On Attacker Models and Profiles for Cyber-Physical Systems

Marco Rocchetto[1] and Nils Ole Tippenhauer[2]

[1] iTrust, Singapore University of Technology and Design
[2] ISTD, Singapore University of Technology and Design

Abstract. Attacker models are a fundamental part of research on security of any system. For different application scenarios, suitable attacker models have to be chosen to allow comprehensive coverage of possible attacks. We consider Cyber-Physical Systems (CPS), that typically consist of networked embedded systems which are used to sense, actuate, and control physical processes. The physical layer aspects of such systems adds novel attack vectors and opportunities for defenses, that require extended models of attackers' capabilities. We develop a taxonomy and show that, so far, there are no commonly used attacker models for such CPS. In addition, concepts of what information belongs into an attacker model are widely different among the community. To address that problem, we develop a framework to classify attacker models, and use it to review related work on CPS Security. Using that framework, we propose a set of attacker profiles and show that those profiles capture most types of attackers described in the related work. We hope that our contributions will enable the use of more well-defined and uniform attacker models in the future.

1 Introduction

In recent years, security of Cyber-Physical Systems (CPS) has received increasing attention by researchers from the domain of computer science, electrical engineering, and control theory [25, 30, 16, 29]. We use the term CPS to refer to systems that consist of networked embedded systems, which are used to sense, actuate, and control physical processes. Examples of such CPS include industrial water treatment facilities, electrical power plants, public transportation infrastructure, or even smart cars. All those systems have seen a rapid increase in automation and connectivity, which threatens to increase vulnerability to malicious attacks.

In contrast to the domain of information security, where the Dolev-Yao attacker model [12] (DY) is widely used for protocol analysis, the state-of-the-art for CPS security does not have a common terminology for attacker models. Instead, attacker-models are usually defined ad-hoc for the specific setting considered. Even if the topic has been broadly discussed in the CPS research community (e.g., in [30, 16, 29]) only a small number of tentative works [35, 19] have tried to overcome this problem.

In this work, we provide a comprehensive overview of work on CPS attacks. We find that in most cases, the authors show how to attack the system without defining attacker models [2, 3, 17] or propose their own attacker model(s) depending on the system they are considering (e.g., [8]) or leave to the users the (non trivial) problem of defining their

own attacker models [2]. In general, authors often prefer to use their own attacker model with a *customized* (reduced) set of constraints on the attacker. In addition, we review attempts to provide more general attacker models or frameworks, e.g., [35, 19].

Based on our findings of related work, we show commonalities and differences in existing attempts to generalize attacker models for CPS, and provide recommendations for future attacker models in that direction.

We summarize our main contributions as follows:

- We define and apply a taxonomy of 10 different features to classify and compare attacker models in related work
- We provide a detailed overview of work discussing attacks and attackers on CPS
- We propose an attacker framework and a more formal and standardized definition of attacker model for CPS
- Using that framework, we extract attacker profiles from related work, analyze those profiles, and propose six attacker profile archetypes that distill common intuition behind related work

In addition, we developed a complementary tool to support our review of the related work. The tool allows the application of our taxonomy to classify related work, comparisons between profiles and export filters, and contains the results of our analysis. The tool is called APE available at scy-phy.net/ape.

Structure. In Section 2, we describe the scope of our review and aspects considered. We review the literature in Section 3 (attacks on CPS, categorizations of attackers, (semi-) formal model of attacker). In Section 4 we analyze the related works, showing a list of commonalities and metrics that we use to categorize the related work accordingly. We propose an attacker model framework in Section 5, apply it to the related work to obtain their attacker profiles, and analyze those profiles and propose a set of own profiles. We conclude the paper in Section 6.

2 Scope and Taxonomy for Related Work Review

We start by defining the scope of our related work review, we then provide definitions which help us to classify the related work. Finally we present our taxonomy that we use to summarize related work.

2.1 Scope of our review

We review the related work on: i) attacks on CPS and their ad-hoc attacker models, ii) works which profile attackers for CPS and iii) works on generic attacker models for CPS. We start by reviewing works that discuss specific attackers who target or leverage the physical layer in their attacks (mechanical, electrical interactions). These works are found in the domain of public infrastructure (e.g, water [2] and power [21]). To limit the scope, we do not focus on attacks that are only related to physical-layer wireless communication (e.g., key establishment, jamming, anti-jamming, friendly jamming).

In addition, we review works that provide profiles for different attackers on CPS, and works that consider more generic attacker models that include the physical layer.

2.2 Terminology

Interestingly, we did not find general definitions of central terms related to models for CPS attackers. In [7], there is a first attempt in providing general definitions for CPS security but the authors focus on attacks and properties rather than attacker model. For that reason, we now provide a short description of the central terminology we use in the remainder of this work.

A *System under attack* is an interacting or connected group of components (soft- and hardware, humans) forming a unified whole and serving a common purpose.

An *Attacker* is a group of human actors that collaborate to achieve a goal related the system under attack. An *Attacker Profile* describes informal templates or classes of attackers. These profiles are more a description of the setting and intuition, and not an exhaustive listing of possible actions, motivations, or capabilities of the attacker.

An *Attacker Model* (together with compatible system models) will ideally fully define the possible interactions between the attacker and the attacked system. In particular, the model will define constraints for the attacker (e.g. finite computational resources, no access to shared keys)

A *System Model* is characterizing all relevant components of the system under attack, to a level of detail that allows to determine all possible interaction of the attacker with the system. We remark that we will not go into the details of the system model since our work focuses on the attacker. Therefore, we will not distinguish between system models which consider (or not) risk or threat linked to components of the system.

An *Attack Model* is characterizing all potential interactions between the attacker and a specific system under attack and the specification of the goal that the attacker wants to achieve with respect to the system under attack. One can consider an attack model as an instantiation of the attacker model on a specific scenario.

2.3 Taxonomy

In our review of related work, we will systematically analyze and summarize the attacker models (or related models) that are used to describe the attacker. In particular, we focus on the following aspects of the related work:

1. If different attacker profiles are discussed, and how many
2. The dimensions used by authors to define the attacker
3. The number of actions types available to the attacker
4. Use of system model (or constraints on type of system)
5. Validation of attack(er) models
6. Generality of model (i.e., specific for one CPS or general attacker model)
7. Supporting case studies (and if they are ad-hoc, real)
8. Whether the authors considered time in their models
9. Terminology used by authors for the model, and how it fits to our terminology
10. The main research goal of the reviewed work.

3 Review of Attacker Definitions in Related Work

The idea of attacker models for CPS has been explored from different perspectives. In this section, we provide a review of related works that focus on a specific case study or attacks which can be exploited on a CPS or a class of CPS (Section 3.1). We then review common informal and semi-formal *attacker profiles* that are often used in research and by the public (Section 3.2). Afterwards, we review CPS-related attacker models (Section 3.3). We *emphasize* the key points we have used in our taxonomy. For a more detailed description of our review we refer to our tool available at scy-phy.net/ape.

