360] Rafał Leszczyna and Michał R Wróbel. Threat intelligence platform for the energy sector. *Software: Practice and Experience*, 49(8):1225–1254, 2019.

[361] Marc-Etienn M. Léveillé and Ignacio Sanmillan. A wild kobalos appears. tricksy linux malware goes after HPCs. Technical report, ESET, January 2021.

[362] James A Lewis. Cybersecurity and critical infrastructure protection. *Center for Strategic and International Studies*, 1:12, January 2006.

[363] Xiang Li, Yan Wen, Min Huan Huang, and Qiang Liu. An overview of bootkit attacking approaches. In *2011 Seventh International Conference on Mobile Ad-hoc and Sensor Networks*, pages 428–431. IEEE, 2011.

[364] Xuran Li, Hong-Ning Dai, Hao Wang, and Hong Xiao. On performance analysis of protective jamming schemes in wireless sensor networks. *Sensors*, 16(12):1987, 2016.

[365] Michael Ligh, Steven Adair, Blake Hartstein, and Matthew Richard. *Malware analyst's cookbook and DVD: tools and techniques for fighting malicious code*. Wiley Publishing, 2010.

[366] Herbert Lin. Attribution of malicious cyber incidents: From soup to nuts. *Journal of International Affairs*, 70(1):75–137, 2016.

[367] Herbert Lin and Trisha E Wyman. *Special Operations Forces and Cyber–Enabled Influence Operations*, chapter 24, pages 333–351. Center for Global Security Research. Lawrence Livermore National Laboratory, January 2021.

[368] Otto Lindström. Next generation security operations center. Technical report, Metropolia University of Applied Sciences, Vantaa, Finland, November 2018.

[369] Eric Lipton, David E Sanger, and Scott Shane. The perfect weapon: How russian cyberpower invaded the us. *The New York Times*, 13, 2016.

[370] Eric G Little and Galina L Rogova. An ontological analysis of threat and vulnerability. In *2006 9th International Conference on Information Fusion*, pages 1–8. IEEE, 2006.

[371] Liu Liu, Olivier De Vel, Qing-Long Han, Jun Zhang, and Yang Xiang. Detecting and preventing cyber insider threats: A survey. *IEEE Communications Surveys & Tutorials*, 20(2):1397–1417, 2018.

[372] Steven Loleski. From cold to cyber warriors: the origins and expansion of nsa's tailored access operations (tao) to shadow brokers. *Intelligence and National Security*, 34(1):112–128, 2019.

[373] M Ángeles López-Cabarcos, Domingo Ribeiro-Soriano, and Juan Piñeiro-Chousa. All that glitters is not gold. the rise of gaming in the covid-19 pandemic. *Journal of Innovation & Knowledge*, 5(4):289–296, 2020.

[374] George Loukas, Eirini Karapistoli, Emmanouil Panaousis, Panagiotis Sarigiannidis, Anatolij Bezemskij, and Tuan Vuong. A taxonomy and survey of

cyber-physical intrusion detection approaches for vehicles. *Ad Hoc Networks*, 84:124–147, 2019.

[375] George Loukas, Tuan Vuong, Ryan Heartfield, Georgia Sakellari, Yongpil Yoon, and Diane Gan. Cloud-based cyber-physical intrusion detection for vehicles using deep learning. *Ieee Access*, 6:3491–3508, 2017.

[376] Mark M Lowenthal. *Intelligence: From secrets to policy.* CQ press, 2019.

[377] Mark M Lowenthal and Robert M Clark. *The five disciplines of intelligence collection.* Sage, 2015.

[378] John Lowry, Rico Valdez, and Brad Wood. Adversary modeling to develop forensic observables. In *4th annual digital forensics research workshop*, pages 204–213, 2004.

[379] Kuan-Chu Lu, I-Hsien Liu, and Jung-Shian Li. A survey of the offensive and defensive in industrial control system. *Bulletin of Networking, Computing, Systems, and Software*, 11(1):1–6, 2022.

[380] Parushi Malhotra, Yashwant Singh, Pooja Anand, Deep Kumar Bangotra, Pradeep Kumar Singh, and Wei-Chiang Hong. Internet of things: Evolution, concerns and security challenges. *Sensors*, 21(5):1809, 2021.

[381] Jennifer Mankin. *Classification of malware persistence mechanisms using low-artifact disk instrumentation.* PhD thesis, Northeastern University, Boston, Massachusetts, September 2013.

[382] Steve Mansfield-Devine. Hacktivism: assessing the damage. *Network Security*, 2011(8):5–13, 2011.

[383] Steve Mansfield-Devine. Nation-state attacks: the escalating menace. *Network Security*, 2020(12):12–17, 2020.

[384] Louis Marinos. Enisa threat landscape 2013. overview of current and emerging cyber–threats. Technical report, ENISA, European Union Agency for network and information security, Athens, Greece, December 2013.

[385] Alex Matrosov. Uefi vulnerabilities classification focused on bios implant delivery. https://medium.com/@matrosov/uefi-vulnerabilities-classification-4897596e60af, January 2019.

[386] Alex Matrosov, Eugene Rodionov, and Sergey Bratus. *Rootkits and bootkits: reversing modern malware and next generation threats.* No Starch Press, 2019.

[387] Vasileios Mavroeidis and Siri Bromander. Cyber threat intelligence model: an evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence. In *2017 European Intelligence and Security Informatics Conference (EISIC)*, pages 91–98. IEEE, 2017.

[388] Vasileios Mavroeidis and Siri Bromander. Cyber threat intelligence model: An evaluation of taxonomies, sharing standards, and ontologies within cyber

threat intelligence. In *2017 European Intelligence and Security Informatics Conference (EISIC)*, pages 91–98. IEEE, 2017.

[389] Vasileios Mavroeidis, Ryan Hohimer, Tim Casey, and Audun Jesang. Threat actor type inference and characterization within cyber threat intelligence. In *2021 13th International Conference on Cyber Conflict (CyCon)*, pages 327–352. IEEE, 2021.

[390] Vasileios Mavroeidis and Audun Jøsang. Data-driven threat hunting using sysmon. In *Proceedings of the 2nd International Conference on Cryptography, Security and Privacy*, pages 82–88, 2018.

[391] Fernando Maymí, Robert Bixler, Randolph Jones, and Scott Lathrop. Towards a definition of cyberspace tactics, techniques and procedures. In *2017 IEEE International Conference on Big Data (Big Data)*, pages 4674–4679. IEEE, 2017.

[392] Brian M. Mazanec and Bradley A. Thayer. *Deterring Cyber Warfare. Bolstering Strategic Stability in Cyberspace.* Palgrave Macmillan, December 2014.

[393] Uche M Mbanaso and Eman S Dandaura. The cyberspace: Redefining a new world. *IOSR Journal of Computer Engineering*, 17(3):17–24, 2015.

[394] Charlie McCarthy, Kevin Harnett, Art Carter, and Cem Hatipoglu. Assessment of the information sharing and analysis center model. Technical report, National Academies Transportation Research Board, October 2014.

[395] Stephen McCombie. Threat actor oriented strategy: Knowing your enemy to better defend, detect and respond to cyber-attacks. *Journal of the Australian Institute of Professional Intelligence Officers*, 26(1):24–41, 2018.

[396] Gary McGraw and Greg Morrisett. Attacking malicious code: A report to the infosec research council. *IEEE software*, 17(5):33–41, 2000.

[397] Rob McMillan. Definition: threat intelligence. *Gartner*, 2013.

[398] Dan McWhorter. APT1: Exposing one of China's cyber espionage units. Technical Report APT1, Mandiant, 2013.

[399] Nancy R Mead, Forrest Shull, Krishnamurthy Vemuru, and Ole Villadsen. A hybrid threat modeling method. Technical Report CMU/SEI-2018-TN-002, Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, 2018.

[400] Ylona Meeuwenberg. *Threat intelligence sharing as part of supply chain management enhancing security.* PhD thesis, Eindhoven University of Technology, 2017.

[401] Yoram Meijaard, Peter-Paul Meiler, and Luca Allodi. Modelling disruptive apts targeting critical infrastructure using military theory. In *2021 IEEE Eu-*

*ropean Symposium on Security and Privacy Workshops (EuroS&PW)*, pages 178–190. IEEE, 2021.

