Challenges and Peculiarities of Attack Detection in Virtual Power Plants

Towards an Advanced Persistent Threat Detection System

Robin Buchta Felix Heine Carsten Kleiner
Hochschule Hannover
University of Applied Sciences and Arts
Hanover, Germany
{robin.buchta, felix.heine, carsten.kleiner}@hs-hannover.de

2022 IEEE 29th Annual Software Technology Conference (STC) | 978-1-6654-8864-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/STC55697.2022.00019

***Abstract*—Currently, there are no mission-capable systems that can successfully detect advanced persistent threats (APTs). These types of threats are hazardous in critical infrastructures (CIs). Due to the integration of operational technology (OT) and information communication technology (ICT), CI systems are particularly vulnerable to cyberattacks. In addition, power systems, in particular, are an attractive target for attackers, as they are responsible for the operation of modern infrastructures and are thus of great importance for modern warfare or even for strategic purposes of other criminal activities. Virtual power plants (VPPs) are a new implementation of power plants for energy management. The protection of virtual power plants against APTs is not yet sufficiently researched. This circumstance raises the research question - What might an APT detection system architecture for VPPs look like?**

Our methodology is based on intensive literature research to bundle knowledge from different sub-areas to solve a superordinate problem. After the literature review and domain analysis, a synthesis of new knowledge is provided in the presentation of a possible architecture. The in-depth proposal for a potential system architecture relies on the study of VPPs, APTs, and previous prevention mechanisms. The architecture is then evaluated for its effectiveness based on the challenges identified.

The proposed architecture combines concepts such as defense-in-depth and breath with situation awareness, and the observe, orient, decide, and act loop. Furthermore, a combination of traditional detection methods with graph analysis in the architecture is targeted to meet the challenges and peculiarities of VPPs and APTs.

***Index Terms*—Attack Detection, Virtual Power Plants (VPP), Advanced Persistent Threats (APT), Critical Infrastructure**

I. Introduction

The increasing spread of decentralized energy resources has led to the concept of Virtual Power Plants (VPP) gaining acceptance and pointing the way to the future [1]. VPPs bundle decentralized energy resources (DERs) and assume the role of central control and monitoring [2]. VPPs are part of critical infrastructures (CI) and, thus, sensitive and strategic targets [3], [4]. Attacks on CI systems can cause immense damage, such as economic slowdowns [5]. There is also an increasing awareness that there is no solution that prevents entirely from cyberattacks; some can only be mitigated [6]. A resilient system needs efficient mitigation and robust prevention.

Recent reports [7], [8], e.g., one from the European Council on Foreign Relations (ECFR) [9], highlighted the critical security threat, especially for the energy sector. A wide variety of attackers are conceivable, each pursuing different goals [10].

The Canadian Centre for Cyber Security report shows examples of attacks by both cybercriminals and state-sponsored actors on the electricity sector [11]. The case of Edesur S.A., dated July 07, 2020, can be cited as a local energy distributor in Buenos Aires from Argentina for cyber-criminal activities. In this case, ransomware disrupted some IT system functions [12]. So far, state-sponsored activities have increased espionage [11], and if an attack meets specific criteria, it can also be considered a declaration of war [13], [14].

An excerpt of meaningful activities is a case from 2019 where actors sponsored by the Chinese state attacked several US utility providers. Remote access trojans (RAT) were installed, allowing the infected systems to be controlled from the outside and the possibility to exfiltrate information. [15] In 2018, Russian state-sponsored actors spied on several areas of CI, including the energy sector. The operation aimed to obtain information about cyber-physical systems (CPS). Among other things, the attackers tapped the configuration and profile files of ICS or SCADA systems. [16] The examples show, that various techniques, tactics, and procedures (TTP) will enable it to infiltrate

978-1-6654-8864-8/22/$31.00 ©2022 IEEE
DOI 10.1109/STC55697.2022.00019

the target's systems.

The use of sophisticated malware and different TTPs related to a specific target are characteristics of an APT, which poses a particular threat due to its properties. [17], [18] Thus, the capabilities, also outside the previously mentioned espionage examples, of an APT are visible in the famous case of Stuxnet. Attackers exploited multiple zero-day vulnerabilities and created specific malware for the target to sabotage the CI of Iran's uranium enrichment program over a long time. Stuxnet was also discovered forensically in the post-attack phase after the damage to the natural environment had occurred. [19], [20]

In particular, DERs have contributed to the establishment of a new type of power plant, the VPPs [21], which brings new and expanded requirements for attack detection.

The points just mentioned, the APTs and the particular security situation of VPPs, lead to the following research question:

"RQ: What does an APT detection system architecture for VPP look like?"