3.1 Attacks on CPS

In the following, we provide a review of works which focuses on specific attacks or (class of) CPS.

Amin et al. In [2], the authors perform security *threat assessment* of networked control systems and Supervisory Control and Data Acquisition (SCADA) systems with regulatory and supervisory control layers. Authors *do not define a model of the system* and their technique is *specific* for one case study. Even if no dimensions are explicitly considered, some assumptions are made on the *knowledge* of the attacker and his *resources*. There is no discussion on specific actions for the attacker but authors consider *attacks* as a general action.

Esfahani et al. In [13], the authors propose an approach for the identification of security flaw in of electric power transmission systems. The authors do not discuss *profiles*, *dimensions* or *actions* of the attacker model because the study is focused on the *modeling of the system* and the aim is to perform *risk analysis*.

Krotofil et al. In [18], the authors discuss the importance of *time* in security attacks to CPS. The discussion is *specific* for electric power grid. An *attacker model* (called adversary model in the paper) is defined with DoS and false data injection attacks as only *actions*. The only *dimension* of the attacker model is identified with his goals. Authors apply their results to *one case study*.

Lin et al. In [20], the authors study vulnerabilities of distributed energy routing processes by *attack simulation*. Authors focus on false data injection attacks and analyze their impact on the *system model*. The *attacker model* (called threat model) can modify data and compromise component injecting malicious codes. Authors consider the attacker's *knowledge of the system* and ability to *attack* the system (node compromise).

Liu et at. In [21], the authors define a new attack class against electric power systems. The basic idea is that an attacker can inject malicious measurements (attack) *without being detected* by any of the existing techniques for bad measurement detection. Authors describe how to formally represent a *system model* of a power grid and test their attacks against *two ad-hoc examples*.

Taormina et al. In [31], the authors define how to *simulate cyber-physical attacks* on water distribution systems. EPANET[33] (a numerical modeling environment) is used to define the *system model* along with the properties of each component. The *attacker model* is informally defined by two *actions*: direct and indirect attacks. These actions

represent the *knowledge of physical and virtual attacks*. The effectiveness of the technique is motivated on *one case study*.

Urbina et al. In [34], the authors discuss practical MitM (Man in the Middle) *attacks* on ICS Fieldbus communications. They concretely perform such an attack on a water treatment testbed called SWaT (Secure Water Treatment). The *attacker model* consists in a description of his main characteristics which are divided into two different *dimensions*: objective and resources.

3.2 Attacker Profiles

A number of authors defined, formally or semi-formally, attacker profiles. In the following, we provide a summary of that related work.

Cardenas et al. In [5], the authors informally discuss some challenges for securing CPS. They start by identifying the lack of terminology and attacker models for CPS. Authors informally define *four attacker profiles* (*adversary models* in the paper) with respect to *two dimensions*. The authors highlight the importance of defining which are the specific *attacks* targeting CPS. No formal attack model or case studies are provided.

Cardenas et al. In [7], the authors address the problem of sensor network security focusing on SCADA systems. They propose a taxonomy for security of sensor networks discussing security *properties*, the *attacker model* (*threat model* in the paper). They distinguish between *insider and outsider attacks* and several dimensions and subdimensions to rank the attacker (attacker profiles). *Skills*, *costs* and *distance* are discussed in the paper. Authors do not define specific *actions* of the attacker since they focus on the system model (called threat model) but they discuss a number of *attacks*.

The physical distance between the attacker and the target (for wireless networks) is discussed in [28, 10, 8]. In particular, in [8], the authors define an insider attacker and locality dimension to describe attackers for securing wireless authentication.

Corman et al. In a talk at the RSA conference [9] in 2012, the authors presented an high-level definition of several *attacker profiles* which they call adversaries. The authors defined several *dimensions*. Authors do not define a set of possible attacks but provide some examples. Finally, we highlight that the work is not published and no or a few details about profiles and the model in general are provided.

Heckman. In [16] a comprehensive informal proposal of several attacker profiles is presented in an industrial white paper. The author shows several *dimensions* to categorize different *attacker profiles*. Even if the categorization describes several different attacker profiles, there is no formal definition of attacker model and there is no clear distinction between terrorists and hacktivists, and between basic users and cybercriminals. Furthermore, the categorization does not consider any physical aspect of the attacker focusing on cyber actions only. The author uses the dimensions to *rate the threat risk* of each attacker profiles over one ad-hoc case study.

A similar categorization for cyber attacker is defined in [29]. The authors performed an extensive description of concrete metrics to categorize an attacker. The work focuses on a subset of profiles without going too much into the details of the dimensions distinguishing these profiles.

3.3 Formal Models for Attackers

We now provide an overview of related work on formal models for CPS attackers.

Adepu et al. In [1], the authors defines how to model a CPS along with an attacker. The study focuses on a *specific* CPS (a water treatment system) and the *dimensions* that are used to define the attacker can be summarized as: components of the CPS (the target), the property an attacker wants to violate and performance (impact of the attack). *Actions* are defined as steps of the attack model.

Basin et al. In [4], the authors present a formal model for modeling and reasoning on security protocols that are using physical-layer properties such as the *distance* between communication partners. The authors define several dimensions (*time*, agent locations, and physical properties of the communication network) to describe physical properties of CPS (such as the physical distance between communication partners). They then define the intruder as set of nodes of the formalized CPS. Authors apply their model to *four case studies*.

Le May et al. In [19, 14], the authors formally define a framework for the identification of attacks in CPS. The authors define a set of abstract components to describe an attack execution graphs (AEG) and an *attacker model*. The AEG represents potential attack steps against the system, together with a formal definition of the attacker using a set of six *dimensions*. This formalization has been implemented in a framework called ADVISE where users can define their own attacker models, e.g., defining the *knowledge* of the attacker with respect to the AEG. Some *attacker profiles* are defined with respect to cost, payoff and detection.

McEvoy et al. In [23], the authors present a variant of π-calculus to prove security properties in the context of intrusion detection for SCADA systems. Authors define how to *model a SCADA network* along with an *attacker model* (called agent-based adversary capability model in the paper). In contrast with the DY model, the intruder is not the source and sink of all communications but he can communicate by request. The dimensions considered are: *distance*, topology of the network related to attacker actions and *skills*, the attacker can subvert any process.

Mo et al. A survey on CPS security, but specific for power grid, is presented in [25]. The authors formally describe how to *model a power grid* and provide a description of possible *attackers' actions* and goals. There is no mention to attacker profiles but several general actions to describe the attacker are provided. Finally, we notice that the terms adversary, attack and attacker model are used as synonyms in the paper.