[402] Per Håkon Meland, Yara Fareed Fahmy Bayoumy, and Guttorm Sindre. The ransomware-as-a-service economy within the darknet. *Computers & Security*, 92:101762, 2020.

[403] Kathryn Merrick, Medria Hardhienata, Kamran Shafi, and Jiankun Hu. A survey of game theoretic approaches to modelling decision-making in information warfare scenarios. *Future Internet*, 8(3):34, 2016.

[404] Ioan-Cosmin Mihai, Stefan Pruna, and Ionut-Daniel Barbu. Cyber kill chain analysis. *International Journal of Information Security & Cybercrime*, 3:37, 2014.

[405] John F Miller. Supply chain attack framework and attack patterns. Technical report, MITRE CORP MCLEAN VA, 2013.

[406] Natalia Miloslavskaya. Stream data analytics for network attacks prediction. *Procedia Computer Science*, 169:57–62, 2020.

[407] Jose David Mireles, Jin-Hee Cho, and Shouhuai Xu. Extracting attack narratives from traffic datasets. In *2016 International Conference on Cyber Conflict (CyCon US)*, pages 1–6. IEEE, 2016.

[408] Qublai K Ali Mirza, Martin Brown, Oliver Halling, Louie Shand, and Abu Alam. Ransomware analysis using cyber kill chain. In *2021 8th International Conference on Future Internet of Things and Cloud (FiCloud)*, pages 58–65, Rome, Italy, 23-25 August 2021. IEEE.

[409] Qublai Khan Ali Mirza, Ghulam Mohi-Ud-Din, and Irfan Awan. A cloud-based energy efficient system for enhancing the detection and prevention of modern malware. In *2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA)*, pages 754–761. IEEE, 2016.

[410] Rahul Mishra and Sudhanshu Kumar Jha. Survey on botnet detection techniques. In *Internet of Things and Its Applications*, pages 441–449. Springer, 2022.

[411] Aziz Mohaisen, Omar Al-Ibrahim, Charles Kamhoua, Kevin Kwiat, and Laurent Njilla. Rethinking information sharing for threat intelligence. In *Proceedings of the Fifth ACM/IEEE Workshop on Hot Topics in Web Systems and Technologies*, pages 1–7, 2017.

[412] Abhijit Mohanta and Anoop Saldanha. Persistence mechanisms. In *Malware Analysis and Detection Engineering*, pages 213–236. Springer, 2020.

[413] K A Monnappa. *Learning Malware Analysis: Explore the concepts, tools, and techniques to analyze and investigate Windows malware.* Packt Publishing Ltd, June 2018.

[414] Matthew Monte. *Network Attacks and Exploitation. A Framework.* John Wiley and sons, Hoboken, NJ, USA, July 2015.

[415] Matthew Monte. *Network Attacks and Exploitation. A Framework.* John Wiley and sons, July 2015.

[416] Daesung Moon, Hyungjin Im, Ikkyun Kim, and Jong Hyuk Park. Dtb-ids: an intrusion detection system based on decision tree using behavior analysis for preventing apt attacks. *The Journal of supercomputing*, 73(7):2881–2895, 2017.

[417] Daesung Moon, Hyungjin Im, Jae Dong Lee, and Jong Hyuk Park. Mlds: multi-layer defense system for preventing advanced persistent threats. *Symmetry*, 6(4):997–1010, 2014.

[418] Philipp Morgner, Stefan Pfennig, Dennis Salzner, and Zinaida Benenson. Malicious iot implants: Tampering with serial communication over the internet. In *International Symposium on Research in Attacks, Intrusions, and Defenses*, pages 535–555. Springer, 2018.

[419] John Moteff, Claudia Copeland, and John Fischer. Critical infrastructures: What makes an infrastructure critical? Technical report, Congressional Research Service. The Library of Congress, January 2003.

[420] John Moteff and Paul Parfomak. Critical infrastructure and key assets: definition and identification. Technical report, Congressional Research Service. The Library of Congress, October 2004.

[421] J Mtsweni, Muyowa Mutemwa, and Njabulo Mkhonto. Development of a cyber-threat intelligence-sharing model from big data sources. *Journal of Information Warfare*, 15(3):56–68, 2016.

[422] Milton Mueller, Karl Grindal, Brenden Kuerbis, and Farzaneh Badiei. Cyber attribution. *The Cyber Defense Review*, 4(1):107–122, 2019.

[423] Andrew Mumford. *Proxy warfare.* John Wiley & Sons, 2013.

[424] Joseph Muniz, Gary McIntyre, and Nadhem AlFardan. *Security operations center: Building, operating, and maintaining your SOC.* Cisco Press, 2015.

[425] Dennis M. Murphy. Information operations primer. fundamentals of information operations. Technical Report AY12, U.S. Army War College, Department of Military Strategy, Planning, and Operations, November 2011.

[426] Muyowa Mutemwa, Jabu Mtsweni, and Njabulo Mkhonto. Developing a cyber threat intelligence sharing platform for south african organisations. In *2017 Conference on Information Communication Technology and Society (ICTAS)*, pages 1–6. IEEE, 2017.

[427] Muyowa Mutemwa, Jabu Mtsweni, and Lukhanyo Zimba. Integrating a security operations centre with an organization's existing procedures, policies and information technology systems. In *2018 International Conference on*

*Intelligent and Innovative Computing Applications (ICONIC)*, pages 1–6. IEEE, 2018.

[428] Henry Mwiki, Tooska Dargahi, Ali Dehghantanha, and Kim-Kwang Raymond Choo. Analysis and triage of advanced hacking groups targeting western countries critical national infrastructure: Apt28, red october, and regin. In *Critical infrastructure security and resilience*, pages 221–244. Springer, 2019.

[429] Suvda Myagmar, Adam J Lee, and William Yurcik. Threat modeling as a basis for security requirements. In *Symposium on requirements engineering for information security (SREIS)*, volume 2005, pages 1–8, 2005.

[430] Lisa Myers. The practicality of the cyber kill chain approach to security. *CSO Online*, October 2013.

[431] Centro Criptológico Nacional. Ciberamenazas y tendencias 2021. Technical report, Centro Nacional de Inteligencia, October 2021.

[432] Ramin Nafisi and Andrea Lelli. Goldmax, goldfinder, and sibot: Analyzing NOBELIUM's layered persistence. https://www.microsoft.com/security/blog/2021/03/04/goldmax-goldfinder-sibot-analyzing-nobelium-malware/, March 2021.

[433] Anitta Patience Namanya, Andrea Cullen, Irfan U Awan, and Jules Pagna Disso. The world of malware: An overview. In *2018 IEEE 6th International Conference on Future Internet of Things and Cloud (FiCloud)*, pages 420–427. IEEE, 2018.

[434] Dipika Narsingyani and Ompriya Kale. Optimizing false positive in anomaly based intrusion detection using genetic algorithm. In *2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE)*, pages 72–77. IEEE, 2015.

[435] Hiran V Nath and Babu M Mehtre. Static malware analysis using machine learning methods. In *International Conference on Security in Computer Networks and Distributed Systems*, pages 440–450. Springer, 2014.

[436] David Nathans. *Designing and Building Security Operations Center*. Syngress, 2014.

[437] TM Navamani. A review on cryptocurrencies security. *Journal of Applied Security Research*, pages 1–21, 2021.

[438] NCSC. Reckless campaign of cyber attacks by russian military intelligence service exposed. https://www.ncsc.gov.uk/news/reckless-campaign-cyber-attacks-russian-military-intelligence-service-exposed, October 2018. Accessed: 20190314.

[439] Tjada Nelson and Houssain Kettani. Open source powershell-written post exploitation frameworks used by cyber espionage groups. In *2020 3rd Inter-*

*national Conference on Information and Computer Technologies (ICICT)*, pages 451–456, San Jose, CA, USA, 9-12 March 2020. IEEE.

[440] Krisztián Németh. *Detection of persistent rootkit components on embedded IoT devices.* PhD thesis, Budapest University of Technology and Economics, December 2020.