The RQ leads to the following three sub-questions:

- SQ1: What are the characteristics and unique features of VPPs in contrast to pure IT-System?
- SQ2: What is the threat landscape for VPPs, and what are the unique characteristics of APTs in this domain?
- SQ3: What makes it so difficult to prevent a APT in general?

After answering these sub-questions, the challenges and particularities are identified, which serve as the basis for presenting a possible architecture for answering the research question.

The contribution of this paper includes the following main aspects:

- Based on an extensive literature review, a detailed overview of the concept of VPPs, their threat landscape, and possible detection measures of APTs in the domain. (Section II, III, IV)
- A requirements elicitation for an attack detection system architecture from three perspectives. The VPPs, the APTs, and previous detection systems. (Section II, III, IV)
- A proposal for an attack detection system architecture that meets the requirements raised in the paper. (Section V

The methodology is based on literature research to bundle knowledge from different sub-areas to solve a superordinate problem; this reflects in the structure of the RQ and SQ. After the literature review and domain analysis, a synthesis of new knowledge is provided in the presentation of a proposed architecture.

The structure of the work consists of answering the sub-questions and then the research question, starting with Section II. Section III examines the drivers of VPPs and then focuses on APTs, defining and describing them and analyzing their characteristics. Section IV answers the third sub-question (SQ3) by considering the characteristics and challenges of the previous sections along with the defensive methods. Section V focuses on answering the research question (RQ) by combining the previously elaborated specifics with the defense possibilities. A theoretical evaluation follows this in section VI of the architecture to determine whether it meets the challenges raised. A description of further planned work follows in section VII. Section VIII reviews known work in this area, and section IX summarizes the results of this research.

II. Subject Domain to be considered

A clear definition of VPPs is needed to identify the particularities and challenges (C) of the subject domain (SD). The presented definition focuses on the domain's essential aspects, leaving the VPP operational details aside.

> A virtual power plant (VPP) consists of a portfolio of decentralized energy resources (DERs) connected via information and communication technology (ICT). The interconnection of DERs is monitored and coordinated by an energy management system (EMS).

Descriptions and illustrations from [2], [22], and [23] forms our definition.

The DER portfolio is based upon the operational technology (OT) layer (SD_C1). The EMS systems and ICT represent the information technology (IT) layer (SD_C2). The EMS is the central monitoring and control element for the entire VPP (SD_C3). OT and IT interconnect, meaning that the VPPs assign to the CPSs (SD_C4).

CPSs have the particularity to interact with the natural world by integrating computing resources and networks [24] (SD_C5). An EMS is a type of supervisory control and data acquisition (SCADA) system, which are industrial control systems (ICSs) [13].

The architecture, in figure 1, consists of a physical layer and three system layers: the sensor/actuator, the network, and the control layer (SD_C6). In addition, there is the information layer, which consists of two information flows: The lower layers send measured values to the upper ones and vice versa for the operation commands.

General factors in which OT and IT differ are listed below and adapted from the work of [13] (SD_C7).

- Operational objectives: IT systems have data processing as their goal, while OT systems have control of processes as their goal.
- High availability: The availability differs between IT and OT as many OT systems, especially critical infrastructure, may have almost no downtime.
- Geographical position: While IT systems are often centrally operated at a geographically imposed location, OT systems are often distributed and can span

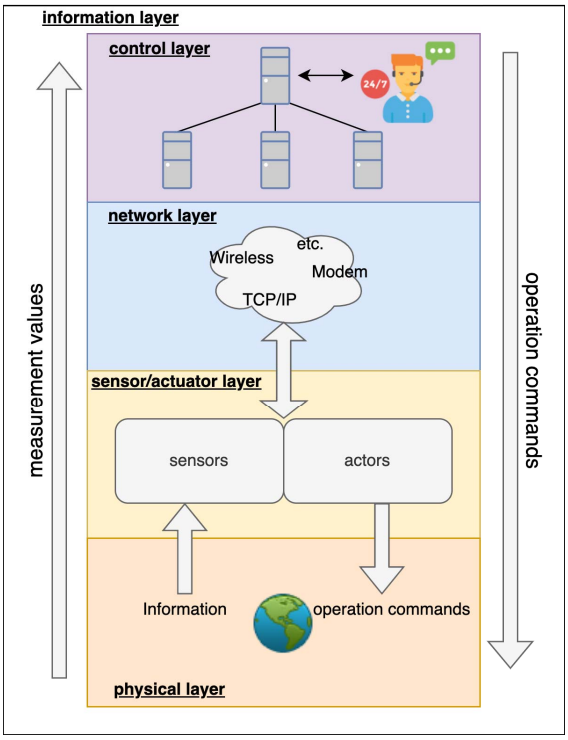

Fig. 1. Architecture of a CPS. Adopted from [25].

long distances, requiring different communications technologies.