Orojloo et al. In [26], authors define an approach for *modeling* and evaluating the security of CPS. They propose a model, based on semi-Markov chain, which aims at predicting possible attacker's decisions with respect to the search of both cyber and physical attacks. The authors define five different *dimensions*. Finally, they show how their technique can concretely be used against a simple ad hoc case study.

Teixeira et al. In [32], authors define an approach for the modeling of attacks and scenarios in network controlled system. They describe how to define an *attacker model* using three main *dimensions* (along with several sub-dimensions): knowledge, disclosure resources and disruption resources available to the attacker. The authors take into

Table 1: Summary of taxonomy of related work on attacker models and profiles for CPS

Publication	#Profiles	#Dimensions	#Actions	System Modeling	Validation	Generic/Specific	#Test Cases	Time	Terminology used	Our terminology	Research Goal
Amin et al. [2]	1	2	1	○	○	S	1	●	AtkM	AtkM	Threat Assessment
Esfahani et al. [13]	0	0	0	●	●	S	1	●	SM	SM	Risk Analysis
Krotofil et al. [18]	0	1	1	○	○	S	1	●	AdM	AtM	Security Analysis
Lin et al. [20]	0	1	1	●	●	S	1	●	TM	AtM	Attack Simulation
Liu et al. [21]	0	3	1	●	●	S	2	●	SM	SM	Attack Simulation
Taormina et al. [31]	0	2	1	●	●	S	1	●	AtM	AtM	Attack Simulation
Urbina et al. [34]	1	4	1	○	○	S	1	●	AtM	AtkM	Testing
Adepu et al. [1]	0	1	1	●	○	S	1	○	AtM	AtM	Security Analysis
Cardenas et al. [5]	4	2	1	○	○	G	0	●	AdM	AtP	Overview
Cardenas et al. [7]	2	4	1	○	○	G	0	●	TM	AtM	Risk Analysis
Corman et al. [9]	4	4	0	○	○	G	0	○	Ad	AtP	Risk Analysis
Heckman [16]	9	5	0	○	○	G	1	○	TM	AtP	Risk Analysis
Basin et al. [4]	0	2	2	●	○	G	4	●	IM	AtM	Security Analysis
Le May et al. [19]	4	8	0	●	●	G	2	●	AdP	AtM	Risk Analysis
McEvoy et al. [23]	0	2	3	●	●	G	1	○	Ad	AtM	Intrusion Detection
Mo et al. [25]	0	0	8	●	○	G	0	●	AtM	AtM	Survey
Orojloo et al. [26]	0	5	0	●	○	G	1	●	SM	SM	Quantitative Evaluation
Teixeira et al. [32]	0	4	0	●	○	G	1	●	AdM	AtM	Security Analysis
Vigo [35]	0	2	5	●	○	G	0	○	AtM	AtM	Definition

●= argument discussed, ○= not discussed, At=Attacker, I=Intruder, Ad=Adversary, T=Threat, S=System, Atk=Attack, M=Model, P=Profile

account the *stealthiness* of an attacker. The attacker model is *general* but constrained to networked controlled systems and is tested on *one test case*.

Vigo. In [35], the author presents a formal definition of an *attacker model* for CPS. The attacker model is presented along with a *system model*. The attacker is define as a set of pairs representing locations in the network topology and capabilities. *Capabilities* are defined as a set of tuples expressing actions, cost (energy/time) and range (with respect to the topology) of the attacker. The attacker is believed to perform two types of attacks: *physical*, against a device and *cyber* against the communications

4 Discussion of Attackers in Related Work

We now summarize our findings, and show the results of applying our taxonomy to the related work in Table 1. Then, we discuss each aspect of the taxonomy in detail.

Profiles, dimensions, and actions. Seven works explicitly use different attacker profiles, seventeen define dimensions and the vast majority use actions to characterize the

attacker. Just two works define a system model and perform risk analysis without explicitly considering an attacker model. This shows the trend of defining an attacker model to perform security analysis on CPS and, at the same time, that there exist various way to model the attacker.

One common aspect of the related work is that the attacker actions should consider all the actions of the usual cyber attacks, e.g., read the network communication (sniffing) and modifying all or some of the messages (spoofing) with the ability of injecting new values. The authors commonly assume that the attacker should not be identified with the network itself but, instead, be located somewhere in the network. In other words, following the rules defined by the topology of the network. This gives to the attacker the possibility to divert a node and then to decrypt (encrypt) the network traffic if the node contains the proper key.

System modeling, validation, and test cases. Roughly half of the reviewed papers define how to create a model of a CPS, but only a few (six) validate their model against an attacker model. Considering validation, simulation, and implementation, we note that in general, only three papers show their results on more than one test case.

Time. Most of the works take into account the notion of time as an important feature to perform attacks since a CPS very often has different (sequential and/or parallel) phases. An attack then has to be carefully timed to go through some of these phases in a particular order or to not be detected by intrusion detection systems.

Terminology. We note that there is no common terminology and attacker, attack and threat model are usually used as synonyms. In Table 1 we propose a mapping from the various terminologies used in the papers to the one we proposed in Section 2.2.

Summary. From our review, we notice that the actions for the attacker model for CPS have been defined in a common way, i.e., all the papers share the same actions or the same intuitions on this aspect. However, they apply those actions to different definitions of the concept of attacker models. We can group the reviewed papers into two different categories, i) the ones which use different attacker profiles with different properties (e.g., to distinguish between insider and a nation-state attackers) and ii) the ones which define a set of dimensions, e.g., knowledge, to define one specific attacker model. Both groups aim at identifying a set of useful characteristics of the attacker but the former, as showed in Table 1 is more focused on risk analysis and tries to handle several different attacker instantiations while the latter is more system-specific and focuses on one generic description of the attacker model.

One might ask which is the best way to define an attacker model or if there exist a way to define *one* general attacker model in the context of CPS (e.g., as the DY model for security protocols); or if CPS are so heterogeneous that we should define a variety of different profiles for the attacker. In the remainder of this section, we provide insights to answer to these questions by discussing different attacker profiles and dimensions found in the related work. We believe that a common understanding of what are thought to be the key aspects of the attacker model (in the context of CPS) can be useful for the identification of a common definition. In Section 5 we propose a first steps in this direction by providing an attacker framework and a more formal definition of attacker model and profile.

4.1 Profiles

The following classification is a collection of all the attacker profiles we have found in the literature. The boundary between the different attacker profiles are not well defined, and sometimes it is hard to classify a specific real-life attacker as one specific profile.

Basic user [9, 16], also known as *script kiddie*, *unstructured hacker*, *hobbyist* or even *crackers*. Someone who uses already established and potentially automated techniques to attack a system. This attacker has average access to hardware, software, and Internet connectivity, similar to what an individual can obtain through purchase with personal funds or by theft from an employer.