[441] Amirreza Niakanlahiji, Lida Safarnejad, Reginald Harper, and Bei-Tseng Chu. Iocminer: Automatic extraction of indicators of compromise from twitter. In *2019 IEEE International Conference on Big Data (Big Data)*, pages 4747–4754. IEEE, 2019.

[442] Umara Noor, Zahid Anwar, and Zahid Rashid. An association rule mining-based framework for profiling regularities in tactics techniques and procedures of cyber threat actors. In *2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)*, pages 1–6. IEEE, 2018.

[443] Aleksey Novokhrestov, Anton Konev, and Alexander Shelupanov. Model of threats to computer network software. *Symmetry*, 11(12):1506, 2019.

[444] Lydia Novoszel and Tina Wakolbinger. Meta-analysis of supply chain disruption research. In *Operations Research Forum*, volume 3, pages 1–25, Berlin/Heidelberg, Germany, 2022. Springer.

[445] Jordan Nuce, Jeremy Kennelly, Kimberly Goody, Andrew Moore, Alyssa Rahman, Matt Williams, Brendan McKeague, and Jared Wilson. Shining a light on darkside ransomware operations. *FireEye Blogs*, 2021.

[446] Jacob G Oakley. Cyber collection. In *Waging Cyber War*, pages 57–70. Springer, 2019.

[447] US Department of Defense. Joint publication 3-13. information operations. Technical report, US Department of Defense, November 2014.

[448] US Department of Defense. Joint publication 2-03. geospatial intelligence in joint operations. Technical report, US Department of Defense, July 2017.

[449] Joint Chiefs of Staff. Joint publication 2-0. joint intelligence, October 2013.

[450] Joint Chiefs of Staff. Joint Publication 3-12. Cyberspace Operations, June 2018.

[451] US Department of the Army. *FM 2-22.3. Human Intelligence Collector Operations.* US Army, September 2006.

[452] Office of the Chairman of the Joint Chiefs of Staff. *DOD Dictionary of Military and Associated Terms.* Department of Defense, January 2021.

[453] Office of the Director of National Intelligence. A guide to cyber attribution. Technical report, September 2018.

[454] NATO Standarization Office. *NATO glossary of terms and definitions (English and French).* NSO, 2018.

[455] Gavin O'Gorman and Geoff McDonald. Ransomware: A growing menace. Technical report, Symantec Corporation, 2012.

[456] Marc Ohm, Henrik Plate, Arnold Sykosch, and Michael Meier. Backstabber's knife collection: A review of open source software supply chain attacks. In *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, pages 23–43, Lisbon, Portugal, 24-26 June 2020. Springer.

[457] Kenneth Okereafor. *Cybersecurity in the COVID-19 Pandemic*. CRC Press, 2021.

[458] Ebenezer A Oladimeji, Sam Supakkul, and Lawrence Chung. Security threat modeling and analysis: A goal-oriented approach. In *Proc. of the 10th IASTED International Conference on Software Engineering and Applications (SEA 2006)*, pages 13–15, 2006.

[459] Idowu Dauda Oladipo, Muyideen AbdulRaheem, Joseph Bamidele Awotunde, Akash Kumar Bhoi, Emmanuel Abidemi Adeniyi, and Moses Kazeem Abiodun. Machine learning and deep learning algorithms for smart cities: A start-of-the-art review. *IoT and IoE Driven Smart Cities*, pages 143–162, 2022.

[460] Mohammad N Olaimat, Mohd Aizaini Maarof, and Bander Ali S Al-rimy. Ransomware anti-analysis and evasion techniques: A survey and research directions. In *2021 3rd international cyber resilience conference (CRC)*, pages 1–6. IEEE, 2021.

[461] Cyril Onwubiko. Cyber security operations centre: Security monitoring for protecting business and supporting cyber defense strategy. In *2015 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, pages 1–10. IEEE, 2015.

[462] Kris Oosthoek and Christian Doerr. Sok: Att&ck techniques and trends in windows malware. In *International Conference on Security and Privacy in Communication Systems*, pages 406–425. Springer, 2019.

[463] Hilarie Orman. Evil offspring-ransomware and crypto technology. *IEEE Internet Computing*, 20(5):89–94, 2016.

[464] Luz Ortiz, Hector Tillerias, Christian Chimbo, and Veronica Toaza. Impact on the video game industry during the covid-19 pandemic. *Athenea Engineering sciences journal*, 1(1):5–13, 2020.

[465] Opeyemi Osanaiye, Attahiru S Alfa, and Gerhard P Hancke. A statistical approach to detect jamming attacks in wireless sensor networks. *Sensors*, 18(6):1691, 2018.

[466] Rain Ottis. Analysis of the 2007 cyber attacks against estonia from the information warfare perspective. In *Proceedings of the 7th European Conference*

*on Information Warfare*, page 163. Academic Publishing Limited Reading, MA, 2008.

[467] Harun Oz, Ahmet Aris, Albert Levi, and A Selcuk Uluagac. A survey on ransomware: Evolution, taxonomy, and defense solutions. *ACM Computing Surveys (CSUR)*, 2021.

[468] Mike O'Leary. Malware and persistence. In *Cyber Operations*, pages 507–566. Springer, 2019.

[469] Timea Pahi and Florian Skopik. Cyber attribution 2.0: Capture the false flag. In *Proceedings of the 18th European Conference on Cyber Warfare and Security (ECCWS 2019)*, pages 338–345, 2019.

[470] Changwook Park, Hyunji Chung, Kwangseok Seo, and Sangjin Lee. Research on the classification model of similarity malware using fuzzy hash. *Journal of the Korea Institute of Information Security and Cryptology*, 22(6):1325–1336, 2012.

[471] Joshua Park. The lazarus group: The cybercrime syndicate financing the north korea state. *Harvard International Review*, 42(2):34–39, 2021.

[472] So-Hyun Park, Sun-Woo Yun, So-Eun Jeon, Na-Eun Park, Hye-Yeon Shim, Yu-Rim Lee, Sun-Jin Lee, Tae-Rim Park, Na-Yeon Shin, Min-Jin Kang, et al. Performance evaluation of open-source endpoint detection and response combining google rapid response and osquery for threat detection. *IEEE Access*, 10:20259–20269, 2022.

[473] Youngsam Park, Jackie Jones, Damon McCoy, Elaine Shi, and Markus Jakobsson. Scambaiter: Understanding targeted nigerian scams on craigslist. *system*, 1:2, 2014.

[474] Bimal Parmar. Protecting against spear-phishing. *Computer Fraud & Security*, 2012(1):8–11, 2012.

[475] Javier Pastor-Galindo, Pantaleone Nespoli, Félix Gómez Mármol, and Gregorio Martínez Pérez. The not yet exploited goldmine of osint: Opportunities, open challenges and future trends. *IEEE Access*, 8:10282–10304, 2020.

[476] P Pawlinski, Przemylaw Jaroszewski, Piotr Kijewski, Lukasz Siewierski, Pawel Jacewicz, Przemyslaw Zielony, and Radoslaw Zuber. Actionable information for security incident response. Technical report, European Union Agency for Network and Information Security, 2014.

[477] Sean Peisert, Bruce Schneier, Hamed Okhravi, Fabio Massacci, Terry Benzel, Carl Landwehr, Mohammad Mannan, Jelena Mirkovic, Atul Prakash, and James Bret Michael. Perspectives on the solarwinds incident. *IEEE Security and Privacy*, 19(2):7–13, 2021.

[478] Kai Peng, Meijun Li, Haojun Huang, Chen Wang, Shaohua Wan, and Kim-Kwang Raymond Choo. Security challenges and opportunities for smart

contracts in internet of things: A survey. *IEEE Internet of Things Journal*, 2021.

[479] Adam Pennington, Andy Applebaum, Katie Nickels, Tim Schulz, Blake Strom, and John Wunder. Getting started with att&ck. Technical report, The MITRE Corporation, 2019.

[480] Dale Peterson. Offensive cyber weapons: construction, development, and employment. *Journal of Strategic Studies*, 36(1):120–124, 2013.

[481] Pythagoras Petratos. Definition and importance of cyberintelligence: An introduction. *SSRN 1977061*, 2011.

[482] Mark Phythian. *Understanding the intelligence cycle*. Routledge, 2013.