- Technical distinctions: IT systems are pretty standardized from the hardware, protocols, and operating systems, while OT systems widely use case-specific hardware and communication methods and lack standards.
- Real-time performance: OT systems often require real-time performance since it makes a big difference when, for example, a command arrives at the plant. The command can be benign out one second and cause a critical error the next.
- Life cycle: IT systems were rapidly changing and replaced more frequently because it pays to do so. On the other hand, OT systems are often expensive to purchase and risky to replace due to the expected high availability. Therefore, the life cycles of OT systems are often high and can be in operation for more than 20 years [13], [26].
- Financial possibilities, manufacturers, suppliers, and the different responsibilities and expertise in the areas are further factors affecting OT, often differentiated from IT.

From a security perspective, the differences are even more significant. OT systems' design is usually unsuitable for operation in insecure networks [13], [26], so the security situation for CPSs is often highly challenging (SD_C8). Security in the IT environment mainly uses the CIA-Triad (confidentiality, integrity, and availability) [26], [27]. There is a need for an extended security model for OT security, which considers the physical processes and the specifics of the environment. Resilience is one measure of this [26]. Resilience also applies well to the National Institute of Standards and Technology (NIST) defined security goals for smart grids, which address: Maintaining safety, power system reliability, resilience, and supporting grid modernization [28]. The same applies to the shift in the importance of the CIA-Triad criteria. With OT, availability is the most important, followed by integrity, and confidentiality is the least important. [27] Operational resilience (OR) is defined by CERT as follows: [29]

> "The emergent property of an organization that can continue to carry out its mission in the presence of operational stress and disruption that does not exceed its operational limit."

The authors make a deliberate distinction between resilience and OR. OR can be attributed to the fact that resilience defines a physical property that states that the material returns to its original shape [29].

The presented concepts and characteristics of the systems contribute to the answer to SQ1. In conclusion, VPPs are composed of OT and IT systems. Therefore, the specifics of both worlds must get addressed in combination.

III. Threat Environment of the Domain

In addition to the specific threat constraints inherent in the design of a VPP described in section II, this section addresses the CI domain's threat environment (TE). The CI domain, especially the critical energy infrastructure, is, according to [7], [8], one of the most popular targets of global cybercrime, also popular for APTs [30] (TE_C1). All conceivable forms occur as attackers, from individuals who want to try their hand to state-sponsored/organized groups [10] (TE_C2). These include corporations, organizations, "common" cybercriminals, insider threat agents, hacktivists, nation-states, terrorists, and cyber fighters (e.g., APT groups) [18]. The typical motives are: creating financial gains, capability demonstration, securing competitive advantages, private, social, political, or national interests, intimidation, or even destruction [18]. In a potential cyberwar, critical infrastructure is an important strategic target [31]. In the sense of war, the motivation can also be self-defense [18], respectively, the legitimacy for full-scale war, to justify using force [31].

Often, the system landscapes have grown and secured with outdated security concepts, such as *security by obscurity* (TE_C3). In addition, secure communication channels, authentication and authorization mechanism, and systems monitoring are often lacking. [13], [32], [33] In addition to cyberattacks, CPSs are also vulnerable to physical attacks, mainly due to the geographic distribution of the systems [13] (TE_C4).

TABLE I
COMPARISON OF TRADITIONAL AND APT ATTACKS. ADOPTED FROM [40].

	Traditional attacks	APT attacks
Attacker	single person	powerful groups
Target	unspecific	single specific target
Purpose	financial benefits, demonstrating abilities	competitive advantages, strategic benefits
Approach	single-run, "smash and grab"	repeated attempts, stays "low and slow", adapts to resist defenses

A. Advanced Persistent Threats

APTs severely threaten CIs [34], [35].

The name of the threat is composed of three parts:

- Advanced: Stands for the fact that the attack has more resources than conventional attacks. This resource advantage extends to unknown vulnerabilities - zero-day exploits and highly sophisticated tools. The knowledge base created by an interdisciplinary team or the combination can also fill this attribute. [17], [36] An attack scenario known to the target is no longer part of the attack repertoire of new APT campaigns.
- Persistent: The threat does not stop until the attacker reaches his target. He will try to penetrate the system again and again [17].
- Threat: APT campaigns typically threaten the loss of sensitive data or compromise critical components and processes [17]. In addition, some targets, such as CIs, also offer the threat of disruptive actions [35].

The NIST assigns the following attributes to an APT [37] (TE_C5):

- Constantly pursues his goals and also repeats his efforts.
- Adapts to the defensive actions of the target.
- Is not deterred from achieving its goal.