Insider [16, 5, 7, 19], which for example can be *disgruntled employees* or a *social engineering victims*. The employment position or the system privileges he owns (e.g., user, supervisor, administrator) are tightly related to the damage he can cause to the target. This type of attacker is of high importance for systems that are mainly protected through air-gaps between the system network and the outside world (often used in CPS).

Hacktivist [5, 9, 16]. A portmanteau word which combines hacker and activist, as defined in [11]. This class of attackers uses their hacking abilities to promote a political agenda. Often related to freedom of information (e.g., Anonymous).

Terrorist [5, 16, 19], also known as *cyber-terrorist*. Is a politically motivated attacker who uses computers and information technology in general to cause severe disruption or widespread fear [22, 11].

Cybercriminal [5, 7, 9, 16, 19], sometimes generally called *black hat hacker* or *structured hacker*. An attacker with an extensive security knowledge and skills. This category of attackers takes advantage of known vulnerabilities, and potentially has the knowledge and intention of finding new zero-day vulnerabilities. The cyber-criminals' goals can range from blackmailing to espionage (industrial, foreign) or sabotage.

Nation-State [16, 9, 16, 19], an attacker sponsored by a nation/state. Possibly belonging to (or that used to belong to) a state organization for carrying out offensive cyber operations [27]. His targets usually are public infrastructure systems, mass transit, power or water systems, and general intelligence.

4.2 Dimensions

By assigning quantitative or qualitative scores on the dimensions, a large set of potential attacker configurations could be described. We now define a set of dimensions extracted from the related work. The application of those definitions to the related work is summarized in Table 2. Note that we have standardized the names used for dimensions. Therefor, the names of the dimensions in Table 2 might be different from the one used in the related work. For readability and lack of space we do not go into the details of the mapping which is defined in the APE.

- *Financial support*, expresses the budget that an attacker has to perform his attacks.
- *Manpower available*, is used to differentiate between lone attackers and (small to large) groups. This dimension expresses quantitatively the human resources available to perform the attack.

Table 2: Dimensions proposed in the related work

	Aim-Physical	Aim-Virtual	Resources	Offensive	Distance	System	Manpower	Tools	Credentials	Camouflage	Motivation	Target Asset	Aim	Honesty	Determination	Likelihood	Knowledge	Attack Step	Financial Support	Psychology	Reward	Easy Of Access	Physical	Network	Disclosure Resources	Disruption Resources	Aim-Virtual (Availability)	Protocols
Adepu et al.[1]	●	●																										
Amin et al.[2]			●	●	●																							
Basin et al.[4]					●	●																						
Cardenas et al.[5]					●		●																					
Cardenas et al.[7]					●			●	●	●																		
Corman et al.[9]								●			●	●	●															
Esfahani et al.[13]																												
Heckman[16]			●	●										●	●	●												
Krotofil et al.[18]	●	●																										
Le May et al.[19]	●	●	●	●					●	●							●	●	●	●								
Lin et al.[20]						●																						
Liu et al.[21]						●				●																		
McEvoy et al.[23]				●	●																							
Mo et al.[25]																												
Orojloo et al.[26]					●	●													●		●	●						
Taormina et al.[31]																							●	●				
Teixeira et al.[32]						●				●															●	●		
Urbina et al.[34]					●			●																			●	●
Vigo[35]					●	●																						

- *Tools (Resources) available*, also known as attacklets, or actions in abstract definition of attacker model, defines which types of tools are available to the attacker. This dimension can be used to better understand which are the countermeasures needed to protect a CPS.
- *Camouflage* or preference to stay hidden, expresses the aim and/or the ability of the attacker to not be tracked down after or while performing an attack.
- *Distance* to the CPS. An attacker can be located in another country, within WiFi range or possibly have direct access to the system.
- *Knowledge*, defines the knowledge of the attacker. It may refer to the knowledge of the *System*, the technical knowledge (distinguish between *Physical*, *Network* and *Protocols*) and attack knowledge (*Offensive*) which can be considered as subdimensions. In addition, some of the authors consider the *Credentials* dimension as related to the knowledge of the system.
  Note here that the knowledge of the attacker is intuitively always considered. However, sometimes the knowledge (of the system or attacks) is hard-coded into the system model and not explicitly considered as part of the attacker model.

– *Attack*, defines which type of attack an attacker can perform, e.g., white, gray or black box attack. This dimension can be used to determine whether obfuscation should be take into consideration as a protection against a particular attacker profile.
– *Target* (e.g., CPS, valves, pumps, access points, information) identifies which physical and logical parts of the system under attack are targeted by an attacker profile.
– *Motivations* and *Aim*, which can be considered as a sub-dimension of target, refer to the objective of the attacker. In some work, the authors details the aim distinguishing between *Physical* or *Virtual* components of the system.

5 Profiles and A Generic Attacker Framework

In this section, we propose the draft of a formalized *attacker framework* that is designed to encompass commonly used informal attacker models in other works. The framework allows to define *attacker profiles* characterized by a number of dimensions.

The idea behind our framework is that an *attacker model* can be described by a set of dimensions. These dimensions can be instantiated to define an *attacker profile* which characterize the key aspects of an attacker. We cannot prove that our framework is complete, however, we have considered, expanded and structured all the aspects extrapolated from our review, i.e., the ones in Table 2.

5.1 Attacker framework, profile, model, and system model

From our literature review in Section 3, we found that attacker models are often defined on different layer of abstractions. Before going into the details of our framework, we propose a terminology to differentiate between those different layers, and show how they are related (see Figure 1).

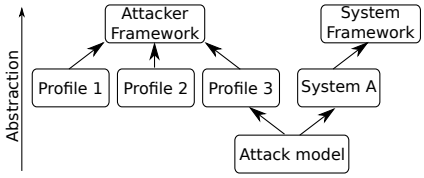

Fig. 1: Proposed hierarchy of attacker framework, profiles, attack models and system models.

An *attacker framework* is defined as a set of different, structured *dimensions* which quantitatively represent a characteristic of an attacker. A *metric* is associated to each dimensions and when the dimensions are instantiated, the framework produces an *attacker profile*. An attacker profiles is then an instantiation of the set of dimensions defined by the attacker framework. For the sake of readability, we provide the details of each dimension and sub-dimension in the Appendix A.

Correspondingly, we define a *system framework* as a paradigm which provides different aspects (dimensions) of a CPS. By instantiating these aspects we produce a *system model*. In practice, system models are often only considering a small subset of

the system under consideration due to the involved complexity. Such reduced system models are nevertheless useful to define the scope of the analysis. When we combine attacker profiles and system model (e.g., we run the attacker profiles against a system model) searching for attacks, we obtain an *attack model*.