[483] Heloise Pieterse, Martin Olivier, and Renier van Heerden. Detecting manipulated smartphone data on android and ios devices. In *International Information Security Conference*, pages 89–103. Springer, 2018.

[484] Mauno Pihelgas. Mitigating risks arising from false-flag and no-flag cyber attacks. Technical report, NATO Cooperative Cyber Defence Centre of Excellence, 2015.

[485] Nikolaos Pitropakis, Emmanouil Panaousis, Alkiviadis Giannakoulias, George Kalpakis, Rodrigo Diaz Rodriguez, and Panayiotis Sarigiannidis. An enhanced cyber attack attribution framework. In *International Conference on Trust and Privacy in Digital Business*, pages 213–228. Springer, 2018.

[486] Michael L Pittaro. Cyber stalking: An analysis of online harassment and intimidation. *International journal of cyber criminology*, 1(2):180–197, 2007.

[487] Dirk Pleiter, Sébastien Varrette, Ezhilmathi Krishnasamy, Enver Özdemir, and Michal Pilc. Security in an evolving european hpc ecosystem. Technical report, PRACE aisbl, 2021.

[488] Jordan J Plotnek and Jill Slay. Cyber terrorism: A homogenized taxonomy and definition. *Computers & Security*, 102:102145, 2021.

[489] Paul Pols. The unified kill chain: Designing a unified kill chain for analyzing, comparing and defending against cyber attacks. Technical report, Cyber Security Academy, 2595 AN The Hague, Netherlands, 2017.

[490] Navneet Kaur Popli and Anup Girdhar. Behavioural analysis of recent ransomwares and prediction of future attacks by polymorphic and metamorphic ransomware. In *Computational Intelligence: Theories, Applications and Future Directions-Volume II*, pages 65–80. Springer, 2019.

[491] Bernardi Pranggono and Abdullahi Arabo. Covid-19 pandemic cybersecurity issues. *Internet Technology Letters*, 4(2):247, 2021.

[492] Davy Preuveneers and Wouter Joosen. Sharing machine learning models as indicators of compromise for cyber threat intelligence. *Journal of Cybersecurity and Privacy*, 1(1):140–163, 2021.

[493] Douglas R. Price. A guide to cyber intelligence. *Journal of US Intelligence Studies*, 21(1):55–60, 2014.

[494] Christian W Probst, Jeffrey Hunker, Dieter Gollmann, and Matt Bishop. Aspects of insider threats. In *Insider Threats in Cyber Security*, pages 1–15. Springer, Berlin/Heidelberg, Germany, 2010.

[495] Tony Proctor. The development of warning, advice and reporting points (warps) in uk national infrastructure. In *International Workshop on Critical Information Infrastructures Security*, pages 164–174. Springer, 2011.

[496] Niels Provos, Panayiotis Mavrommatis, Moheeb Abu Rajab, and Fabian Monrose. All your iframes point to us. In *Proceedings of the 17th Conference on Security Symposium*, SS'08, page 1–15, USA, 2008. USENIX Association.

[497] Hank Prunckun. A grounded theory of counterintelligence. *American Intelligence Journal*, 29(2):6–15, 2011.

[498] Hank Prunckun. *Counterintelligence theory and practice.* Rowman & Littlefield, 2019.

[499] Harikrishnan Pushpakumar. *Understanding the threat landscape in e-government infrastructure for business enterprises.* PhD thesis, Delft University of Technology, September 2015.

[500] Attia Qamar, Ahmad Karim, and Victor Chang. Mobile malware attacks: Review, taxonomy & future directions. *Future Generation Computer Systems*, 97:887–909, 2019.

[501] Li Qiang, Yang Zeming, Liu Baoxu, Jiang Zhengwei, and Yan Jian. Framework of cyber attack attribution based on threat intelligence. In *Interoperability, Safety and Security in IoT*, pages 92–103. Springer, 2016.

[502] Shuwang Qin, Chao Zhang, Tao Zhao, Wei Tong, Qiliang Bao, and Yao Mao. Dynamic high-type interval type-2 fuzzy logic control for photoelectric tracking system. *Processes*, 10(3):562, 2022.

[503] Santiago Quintero-Bonilla and Angel Martín del Rey. A new proposal on the advanced persistent threat: a survey. *Applied Sciences*, 10(11):3874, 2020.

[504] Sabina Georgiana Radu. Comparative analysis of security operations centre architectures; proposals and architectural considerations for frameworks and operating models. In *International Conference for Information Technology and Communications*, pages 248–260, Bucharest, Romania, June 2016. Springer.

[505] Robert Radvanovsky and Allan McDougall. *Critical infrastructure: homeland security and emergency preparedness.* CRC Press, October 2018.

[506] Ashwin Ramaswamy. Detecting kernel rootkits. Technical report, Dartmouth College, NH, USA, September 2008.

[507] Andrew Ramsdale, Stavros Shiaeles, and Nicholas Kolokotronis. A comparative analysis of cyber-threat intelligence sources, formats and languages. *Electronics*, 9(5):824, 2020.

[508] Andrew Ramsdale, Stavros Shiaeles, and Nicholas Kolokotronis. A comparative analysis of cyber-threat intelligence sources, formats and languages. *Electronics*, 9(5):824, 2020.

[509] Mark A Randol. *Homeland security intelligence: Perceptions, statutory definitions, and approaches.* DIANE Publishing, 2010.

[510] Konstantinos Rantos, Arnolnt Spyros, Alexandros Papanikolaou, Antonios Kritsas, Christos Ilioudis, and Vasilios Katos. Interoperability challenges in the cybersecurity information sharing ecosystem. *Computers*, 9(1):18, 2020.

[511] Harsha Rao and S Selvakumar. A kernel space solution for the detection of android bootkit. In *ICT and Critical Infrastructure: Proceedings of the 48th Annual Convention of Computer Society of India-Vol I*, pages 703–711. Springer, 2014.

[512] Michael Raska. North korea's evolving cyber strategies: Continuity and change. *SIRIUS–Zeitschrift für Strategische Analysen*, 4(2):1–13, 2020.

[513] Justinas Rastenis, Simona Ramanauskaitė, Justinas Janulevičius, Antanas Čenys, Asta Slotkienė, and Kęstutis Pakrijauskas. E-mail-based phishing attack taxonomy. *Applied Sciences*, 10(7):2363, 2020.

[514] Andrew Rathmell. Cyber-terrorism: The shape of future conflict? *The RUSI Journal*, 142(5):40–45, 1997.

[515] Pratyush Raunak and Prabhakar Krishnan. Network detection of ransomware delivered by exploit kit. *ARPN Journal of Engineering and Applied Sciences*, 12(12):3885–3889, 2017.

[516] Melinda Reed, John F Miller, and Paul Popick. Supply chain attack patterns: Framework and catalog. *Office of the Deputy Assistant Secretary of Defense for Systems Engineering*, 2014.

[517] Kieran Rendall, Antonia Nisioti, and Alexios Mylonas. Towards a multi-layered phishing detection. *Sensors*, 20(16):4540, 2020.

[518] Doug Rhoades. Machine actionable indicators of compromise. In *2014 International Carnahan Conference on Security Technology (ICCST)*, pages 1–5. IEEE, 2014.

[519] Julian Richards. The cyber challenge for intelligence. In *Intelligence in the knowledge society. Proceedings of the XIXth International Conference*, pages 97–108, 2014.

[520] Jeffrey T Richelson. *The US intelligence community.* Routledge, 2018.

[521] Thomas Rid and Ben Buchanan. Attributing cyber attacks. *Journal of Strategic Studies*, 38(1-2):4–37, 2015.

[522] Thomas Rid and Peter McBurney. Cyber-weapons. *the RUSI Journal*, 157(1):6–13, 2012.

[523] Andreas Rieb, Tamara Gurschler, and Ulrike Lechner. A gamified approach to explore techniques of neutralization of threat actors in cybercrime. In *Annual Privacy Forum*, pages 87–103. Springer, 2017.

[524] Thea Riebe, André Kaufhold, Tarun Kumar, Thomas Reinhold, and Christian Reuter. Threat intelligence application for cyber attribution. *Science Peace Security '19*, page 56, 2019.