There is a strong connection between the DY model and the attacker models we have found in our literature review. One intuitive question is how we position the DY model in our definitions. The DY model is defined as a set of actions (e.g., encryption, decryption, concatenation), usually formalized as set of deduction rules. However, a set of constraints over the attacker capabilities is usually defined along with the actions. To give some examples, in the verification of security protocols, the DY is usually identified with the network (i.e., he can read all the messages that are passing through the network) and perfect cryptography is often assumed. In our review, the DY model is always defined along with some constrains. As an example, the attacker's position on the network topology is considered in [4]. These constraints can be defined in one or more profiles of the DY attacker model. Due to lack of space we will not go into the details of the DY profile. In the remainder of this section, we standardize the attacker profiles proposed in the related work.

5.2 Mapping Profiles in Related Work to Our Profiles

In order to standardize the attacker profiles we have first mapped the profiles in the related work into our framework as showed in Table 3.

Using WEKA [15] (a machine learning tool) we have applied several machine learning algorithm for clustering the profiles (results reported at scy-phy.net/ape). However, the results show that there is no general agreement between different authors on the definition of the same or similar profiles (with an incorrectly clustered instances parameter above 47%). The only exceptions are the insider profiles which are correctly clustered together. We have then defined six archetypal profiles, based on the descriptions in the related work, and showed that they are generalization of the ones proposed in the related work.

We now define a *profile distance metric* to measure the distance between two attacker profiles, and analyzed how well the profiles of related work cluster, and fit to our generic profiles as defined in Section 5.3.

5.3 Attacker Profile Archetypes

In Table 4, we give a more rigorous definition of the six common attacker profiles (we described in Section 4.1) using our framework.

As it can be seen in Table 4, the honesty dimension is the same on all the archetype. This is because all but one work [16] only consider dishonest attacker profiles. Our terrorist profile is classified with low knowledge of offensive skills. Changing this metric to an higher metric leads to a mismatch between the terrorist profiles in the literature and the archetype.

1. *Basic User*. Represents the lower bound of our profiles with all the dimensions set to the lowest value. Usually, attacks from this type of profile are believed to be very frequent. However, in the case of CPS might not be the case.

Table 3: Categorization of attacker profiles found in the related work

Dimensions	Cardenas [5] Cybercriminal	Cardenas [5] Insider	Cardenas [5] NationState	Cardenas [5] Terrorist	Corman [9] AdaptivePersistent	Corman [9] Hacktivist	Corman [9] OrganizedCrime	Corman [9] Skiddie	Heckman [16] Hacktivist	Heckman [16] Hobbyist	Heckman [16] Insider	Heckman [16] NationState	Heckman [16] OrganizedCrime	Heckman [16] ScriptKiddie	Heckman [16] StructuredHacker	Heckman [16] Terrorist	Heckman [16] UnstructuredHacker	Le May [19] DisgruntledEmployee	Le May [19] LoneHacker	Le May [19] NationState	Le May [19] SystemAdministrator	Le May [19] Terrorist	Urbina [34] Insider
Knowledge		●							◐	◐	◐	●	●	○	●	●	◐	◐	◐	◐	●	○	●
Offensive									◐	◐	○	●	●	○	●	●	◐		◐	●		○	●
Physical																							
Network																							●
Software																							●
System		●																◐	○	○	●	○	●
Source code																							
Protocols																							●
Credentials		●																◐	○	○	●	○	◐
Resources		◐		●	●	◐	◐	○	●	○	◐	●	●	○	◐	●	◐	◐	○	◐	◐	○	◐
Distance		●																●	○	○	●	○	●
Manpower				●															○				
Effort																							
Tools					●	◐	◐	○															○
Financial support		○							●							●		○		●	○		
Psychology	◐	◐	◐	●	◐	●	●	◐	●	○	●	●	●	●	◐	●	◐	◐	●	●	◐	◐	◐
Honesty	●	●	●	●	●	●	●	●	●	○	●	●	●	●	●	●	●	●	●	●	●	●	●
Periodicity																							
Camouflage																		○	◐	●	○	○	
Aim-Physical				●																			◐
Integrity																							●
Confidentiality																							○
Availability				●																			●
Determination									●	○	◐	●	◐	○	◐	●	◐	●	●	●	●	●	
Strategy	◐														●		○	●	●	●	●	●	●
Aim-Virtual	●		○	●	◐	●	●	◐															◐
Integrity																							●
Confidentiality					●	●	●	◐															○
Availability	●			●	○	●	○	○															○

A metric on each dimensions is expressed on the (strict) partially ordered set [○<◐<●]

2. *Cybercriminal*. Advanced knowledge of network attacks but low of physical layer attacks. Advanced tools and average financial support.

Table 4: Categorization of proposed attacker profile archetypes

	Knowledge	Offensive	Physical	Network	Software	System	Source code	Protocols	Credentials	Resources	Distance	Manpower	Effort	Tools	Financial support	Psychology	Honesty	Periodicity	Camouflage	Strategy	Determination	Aim-Physical	Integrity	Confidentiality	Availability	Aim-Virtual	Integrity	Confidentiality	Availability
B	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○
C	◐	◐	○	●	◐	○	○	○	○	◐	○	○	◐	●	◐	◐	●	◐	◐	●	◐	○	○	○	○	●	◐	●	●
H	◐	◐	○	●	●	○	○	○	○	◐	○	◐	●	◐	○	◐	●	○	◐	◐	●	○	○	○	○	◐	○	●	●
I	◐	○	○	○	◐	●	●	●	●	◐	●	○	◐	◐	○	◐	●	○	◐	●	◐	◐	◐	◐	◐	◐	○	◐	◐
N	◐	●	●	●	●	○	○	○	○	●	○	●	●	●	●	◐	●	○	●	●	●	●	◐	●	●	○	○	◐	○
T	○	○	○	○	○	○	○	○	○	◐	○	◐	●	◐	◐	◐	●	○	○	○	●	◐	○	○	●	◐	○	◐	◐

B=BasicUser, **C**=Cybercriminal, **H**=Hacktivist, **I**=Insider, **N**=NationState, **T**=Terrorist. A metric on each dimensions is expressed on the (strict) partially ordered set [○<◐<●]

3. *Hacktivist*. Similar to the cybercriminal but with a lower financial support but higher manpower support.
4. *Insider*. It is the only profile which has an advance knowledge of the system because it has physical access to it. He has a structured strategy to perform his attacks. His aim are physical properties of the system (e.g., damage the system to attack its availability. He acts alone, with low budgets but with dedicated tools.
5. *Nation-State*. On average the most powerful profile between the archetypes. High offensive skills and tools, high resources and determination. The stealthiness of the attacks is very important.
6. *Terrorist*. Low offensive skills and average resources. The attacks mainly targets the physical availability of the system and their stealthiness is not important.

5.4 Validation of Proposed Profiles and Discussion

Motivated by the results obtained by the machine learning clustering phase, we investigated if our archetypes generalize the related work. We used the Euclidean distance on a n-dimensional space to calculate the distance between profiles as $\sqrt{\sum_{i=1}^{n}(q_i - p_i)^2}$, where two profiles p and q are represented as two points in an Euclidean n-dimensional space: $p = (p_1, \ldots, p_n)$ and $q = (q_1, \ldots, q_n)$. Each point is defined by the metric associated to a dimension, mapping the poset [○<◐<●] to $[1 < 2 < 3]$ $(1, 2, 3 \in \mathbb{N})$.

In 21 cases out of 23, our profile archetype correctly matches to the expected profile (see Table 5). That implies that a) attacker models in related work are based on commonly used implicit profiles, and b) our profiles are closely approximating the underlying intuition behind the commonly used profiles. That result now allows to relate attacker profiles from related work with each other, and could be used to complement those profiles with additional missing information based on our archetypes.

There are two cases in which expected mapping is not found. In [16] the authors do not distinguish between a terrorist and a Nation-State profiles. In fact, the nearest

Table 5: Distance of attacker profiles from related work to our proposed six profiles. Columns represent the first, second, ..., sixth best fit and the respective distance metric value.

Profile	#1	#2	#3	#4	#5	#6
Cardenas [5] Cybercriminal	**C** (1.0)	**H** (1.0)	**I** (1.73)	**T** (1.73)	**N** (3.0)	**B** (3.16)
Cardenas [5] Insider	**I** (1.0)	**H** (3.60)	**C** (3.74)	**T** (4.12)	**B** (4.24)	**N** (4.24)
Cardenas [5] NationState	**N** (0.0)	**I** (1.0)	**H** (1.0)	**T** (1.0)	**B** (1.0)	**C** (2.0)
Cardenas [5] Terrorist	**T** (2.44)	**N** (3.0)	**I** (3.16)	**H** (3.46)	**C** (3.74)	**B** (5.29)
Corman [9] AdaptivePersistent	**N** (1.41)	**I** (2.0)	**T** (2.0)	**H** (2.44)	**C** (2.44)	**B** (3.74)
Corman [9] Hacktivist	**H** (1.41)	**C** (1.41)	**I** (2.0)	**T** (2.0)	**N** (3.46)	**B** (4.24)
Corman [9] OrganizedCrime	**T** (2.0)	**I** (2.0)	**C** (2.44)	**H** (2.44)	**N** (2.82)	**B** (3.74)
Corman [9] Skiddie	**B** (1.73)	**T** (1.73)	**I** (1.73)	**H** (2.64)	**N** (3.0)	**C** (3.31)
Heckman [16] Hacktivist	**N** (1.41)	**C** (2.0)	**T** (2.23)	**H** (2.44)	**I** (2.82)	**B** (4.24)
Heckman [16] Hobbyist	**B** (1.73)	**C** (2.0)	**I** (2.23)	**H** (2.64)	**T** (3.0)	**N** (3.31)
Heckman [16] Insider	**I** (1.0)	**C** (1.41)	**H** (1.73)	**T** (1.73)	**B** (2.64)	**N** (2.64)
Heckman [16] NationState	**N** (1.41)	**H** (2.0)	**C** (2.23)	**I** (2.82)	**T** (3.16)	**B** (4.47)
Heckman [16] OrganizedCrime	**N** (1.73)	**C** (2.0)	**H** (2.23)	**I** (2.64)	**T** (3.31)	**B** (4.12)
Heckman [16] ScriptKiddie	**B** (2.0)	**I** (2.0)	**C** (2.23)	**T** (2.44)	**H** (2.82)	**N** (3.74)
Heckman [16] StructuredHacker	**C** (1.41)	**N** (1.73)	**H** (2.0)	**I** (2.23)	**T** (3.60)	**B** (3.87)
Heckman [16] Terrorist	**N** (1.41)	**C** (2.44)	**H** (2.82)	**T** (3.31)	**I** (3.46)	**B** (4.89)
Heckman [16] UnstructuredHacker	**H** (1.41)	**T** (1.73)	**C** (2.0)	**I** (2.23)	**B** (2.23)	**N** (2.64)
Le May [19] DisgruntledEmployee	**I** (2.0)	**H** (2.82)	**C** (3.0)	**T** (3.46)	**N** (3.87)	**B** (4.12)
Le May [19] LoneHacker	**C** (1.73)	**H** (2.0)	**T** (3.16)	**N** (3.31)	**B** (3.87)	**I** (4.0)
Le May [19] NationState	**N** (1.41)	**C** (2.23)	**H** (2.82)	**T** (3.87)	**I** (4.79)	**B** (5.09)
Le May [19] SystemAdministrator	**I** (1.73)	**H** (3.87)	**C** (4.0)	**T** (4.58)	**N** (4.69)	**B** (5.09)
Le May [19] Terrorist	**T** (2.23)	**H** (2.23)	**C** (2.23)	**B** (3.0)	**N** (3.60)	**I** (4.0)
Urbina [34] Insider	**I** (4.58)	**N** (5.56)	**H** (6.16)	**C** (6.24)	**T** (6.63)	**B** (7.0)
#Expected	21	0	0	2	0	0

B=BasicUser, C=Cybercriminal, H=Hacktivist, I=Insider, N=NationState, T=Terrorist, (Float)=Euclidean distance, X(x.x)=Expected mapping

profile to Heckman [16] Terrorist is Nation-State. Furthermore, the difference between an Hacktivist and Nation-State and Terrorist is not well defined. As [16] is an industrial white paper, it could be that the author's views are somewhat diverging from the academic security community.

In addition, we note that there are six cases in which a profile has the same distance to multiple archetypes. In that case, the archetypes cannot be distinguished only by the subset of dimensions considered by the profile analyzed. That could indicate that a) the profiles in the related work are vaguely defined, or b) our dimensions do not yet appropriately capture all aspects intended by the original authors.

5.5 APE

To support our work in this paper, we developed an interactive command-line tool using Python. The tool is called APE (Attacker Profiler Examiner) available as open source at scy-phy.net/ape. APE allows the application of our taxonomy to classify related work,

definition of own attacker profiles using our framework, and comparisons between profiles. Profiles can be exported to several different formats (e.g. WEKA .arff), and the profiles we defined in this paper are part of the tool.

We envision that other researcher can use our framework and APE to define constraints during the security analysis, verification, or testing of CPS. Most of the related work (e.g., in [19, 4, 35]) base their analysis on some constraints (the same applies for security protocols when the DY is assumed to control the network). One relevant example is the physical distance between the attacker and the CPS which has a severe impact on the physical layer interactions of the attacker. This and other dimensions have been used in a number of works (e.g., [1, 19, 4, 16]) to show different security flaws or attacks based on different profiles. Our framework supports the modeler or the security analyst in the generation of such constraints. In addition to theoretical analysis, our tool can be used to output constraints that can be applied when concretely testing a CPS.