[525] Ryan Riley, Xuxian Jiang, and Dongyan Xu. Guest-transparent prevention of kernel rootkits with vmm-based memory shadowing. In *International Workshop on Recent Advances in Intrusion Detection*, pages 1–20. Springer, 2008.

[526] Tim Ring. Threat intelligence: why people don't share. *Computer Fraud & Security*, 2014(3):5–9, 2014.

[527] Michael Robinson, Kevin Jones, and Helge Janicke. Cyber warfare: Issues and challenges. *Computers & security*, 49:70–94, 2015.

[528] John Rollins and Clay Wilson. Terrorist capabilities for cyberattack: Overview and policy issues. Technical report, Library of Congress. Congressional Research Service, 2007.

[529] Matthew Rosenquist. Prioritizing information security risks with threat agent risk assessment. Technical report, Intel Corporation, 2009.

[530] Ronald S. Ross. Managing information security risk. organization, mission, and information system view. Technical report, National Institute of Standards and Technology, March 2011.

[531] Judith EY Rossebo, Frank Fransen, and Eric Luiijf. Including threat actor capability and motivation in risk assessment for smart grids. In *2016 Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids (CPSR-SG)*, pages 1–7. IEEE, 2016.

[532] Hossein Mohammadi Rouzbahani, Hadis Karimipour, Abolfazl Rahimnejad, Ali Dehghantanha, and Gautam Srivastava. Anomaly detection in cyber-physical systems using machine learning. In *Handbook of big data privacy*, pages 219–235. Springer, 2020.

[533] Julian L Rrushi. Scada protocol vulnerabilities. In *Critical Infrastructure Protection*, pages 150–176. Springer, 2012.

[534] Mark E Russinovich and Aaron Margosis. *Troubleshooting with the Windows Sysinternals Tools*. Microsoft Press, 2016.

[535] Mark E Russinovich, David A Solomon, and Alex Ionescu. *Windows® Internals*. O'Reilly Media, Inc, 2009.

[536] Desiree Sacher. Fingerpointing false positives: How to better integrate continuous improvement into security monitoring. *Digital threats: Research and practice*, 1(1):1–7, 2020.

[537] Igor Saenko and Igor Kotenko. Towards resilient and efficient big data storage: Evaluating a siem repository based on hdfs. In *2022 30th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, pages 290–297. IEEE, 2022.

[538] SAFECode. Tactical threat modeling. Technical report, Software Assurance Forum for Excelence in Code, 2017.

[539] Mirko Sailio, Outi-Marja Latvala, and Alexander Szanto. Cyber threat actors for the factory of the future. *Applied Sciences*, 10(12):4334, 2020.

[540] Takamichi Saito, Ryohei Watanabe, Shuta Kondo, Shota Sugawara, and Masahiro Yokoyama. A survey of prevention/mitigation against memory corruption attacks. In *2016 19th International Conference on Network-Based Information Systems (NBiS)*, pages 500–505. IEEE, 2016.

[541] Fatima Salahdine and Naima Kaabouch. Social engineering attacks: A survey. *Future Internet*, 11(4):89, 2019.

[542] Rodrigo Sandoval-Almazan and J Ramon Gil-Garcia. Towards cyberactivism 2.0? understanding the use of social media and other information technologies for political activism and social movements. *Government information quarterly*, 31(3):365–378, 2014.

[543] David E Sanger and Eric Schmitt. Spy agency consensus grows that russia hacked dnc. *New York Times*, 26, 2016.

[544] BN Sanjay, DC Rakshith, RB Akash, and Vinay V Hegde. An approach to detect fileless malware and defend its evasive mechanisms. In *2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS)*, pages 234–239. IEEE, 2018.

[545] Katia Santacà, Matteo Cristani, Marco Rocchetto, and Luca Viganò. A topological categorization of agents for the definition of attack states in multiagent systems. In *Multi-Agent Systems and Agreement Technologies*, pages 261–276. Springer, 2016.

[546] Igor Santos, Jaime Devesa, Felix Brezo, Javier Nieves, and Pablo Garcia Bringas. Opem: A static-dynamic approach for machine-learning-based malware detection. In *International joint conference CISIS'12-ICEUTE´ 12-SOCO´ 12 special sessions*, pages 271–280. Springer, 2013.

[547] Omar Santos. *Network Security with NetFlow and IPFIX: Big Data Analytics for Information Security*. Cisco Press, 2015.

[548] Arif Sari. Context-aware intelligent systems for fog computing environments for cyber-threat intelligence. In *Fog Computing*, pages 205–225. Springer, 2018.

[549] Niklas Särökaari. *Phishing attacks and mitigation tactics.* PhD thesis, University of Jyväskylä, Jyväskylä, Finland, 2020.

[550] Clemens Sauerwein, Christian Sillaber, Andrea Mussmann, and Ruth Breu. Threat intelligence sharing platforms: An exploratory study of software vendors and research perspectives. In *Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017)*, pages 837–851, 2017.

[551] Ben Saul and Kathleen Heath. Cyber terrorism and use of the internet for terrorist purposes. In *Research Handbook on International Law and Cyberspace.* Edward Elgar Publishing, 2021.

[552] Kimberly Saunders. *Open Source Information: A True Collection Discipline.* PhD thesis, Citeseer, 2000.

[553] Neetesh Saxena, Emma Hayes, Elisa Bertino, Patrick Ojo, Kim-Kwang Raymond Choo, and Pete Burnap. Impact and key challenges of insider threats on organizations and critical businesses. *Electronics*, 9(9):1460, 2020.

[554] Michael N Schmitt. *Tallinn manual 2.0 on the international law applicable to cyber operations.* Cambridge University Press, 2017.

[555] Janine Schmoldt. The rising power of cyber proxies. In *ECCWS 2021 20th European Conference on Cyber Warfare and Security*, page 369. Academic Conferences Inter Ltd, 2021.

[556] ETSI Technical Committee Cyber Security. Cyber. implementation of the network and information security (nis) directive. Technical Report ETSI TR 103 456, European Telecommunications Standards Institute, October 2017.

[557] Christopher Seedyk. Characterizing cyber intelligence as an all-source intelligence product. *DSIAC Journal*, 5(3), Summer 2018.

[558] S Sandeep Sekharan and Kamalanathan Kandasamy. Profiling siem tools and correlation engines for security analytics. In *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pages 717–721. IEEE, 2017.

[559] Vladimir Shakhov and Insoo Koo. Depletion-of-battery attack: Specificity, modelling and analysis. *Sensors*, 18(6):1849, 2018.

[560] Scott Shane. The fake americans russia created to influence the election. *The New York Times*, 7(09), 2017.

[561] Arushi Sharma, Ekta Gandotra, Divya Bansal, and Deepak Gupta. Malware capability assessment using fuzzy logic. *Cybernetics and Systems*, 50(4):323–338, 2019.

[562] Kalpana Sharma, MK Ghose, Deepak Kumar, Raja Peeyush Kumar Singh, and Vikas Kumar Pandey. A comparative study of various security approaches used in wireless sensor networks. *International journal of advanced science and technology*, 17(2):31–44, 2010.

[563] Kamran Shaukat, Talha Mahboob Alam, Ibrahim A Hameed, Wasim Ahmed Khan, Nadir Abbas, and Suhuai Luo. A review on security challenges in internet of things (iot). In *2021 26th International Conference on Automation and Computing (ICAC)*, pages 1–6. IEEE, 2021.

[564] Kamran Shaukat, Suhuai Luo, Shan Chen, and Dongxi Liu. Cyber threat detection using machine learning techniques: A performance evaluation perspective. In *2020 International Conference on Cyber Warfare and Security (ICCWS)*, pages 1–6. IEEE, 2020.

[565] Nataliya Shevchenko, Timothy A Chick, Paige O?riordan, Thomas Patrick Scanlon, and Carol Woody. Threat modeling: a summary of available methods. *no. July*, 2018.

[566] Adam Shostack. Experiences threat modeling at microsoft. *MODSEC@ MoDELS*, 2008, 2008.

[567] Adam Shostack. *Threat modeling: Designing for security.* John Wiley & Sons, 2014.