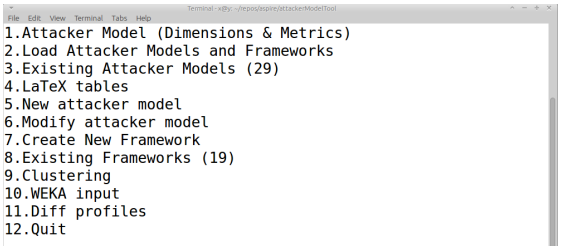

Fig. 2: Features available in Attacker Profiler Examiner

6 Conclusion

In this work, we discussed attacker models for security research, in particular ones for CPS. We started with a literature review, and defined a taxonomy of 10 different features that we applied to the literature. This lead us to the identification of discrepancies and commonalities between different works. In general, we grouped the reviewed papers into two main classes: publications that aim at profiling attackers, and publications that propose an attacker model. To describe both classes in more detail, we discussed profiles and dimensions we found in the literature. We argued that these classes and dimensions should be the starting point for a definition of a comprehensive attacker model.

We then defined a general attacker framework that attempts to capture the information provided by related works on their model. We mapped the 23 attacker profiles from related work into that framework, and defined a distance metric that allows us to compute overlap and discrepancies between attacker models in related work. We attempted to use machine learning approaches to cluster the attacker models from related work, but did not obtain good results so far. Motivated by that, we manually constructed 6 attacker profiles, and show that they match the profiles from the literature in 21/23 cases.

In addition to our theoretical results, we wrote a software tool to capture our attacker framework, and profiles proposed by us and the related work. The tool showcases some of the benefits of more structured approaches to attacker models: we use it to compare different profiles, export profiles to tools such as WEKA, and produce structured representations such as the tables in this work.

References

1. S. Adepu and A. Mathur. An investigation into the response of a water treatment system into cyber attacks. In *IEEE International Symposium on High Assurance Systems Engineering (HASE)*, 2015.
2. S. Amin, X. Litrico, S. Sastry, and A. Bayen. Cyber security of water SCADA systems; Part I: Analysis and experimentation of stealthy deception attacks. *Control Systems Technology, IEEE Transactions on*, 21(5):1963–1970, 2013.
3. S. Amin, X. Litrico, S. Sastry, and A. Bayen. Cyber security of water SCADA systems; Part II: Attack detection using enhanced hydrodynamic models. *Control Systems Technology, IEEE Transactions on*, 21(5):1679–1693, 2013.
4. D. Basin, S. Capkun, P. Schaller, and B. Schmidt. Formal reasoning about physical properties of security protocols. *ACM Transactions on Information and System Security (TISSEC)*, 14(2):16, 2011.
5. A. A. Cárdenas, S. M. Amin, B. Sinopoli, A. Giani, A. Perrig, and S. S. Sastry. Challenges for securing cyber physical systems. In *Workshop on Future Directions in Cyber-physical Systems Security*. DHS, July 2009.
6. A. A. Cárdenas and J. S. Baras. Evaluation of classifiers: practical considerations for security applications. In *AAAI Workshop on Evaluation Methods for Machine Learning*, 2006.
7. A. A. Cárdenas, T. Roosta, and S. Sastry. Rethinking security properties, threat models, and the design space in sensor networks: A case study in SCADA systems. *Ad Hoc Networks*, 7(8):1434–1447, 2009.
8. S.-Y. Chang, Y.-C. Hu, and Z. Liubook. Securing wireless medium access control against insider denial-of-service attackers. In *Proceedings of Conference on Communications and Network Security (CNS)*, 2015.
9. J. Corman and D. Etue. Adversary ROI: Evaluating security from the threat actor's perspective, 2012.
10. S. Daniel, S. Matthias, and H. Matthias. Lockpicking physical layer key exchange: Weak adversary models invite the thief. In *Proc. ACM Conference Wireless Security (WiSeC)*, 2015.
11. D. E. Denning. Activism, hacktivism, and cyberterrorism: The internet as a tool for influencing foreign policy. In *Networks and Netwars: The Future of Terror, Crime, and Militancy*. RAND Corporation, 2001.
12. D. Dolev and A. C. Yao. On the security of public key protocols. *IEEE Transactions on Information Theory*, 29(2):198–207, 1983.
13. P. Esfahani, M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson. Cyber attack in a two-area power system: Impact identification using reachability. In *American Control Conference (ACC)*, pages 962–967, June 2010.
14. M. D. Ford, K. Keefe, E. LeMay, W. H. Sanders, and C. Muehrcke. Implementing the ADVISE security modeling formalism in möbius. In *IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, 2013.
15. M. A. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The WEKA data mining software: an update. *SIGKDD Explorations*, 11(1):10–18, 2009.
16. R. Heckman. Attacker classification to aid targeting critical systems for threat modelling and security review, 2005. www.rockyh.net/papers/AttackerClassification.pdf, last visited Oct 23 2015.
17. E. D. Knapp and R. Samani. *Applied Cyber Security and the Smart Grid*. Elsevier Syngress, 2013.
18. M. Krotofil, A. A. Cárdenas, B. Manning, and J. Larsen. CPS: Driving cyber-physical systems to unsafe operating conditions by timing dos attacks on sensor signals. In *Proceedings*

of the Computer Security Applications Conference (ACSAC), pages 146–155, New York, NY, USA, 2014. ACM.
19. E. LeMay, M. D. Ford, K. Keefe, W. H. Sanders, and C. Muehrcke. Model-based security metrics using adversary view security evaluation (ADVISE). In *Proceedings of Conference on Quantitative Evaluation of Systems, QEST*, 2011.
20. J. Lin, W. Yu, X. Yang, G. Xu, and W. Zhao. On false data injection attacks against distributed energy routing in smart grid. In *Proc. of the Conference on Cyber-Physical Systems (ICCPS)*, 2012.
21. Y. Liu, P. Ning, and M. K. Reiter. False data injection attacks against state estimation in electric power grids. *ACM Transactions on Information and System Security (TISSEC)*, 14(1):13, 2011.
22. J. Matusitz. Cyberterrorism: Postmodern state of chaos. *Information Security Journal: A Global Perspective*, 17(4):179–187, 2008.
23. T. R. McEvoy and S. D. Wolthusen. A formal adversary capability model for SCADA environments. In *Critical Information Infrastructures Security, (CRITIS)*, pages 93–103, 2010.
24. MITRE. Common attack pattern enumeration and classification (capec).
25. Y. Mo, T.-H. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, and B. Sinopoli. Cyber-physical security of a smart grid infrastructure. *Proceedings of the IEEE*, 100(1):195–209, 2012.
26. H. Orojloo and M. A. Azgomi. A method for modeling and evaluation of the security of cyber-physical systems. In *ISC Conference on Information Security and Cryptology (ISCISC)*, 2014.
27. R. Ottis. Theoretical model for creating a nation-state level offensive cyber capability. In *European Conference on Information Warfare and Security*, 2009.
28. P. Papadimitratos, M. Poturalski, P. Schaller, P. Lafourcade, D. Basin, S. Capkun, and J.-P. Hubaux. Secure neighborhood discovery: a fundamental element for mobile ad hoc networking. *Communications Magazine, IEEE*, 46(2):132–139, 2008.
29. T. Parker, E. Shadow, E. Stroz, M. G. Devost, and M. H. Sachs. *Cyber Adversary Characterization: Auditing the Hacker Mind*. Syngress Publishing Inc., 2004.
30. SPaCIoS. Methodology and technology for vulnerability-driven security testing (final version), 2014.
31. R. Taormina, S. Galelli, N. O. Tippenhauer, E. Salomons, and A. Ostfeld. Simulation of cyber-physical attacks on water distribution systems with EPANET. In *Proceedings of Singapore Cyber Security R&D Conference (SG-CRC)*, Jan. 2016.
32. A. Teixeira, D. Pérez, H. Sandberg, and K. H. Johansson. Attack models and scenarios for networked control systems. In *Proceedings of the Conference on High Confidence Networked Systems (HiCoNS)*, pages 55–64. ACM, 2012.
33. United States Environmental Protection Agency. Epanet: Software that models the hydraulic and water quality behavior of water distribution piping systems. www.epa.gov/nrmrl/wswrd/dw/epanet.html.
34. D. Urbina, J. Giraldo, N. O. Tippenhauer, and A. Cardenas. Attacking fieldbus communications in ICS: Applications to the SWaT testbed. In *Proceedings of Singapore Cyber Security R&D Conference (SG-CRC)*, Jan. 2016.
35. R. Vigo. The cyber-physical attacker. In *Proceedings of Workshop of Conference on Computer Safety, Reliability, and Security (SAFECOMP)*, 2012.