[568] Abram N Shulsky and Gary James Schmitt. *Silent warfare: understanding the world of intelligence.* Potomac Books, Inc., 2002.

[569] Mason Shuya. Russian cyber aggression and the new cold war. *Journal of Strategic Security*, 11(1):1–18, 2018.

[570] Murtaza A Siddiqi and Naveed Ghani. Critical analysis on advanced persistent threats. *International Journal of Computer Applications*, 141(13):46–50, 2016.

[571] Ragnhild Endresen Siedler. Hard power in cyberspace: CNA as a political means. In *2016 8th International Conference on Cyber Conflict (CyCon)*, pages 23–36. IEEE, 2016.

[572] Johan Sigholm. Non-state actors in cyberspace operations. *Journal of Military Studies*, 4(1):1–37, 2013.

[573] Johan Sigholm and Martin Bang. Towards offensive cyber counterintelligence: Adopting a target-centric view on advanced persistent threats. In *2013 European Intelligence and Security Informatics Conference*, pages 166–171. IEEE, 2013.

[574] Michael Sikorski and Andrew Honig. *Practical malware analysis: the hands-on guide to dissecting malicious software.* No Starch Press, 2012.

[575] Christian Sillaber, Clemens Sauerwein, Andrea Mussmann, and Ruth Breu. Data quality challenges and future research directions in threat intelligence sharing practice. In *Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security*, pages 65–70, 2016.

[576] Daniel B. Silver. Computer network attack as a use of force under article 2 (4) of the united nations charter. *International Law Studies*, 76(1):21, 2002.

[577] Andrew Simmonds, Peter Sandilands, and Louis Van Ekert. An ontology for network security attacks. In *Asian Applied Computing Conference*, pages 317–323. Springer, 2004.

[578] Jagsir Singh and Jaswinder Singh. A survey on machine learning-based malware detection in executable files. *Journal of Systems Architecture*, page 101861, 2020.

[579] Florian Skopik and Stefan Filip. Design principles for national cyber security sensor networks: Lessons learned from small-scale demonstrators. In *2019 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*, pages 1–8. IEEE, 2019.

[580] Florian Skopik and Timea Pahi. Under false flag: Using technical artifacts for cyber attack attribution. *Cybersecurity*, 3(1):1–20, 2020.

[581] Victor A Skormin, Douglas H Summerville, and James S Moronski. Detecting malicious codes by the presence of their "gene of self-replication". In *International Workshop on Mathematical Methods, Models, and Architectures for Computer Network Security*, pages 195–205, St. Petersburg, Russia, 21-23 September 2003. Springer.

[582] Joe Slowik. Anatomy of an attack: Detecting and defeating crashoverride. *VirusBulletin*, October 2018.

[583] Joseph Slowik. Evolution of ICS attacks and the prospects for future disruptive events. Technical report, Threat Intelligence Centre Dragos Inc, Hanover, MD, USA, 2019.

[584] Max Smeets. The strategic promise of offensive cyber operations. *Strategic Studies Quarterly*, 12(3):90–113, 2018.

[585] Max Smeets and Herbert S Lin. Offensive cyber capabilities: To what ends? In *2018 10th International Conference on Cyber Conflict (CyCon)*, pages 55–72. IEEE, 2018.

[586] Ron Smith and Scott Knight. Applying electronic warfare solutions to network security. *Canadian Military Journal*, 6(3):49–58, 2005.

[587] Aditya K Sood and Richard J Enbody. Targeted cyberattacks: a superset of advanced persistent threats. *IEEE security & privacy*, 11(1):54–61, 2012.

[588] Stephen Specht and Ruby Lee. Distributed denial of service: Taxonomies of attacks, tools, and countermeasures. In *Proceedings of the International Workshop on Security in Parallel and Distributed Systems*, pages 543–550, 01 2004.

[589] Robert David Steele. Open source intelligence. In *Handbook of intelligence studies*, pages 147–165. Routledge, 2007.

[590] Timo Steffens. *Attribution of Advanced Persistent Threats.* Springer, 2020.

[591] Alan N Steinberg. An approach to threat assessment. In *2005 7th International Conference on Information Fusion*, volume 2, pages 8–pp. IEEE, 2005.

[592] Alan N Steinberg. Foundations of situation and threat assessment. *Handbook of multisensor data fusion: theory and practice*, pages 437–501, 2009.

[593] Patrick Stewin and Iurii Bystrov. Understanding dma malware. In *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, pages 21–41. Springer, 2012.

[594] Ryan Stillions. The DML model. http://ryanstillions.blogspot.com/2014/04/the-dml-model_21.html, April 2014. Accessed: 20210529.

[595] Ryan Stillions. On ttps. http://ryanstillions.blogspot.com/2014/04/on-ttps.html, April 2014. Accessed: 20200306.

[596] Blake E Strom, Andy Applebaum, Doug P Miller, Kathryn C Nickels, Adam G Pennington, and Cody B Thomas. Mitre att&ck: Design and philosophy. Technical report, The MITRE Corporation, July 2018.

[597] Blake E Strom, Andy Applebaum, Douglas P Miller, Kathryn C Nickels, Adam G Pennington, and Cody B Thomas. Mitre att&ck: Design and philosophy. Technical report, MITRE, July 2018.

[598] Blake E Strom, Joseph A Battaglia, Michael S Kemmerer, William Kupersanin, Douglas P Miller, Craig Wampler, Sean M Whitley, and Ross D Wolf. Finding cyber threats with ATT&CK™-based analytics. Technical report, MITRE Technical Report MTR170202. The MITRE Corporation, 2017.

[599] T Subburaj and K Suthendran. Digitalwatering hole attack detection using sequential pattern. *Journal of Cyber Security and Mobility*, pages 1–12, 2018.

[600] Kannan Subramanian R, Dr Kumar Kattumannil, et al. Errm gap analysis & identification. In *Event-and Data-Centric Enterprise Risk-Adjusted Return Management*, pages 205–283. Springer, Berlin/Heidelberg, Germany, 2022.

[601] S Sullivan, Alessandro Brighente, S Kumar, and M Conti. 5g security challenges and solutions: A review by osi layers. *IEEE Access*, 2021.

[602] Lidia Prudente T., Eleazar Aguirre A., Alba F. Moreno, and Rubén J. García. Dos attacks flood techniques. *International Journal of Combinatorial Optimization Problems and Informatics*, 3(2):3–13, May–Aug 2012.

[603] Teresa M. Takai, Adolpho Tarasiuk Jr., Charles H. Romine, Ron Ross, et al. Guide for conducting risk assessments. Technical Report SP 800–30, National Institute of Standards and Technology, September 2012.

[604] Yuvraj Sanjayrao Takey, Sai Gopal Tatikayala, Satyanadha Sarma Samavedam, PR Lakshmi Eswari, and Mahesh Uttam Patil. Real time early multi stage attack detection. In *2021 7th International Conference on Ad-*

*vanced Computing and Communication Systems (ICACCS)*, volume 1, pages 283–290, Coimbatore, India, 19–20 March 2021. IEEE.

[605] Nenad Taneski, Aleksandar Petrovski, and Dimitar Bogatinov. Geography in geospatial intelligence-c4irs and cyber security. In *Security and crisis management–theory and practice*, pages 65–73, 2019.

[606] Qiang Tang and Moti Yung. Cliptography: post-snowden cryptography. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 2615–2616, 2017.

[607] Robert W Taylor, Eric J Fritsch, John Liederbach, Michael R Saylor, and William L Tafoya. *Cyber crime and cyber terrorism.* Pearson New York, NY, 2019.

[608] Stan A Taylor. The role of intelligence in national security. *Contemporary security studies*, pages 249–67, 2007.

[609] R Osman Tekes. *A common architecture for cyber offences and assaults-(organized advanced multi-vector persistent attack): Cyber war cyber intelligence, espionage, and subversion cyber crime.* PhD thesis, University of London. London, UK, 2011.

[610] ThaiCERT. Threat group cards: A threat actor encyclopedia. Technical report, Thailand Computer Security Incident Response Team, 2020.

[611] Eric C Thompson. Threat intelligence. In *Designing a HIPAA-Compliant Security Operations Center*, pages 37–63. Springer, 2020.