A Appendix: Subdimensions

We now show a section for each top-most dimension: Knowledge Section A.1, resources Section A.2 and Section A.3 and an itemize for each sub-dimensions. Each metric is defined between square brackets with the following order: $[1 < 2 < 3]$.

Finally, in A.4 we clarify the relation between a subset of our dimensions and time.

A.1 Knowledge

The knowledge dimension ([`low`, `medium`, `high`]) represents the understanding of the system under attack and the expertise of the attacker (as in, [12, 35, 19] to give some examples). The dimension is structured as follows.

- *Offensive* ([`basic`, `intermediate`, `advanced`]), determines the expertise of the attacker with regard to the attacks known, e.g., attack methodologies, attack patterns [24]. It is composed by three sub-dimensions: *Physical* ([`basic`, `intermediate`, `advanced`]), *Network* ([`basic`, `intermediate`, `advanced`]) and *Software* ([`basic`, `intermediate`, `advanced`]) which can be used to define the offensive knowledge with a finer granularity considering different expertise of the attacker.
- *System* ([`basic`, `intermediate`, `advanced`]), expresses the knowledge of the system under attack/analysis, for example, the set of components of a CPS [35] or entities in a security protocol [12]. It is composed by three sub-dimensions: *Source code* ([`blackBox`, `grayBox`,`whiteBox`]), *Protocols* ([`blackBox`, `grayBox`,`whiteBox`]) and *Credentials* ([`user`, `supervisor`, `admin`]) which can be used to define the knowledge with respect to the these three general aspects of CPS and systems in general.

A.2 Resources

The resource dimension ([`low`, `medium`, `high`]) represents the resources available to the attacker [16, 10, 29]. It can be used to limit the practical capabilities of the attacker. This dimension is widely accepted in our related work. This dimension is structured in the following different sub-dimensions.

- *Distance* ([`far`, `near`, `physicalAccess`]), expresses the physical distance of the attacker with respect to the target and may limit his interactions with the system. This is particularly important with respect to CPS which can be isolated from the Internet or when using WiFi networks, e.g., [10].
- *Manpower* ([`low`, `medium`, `high`]), represents the human resources available to the attacker. This can be used to distinguish between lone attackers and (small to large) groups.
- *Tools* ([`basic`, `intermediate`, `advanced`]), also know as attacklets, defines which types of tools are available to the attacker for performing the attack.
- *Financial support* ([`low`, `medium`, `high`]), expresses which is the budget that an attacker has in order to perform an attack. Discriminating between attacker with low or high budget can be helpful, e.g., for risk assessments.
- *Effort* ([`low`, `medium`, `high`]), defines the effort an attacker will put into his attacks. How deeply the attacker will explore possible attacks and different ways to attack the system.

A.3 Psychology

The psychology dimension ([`weak`, `average`, `strong`]) represents a set of aspects which are not directly related to the knowledge or resources of the attacker. These aspects are related to the motivations or behavioral aspects of the attacker [19, 6, 5, 16]. This dimension is structured in the following different sub-dimensions.

- *Aim* ([`knowledge`, `manipulation`, `damage`]), identifies which parts of the system are more likely to be interesting for the attacker. There are two sub-dimensions which discriminates between virtual and physical components: *Virtual* ([`knowledge`, `manipulation`, `damage`]) and *Physical* ([`knowledge`, `manipulation`, `damage`])
- *Periodicity* ([`once`, `anytime`, `continuous`]), defines which is the frequency with which an attacker will try to attack the system. Some system are more incline to be attacked than other, for example, if a CPS is exposed on the Internet the periodicity of attacks will be higher with respect to a CPS isolated from the Internet.
- *Determination* ([`firstAttempt`, `severalAttempts`, `untiring`]), Defines how long the attacker will perform the attacks on the system. As an example, the effort of the attacker should grow after each assessment performed on a system.
- *Honesty* ([`malicious`, `benign`]), discriminates between benign (White Hat attackers or "honest but curious" [16]) and malicious attackers (Black Hat).
- *Camouflage* ([`visible`, `stealthy`, `invisible`]), is the ability or preference of an attacker to stay hidden.
- *Strategy* ([`random`, `brute-force`, `structured`]), refers to the attack strategy adopted by the attacker. Random if an attacker will randomly select some attacks or some attack patterns. Brute-force when the attacker tries all possible attack pattern and structured when an optimal subset of attack patter is chosen.
- *Aim-Physical* and *Aim-Virtual* ([`low`, `medium`, `high`]), represent the objective of the attacker with respect to physical and virtual components. They are both divided into the three sub-dimensions: *Integrity*, *Confidentiality*, *Availability*.

A.4 Time

As depicted in Figure 3, different aspects related to time in our framework have been captured as a combination of the three dimensions: effort, periodicity and determination. The effort represents how deeply the attacker will try to attack the system during each attack. The determination is the duration of each attack and the periodicity expresses the distribution of attacks over time.

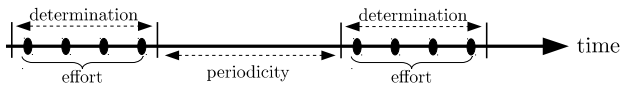

Fig. 3: Time related metrics