[612] Walt Tirenin and Don Faatz. A concept for strategic cyber defense. In *MILCOM 1999. IEEE Military Communications. Conference Proceedings (Cat. No. 99CH36341)*, volume 1, pages 458–463. IEEE, 1999.

[613] Mevlut Serkan TOK and Baris CELİKTAS. Muddywater apt group and a methodology proposal for macro malware analysis. *Bilişim Teknolojileri Dergisi*, 12(3):253–263, 2019.

[614] Inger Anne Tøndel, Maria B Line, and Martin Gilje Jaatun. Information security incident management: Current practice as reported in the literature. *Computers & Security*, 45:42–57, 2014.

[615] Peter Torr. Demystifying the threat modeling process. *IEEE Security & Privacy*, 3(5):66–70, 2005.

[616] Wiem Tounsi. *Cyber-Vigilance and Digital Trust: Cyber Security in the Era of Cloud Computing and IoT.* John Wiley & Sons, 2019.

[617] Wiem Tounsi and Helmi Rais. A survey on technical threat intelligence in the age of sophisticated cyber attacks. *Computers & security*, 72:212–233, 2018.

[618] Troy Townsend, Melissa Ludwick, Jay McAllister, Andrew O Mellinger, and Kate A Sereno. Sei innovation center report: Cyber intelligence tradecraft

project: Summary of key findings. Technical report, Carnegie–Mellon University. Software Engineering Institute, 2013.

[619] Ernesto Troiano, John Soldatos, Ariana Polyviou, Andreas Polyviou, Alessandro Mamelli, and Dimitris Drakoulis. Big data platform for integrated cyber and physical security of critical infrastructures for the financial sector: Critical infrastructures as cyber-physical systems. In *Proceedings of the 11th International Conference on Management of Digital EcoSystems*, pages 262–269, 2019.

[620] Daniele Ucci, Leonardo Aniello, and Roberto Baldoni. Survey of machine learning techniques for malware analysis. *Computers & Security*, 81:123–147, 2019.

[621] UK Ministry of Defence. Joint doctrine note 1/18. cyber and electromagnetic activities, February 2018.

[622] Rusydi Umar, Imam Riadi, and Ridho Surya Kusuma. Analysis of conti ransomware attack on computer network with live forensic method. *IJID (International Journal on Informatics for Development)*, 10(1):53–61, 2021.

[623] Daniel Uroz and Ricardo J Rodríguez. Characteristics and detectability of windows auto-start extensibility points in memory forensics. *Digital Investigation*, 28:S95–S104, 2019.

[624] US Army. Fm 34-45 tactics, techniques, and procedures electronic attack, 2000.

[625] US Army. *Cyberspace and Electronic Warfare Operations*. Army Publishing Directorate, April 2017.

[626] US Army Capabilities Integration Center. The us army concept for cyberspace and electronic warfare operations. 2025-2040. Technical report, 2018.

[627] Martin Ussath, David Jaeger, Feng Cheng, and Christoph Meinel. Advanced persistent threats: Behind the scenes. In *2016 Annual Conference on Information Science and Systems (CISS)*, pages 181–186, Princeton, NJ, USA, 16-18 March 2016. IEEE.

[628] Martin Ussath, David Jaeger, Feng Cheng, and Christoph Meinel. Pushing the limits of cyber threat intelligence: extending stix to support complex patterns. In *Information Technology: New Generations*, pages 213–225. Springer, 2016.

[629] Karl Utterback. *An Analysis of the Cyber Threat Actors Targeting the United States and Its Allies.* PhD thesis, Utica College, 2021.

[630] John R Vacca. *Online terrorist propaganda, recruitment, and radicalization.* CRC Press, 2019.

[631] Iman Vakilinia, Sui Cheung, and Shamik Sengupta. Sharing susceptible passwords as cyber threat intelligence feed. In *MILCOM 2018-2018 IEEE Military Communications Conference (MILCOM)*, pages 1–6. IEEE, 2018.

[632] Veronica Valeros, Maria Rigaki, and Sebastian Garcia. Machete: Dissecting the operations of a cyber espionage group in latin america. In *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, pages 464–473. IEEE, 2019.

[633] Michelle K. Van Cleave. What is counterintelligence. a guide to thinking and teaching about ci. *Journal of US Intelligence Studies*, 20(2), 2013.

[634] Victor Van der Veen, Lorenzo Cavallaro, Herbert Bos, et al. Memory errors: The past, the present, and the future. In *International Workshop on Recent Advances in Intrusion Detection*, pages 86–106. Springer, 2012.

[635] Renier Pelser Van Heerden. *A formalised ontology for network attack classification.* PhD thesis, Rhodes University, Grahamstown, South Africa, April 2014.

[636] Renier Pelser van Heerden, Barry Irwin, and Ivan Burke. Classifying network attack scenarios using an ontology. In *Proceedings of the 7th International Conference on Information-Warfare & Security (ICIW 2012)*, pages 311–324, 2012.

[637] RP Van Heerden, Barry Irwin, Ivan D Burke, and Louise Leenen. A computer network attack taxonomy and ontology. *International Journal of Cyber Warfare and Terrorism (IJCWT)*, 2(3):12–25, 2012.

[638] Rob Van Os. SOC-CMM: Designing and evaluating a tool for measurement of capability maturity in security operations centers. Technical report, Lulea University of Technology, Lulea, Sweden, September 2016.

[639] Ileen E Van Vuuren. It security trust model-securing the human perimeter. *International Journal of Social Science and Humanity*, 6(11):852, 2016.

[640] George Vardangalos. Cyber-intelligence and cyber counterintelligence (cci): General definitions and principles. Technical Report 1, Center for International Strategic Analyses (KEDISA), 2016.

[641] Said Varlioglu, Nelly Elsayed, Zag ElSayed, and Murat Ozer. The dangerous combo: Fileless malware and cryptojacking. In *IEEE Region 3 Technical, Professional and Student Conference*, March 31 - April 02 2022.

[642] Gaurav Varshney, Manoj Misra, and Pradeep K Atrey. A survey and classification of web phishing detection schemes. *Security and Communication Networks*, 9(18):6266–6284, 2016.

[643] Sandor Vegh. Classifying forms of online activism: The case of cyberprotests against the world bank. In *Cyberactivism*, pages 81–106. Routledge, 2013.

[644] Jorge Maestre Vidal, Marco Antonio Sotelo Monge, and Sergio Mauricio Martínez Monterrubio. Espada: Enhanced payload analyzer for malware detection robust against adversarial threats. *Future Generation Computer Systems*, 104:159–173, 2020.

[645] Jorge Maestre Vidal, Marco Antonio Sotelo Monge, Sergio Mauricio Martínez Monterrubio, Lorena Isabel Barona López, and Ángel Leonardo Valdivieso Caraguay. Profits at the dawn of cybercrime-as-a-service. In *2019 International Conference on Information Systems and Software Technologies (ICI2ST)*, pages 71–78. IEEE, 2019.

[646] Manfred Vielberth, Fabian Böhm, Ines Fichtinger, and Günther Pernul. Security operations center: A systematic study and open challenges. *IEEE Access*, 8:227756–227779, 2020.

[647] Sebastian Vogl, Jonas Pfoh, Thomas Kittel, and Claudia Eckert. Persistent data-only malware: Function hooks without code. In *NDSS*. Citeseer, 2014.

[648] Thomas D Wagner. Cyber threat intelligence for “things”. In *2019 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, pages 1–2. IEEE, 2019.

[649] Thomas D Wagner, Khaled Mahbub, Esther Palomar, and Ali E Abdallah. Cyber threat intelligence sharing: Survey and research directions. *Computers & Security*, 87:101589, 2019.

[650] Satohiro Wakabayashi, Seita Maruyama, Tatsuya Mori, Shigeki Goto, Masahiro Kinugawa, and Yu-ichi Hayashi. Poster: Is active electromagnetic side-channel attack practical? In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 2587–2589, 2017.

[651] Jing-he Wang, Jafar Tavoosi, Ardashir Mohammadzadeh, Saleh Mobayen, Jihad H Asad, Wudhichai Assawinchaichote, Mai The Vu, and Paweł Skruch. Non-singleton type-3 fuzzy approach for flowmeter fault detection: Experimental study in a gas industry. *Sensors*, 21(21):7419, 2021.

[652] Liying Wang and Yichao Zhang. Linear approximation fuzzy model for fault detection in cyber-physical system for supply chain management. *Enterprise Information Systems*, 15(7):966–983, 2021.

[653] Yi-Min Wang, Roussi Roussev, Chad Verbowski, Aaron Johnson, Ming-Wei Wu, Yennun Huang, and Sy-Yen Kuo. Gatekeeper: Monitoring auto-start extensibility points (aseps) for spyware management. In *LISA*, volume 4, pages 33–46, 2004.

[654] Gaute Wangen. The role of malware in reported cyber espionage: a review of the impact and mechanism. *Information*, 6(2):183–211, 2015.

[655] Patrick Wardle. Invading the core: Iworm’s infection vector and persistence mechanism. *Virus Bulletin*, October 2014.

[656] Patrick Wardle. Methods of malware persistence on mac os x. In *Proceedings of the virus bulletin conference*, 2014.

[657] Michael Warner. Notes on military doctrine for cyberspace operations in the united states, 1992-2014. *Cyber Defense Review*, 27, August 2015.

[658] Jesse Wasson and Christopher Bluesteen. Cognitive defense: Influencing the target choices of less sophisticated threat actors. *Homeland Security Affairs*, 13, 2017.

[659] Matthew S Webb. *Evaluating tool based automated malware analysis through persistence mechanism detection*. PhD thesis, Kansas State University, 2018.

[660] Fengguo Wei, Yuping Li, Sankardas Roy, Xinming Ou, and Wu Zhou. Deep ground truth analysis of current android malware. In *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, pages 252–276. Springer, 2017.

[661] Yichen Wei, Kam-Pui Chow, and Siu-Ming Yiu. Insider threat prediction based on unsupervised anomaly detection scheme for proactive forensic investigation. *Forensic Science International: Digital Investigation*, 38:301126, 2021.

[662] David A Wheeler and Gregory N Larsen. Techniques for cyber attack attribution. Technical report, Institute for Defense Analyses, 2003.

[663] Charles Wheelus, Elias Bou-Harb, and Xingquan Zhu. Towards a big data architecture for facilitating cyber threat intelligence. In *2016 8th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, pages 1–5. IEEE, 2016.

[664] Geoff White. *The Lazarus Heist*. Penguin Business, June 2022.

[665] Johannes Wiggen. The impact of covid-19 on cyber crime and state-sponsored cyber activities. *V. Facts & Findings*, (391):1–11, 2020.

[666] Heather J Williams and Ilana Blum. Defining second generation open source intelligence (osint) for the defense enterprise. Technical report, RAND Corporation Santa Monica United States, 2018.

[667] Brian Willis. Sharing cyber-threat information: An outcomes-based approach. Technical report, Intel Corporation, 2012.

[668] Clay Wilson. Information operations, electronic warfare, and cyberwar: Capabilities and related policy issues. Technical report, Library of Congress Washington DC Congressional Research Service, 2007.

[669] Clay Wilson. Cyber threats to critical information infrastructure. In *Cyberterrorism*, pages 123–136. Springer, 2014.

[670] Mingtao Wu, Zhengyi Song, and Young B Moon. Detecting cyber-physical attacks in cybermanufacturing systems with machine learning methods. *Journal of intelligent manufacturing*, 30(3):1111–1123, 2019.

[671] Yixin Wu, Cheng Huang, Xing Zhang, and Hongyi Zhou. Grouptracer: Automatic attacker ttp profile extraction and group cluster in internet of things. *Security and Communication Networks*, 2020, 2020.

[672] Wenjun Xiong, Emeline Legrand, Oscar Åberg, and Robert Lagerström. Cyber security threat modeling based on the mitre enterprise att&ck matrix. *Software and Systems Modeling*, 21(1):157–177, 2022.

[673] Dianxiang Xu and Kendall Nygard. A threat-driven approach to modeling and verifying secure software. In *Proceedings of the 20th IEEE/ACM international Conference on Automated software engineering*, pages 342–346, 2005.

[674] Wenyuan Xu, Ke Ma, Wade Trappe, and Yanyong Zhang. Jamming sensor networks: attack and defense strategies. *IEEE network*, 20(3):41–47, 2006.

[675] Tarun Yadav and Arvind Mallari Rao. Technical aspects of cyber kill chain. In *International Symposium on Security in Computing and Communication*, pages 438–452. Springer, 2015.

[676] Abel Yeboah-Ofori, J Abdulai, and Ferdinand Katsriku. Cybercrime and risks for cyber physical systems. *International Journal of Cyber-Security and Digital Forensics (IJCSDF)*, 8(1):43–57, 2019.

[677] Abel Yeboah-Ofori and Shareeful Islam. Cyber security threat modeling for supply chain organizational environments. *Future Internet*, 11(3):63, 2019.

[678] Abel Yeboah-Ofori, Umar Mukhtar Ismail, Tymoteusz Swidurski, and Francisca Opoku-Boateng. Cyberattack ontology: a knowledge representation for cyber supply chain security. In *2021 International Conference on Computing, Computational Modelling and Applications (ICCMA)*, pages 65–70, Brest, France, 14-16 June 2021. IEEE.

[679] Abel Yeboah-Ofori, Haralambos Mouratidis, Umar Ismai, Shareeful Islam, and Spyridon Papastergiou. Cyber supply chain threat analysis and prediction using machine learning and ontology. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*, pages 518–530, Crete, Greece, 25-27 June 2021. Springer.

[680] Jaepil Youn, Haengrok Oh, Jiwon Kang, and Dongkyoo Shin. Research on cyber ipb visualization method based on bgp archive data for cyber situation awareness. *KSII Transactions on Internet and Information Systems (TIIS)*, 15(2):749–766, 2021.

[681] Eric Yu. Modeling strategic relationships for process reengineering. *Social Modeling for Requirements Engineering*, 11(2011):66–87, 2011.

[682] Shuhong Yuan and Chijia Zou. The security operations center based on correlation analysis. In *2011 IEEE 3rd International Conference on Communication Software and Networks*, pages 334–337. IEEE, 2011.

[683] Xiaoyong Yuan. Phd forum: Deep learning-based real-time malware detection with multi-stage analysis. In *2017 IEEE International Conference on Smart Computing (SMARTCOMP)*, pages 1–2. IEEE, 2017.

[684] Wen Zeng and Vasileios Germanos. Modelling hybrid cyber kill chain. In *Proceedings of the International Workshop on Petri Nets and Software Engineering*, pages 143–160, Aachen, Germany, 23-28 June 2019.

[685] Xiaobing Zhang, Shyhtsun Felix Wu, Zhi Fu, and Tsung-Li Wu. Malicious packet dropping: how it might impact the tcp performance and how we can detect it. In *Proceedings 2000 International Conference on Network Protocols*, pages 263–272. IEEE, 2000.

[686] Xiaojun Zhou, Zhen Xu, Liming Wang, Kai Chen, Cong Chen, and Wei Zhang. Kill chain for industrial control system. In *MATEC Web of Conferences*, volume 173, page 01013. EDP Sciences, 2018.

[687] Ziyun Zhu and Tudor Dumitras. Chainsmith: Automatically learning the semantics of malicious campaigns by mining threat intelligence reports. In *2018 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 458–472, London, UK, 24–26 April 2018. IEEE.

[688] Aaron Zimba and Mumbi Chishimba. Exploitation of DNS tunneling for optimization of data exfiltration in malware-free APT intrusions. *Zambia ICT Journal*, 1(1):51–56, 2017.

[689] Aaron Zimba and Zhaoshun Wang. Malware-free intrusions: Exploitation of built-in pre-authentication services for APT attack vectors. *International Journal of Computer Network and Information Security*, 9(7):1, 2017.

[690] Carson Zimmerman. Cybersecurity operations center. *The MITRE Corporation*, 2014.

[691] Anita Rajendra Zope and DR Ingle. Event correlation in network security to reduce false positive. *International Journal of Computer Science & Communication Networks*, 3(3):182, 2013.

[692] Fadi Abu Zuhri. *The illusion of the cyber intelligence era.* zahf. me, 2019.