These attributes derive that protection is almost impossible since the defender must constantly protect himself in all directions. In contrast, the attacker only has to find one possible weak point. [35] The mentioned attributes also apply to the previously, in section I, mentioned hypothesis of some authors that protection may be impossible and mitigation is required. APTs are different from traditional attacks and are subordinate to targeted attacks [38]. Conventional attacks often have a generality and are easy to detect because they do not try to operate covertly [39]. Table I shows other differences in the attack types (TE_C6).

APTs have several stages and belong to multi-stage attacks [41] (TE_C7). The theoretical level number of such an attack often differs in the literature. Often a number between three and eight is mentioned [17]. The actual number used depends on the attacker and the target. In this work, we would like to follow a survey [17] that addresses this problem in conjunction and then decides on five steps. These steps are:

- Step 1: Reconnaissance: Spying on and becoming familiar with the target is the first phase of a more complex attack. The more intensively this phase is pursued, the faster and more successful the attack. There are different ways to go through this stage. Examples are intelligence activities in general, like expert knowledge, insider recruitment, open-source intelligence (OSINT), and phishing. [17], [42], [43]
- Step 2: Establish Foothold: The second phase describes the successful penetration of the target system or network and often includes establishing long-term access [17], [44].
- Step 3: Lateral Movement / Stay Undetected: The third phase describes many activities that can be very different. Typical activities include increasing privileges, spying on other target systems, and exploiting them. [17], [44] Depending on the attack, this phase can also mean securing long-term access to the systems and adapting to changes in the targets. [17]
- Step 4: Exfiltration/Impediment: In this phase, pursue the primary goal of the attack. The goals can vary but usually involve data exfiltration or essential component compromise. [17]
- Step 5: Post-Exfiltration/Post-Impediment: The last phase involves very different activities based on the motivation of the attacker and the previous steps. Further data are collected, components are damaged, or traces are covered. [17] Likewise, for further attacks, a new backdoor is installed [44].

From a general perspective, APT attacks have three goals (TE_C8). They steal information (goal one). They manipulate the functionality or control the system (goal two). Goal three aims to secure a good position for the attacker in the future so that the aggressor can exfiltrate the target quickly and efficiently if needed. [17]

Figure 2 shows the attack tree of an APT. The targets mentioned first are the focus of the figure. The figure includes the third target but would not enter the fourth stage in this scenario. The tree would look different according to the chosen techniques depending on the selected abstraction level. For example, the Cyber Kill Chain, the Mitre ATT&CK, the STRIDE framework, or a combination of these provide the groundwork [45].

This part references one-half of SQ2. The section identified the characteristics of APTs for VPPs and noted that early detection and prevention of these attacks are critical.

B. Target range of VPPs

As shown in figure 1, CPSs are, by architecture, a layered system where one layer is dependent on the other. Based on the particular architecture of VPP, multiple attack vectors can be extracted [2], [46] (TE_C9):

- Direct attacks on IT components. For example, these attacks can occur on modern SCADA systems or the EMS system. Direct attacks can manipulate the operation of the systems. In figure 1, the control layer.

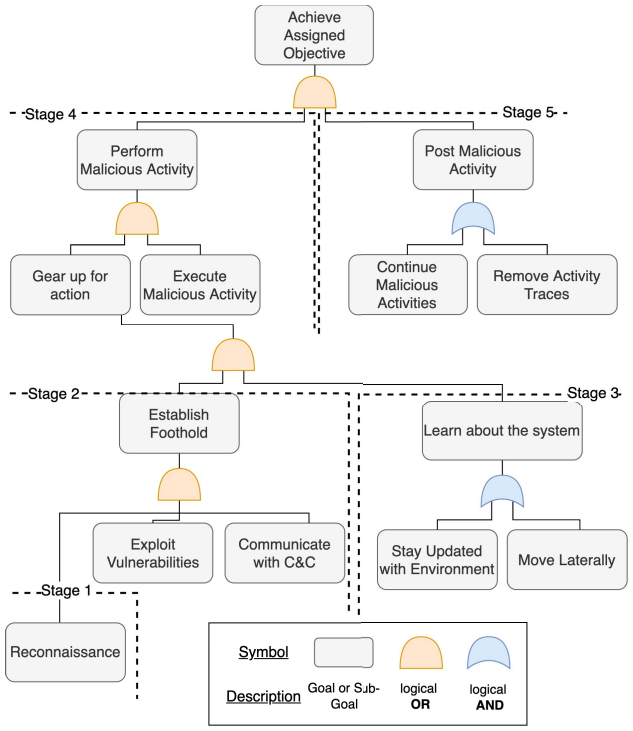

Fig. 2. APT Attack Tree. Adopted from [17].

- Indirect attacks on IT components. Through false data injection, which is possible from the outside, the attacker can control the systems to make decisions in their favor. This attack is located in the architecture's information layer and relates to the network layer, but attackers can launch this from the lower layers.
- Attacks on the ICT. ICT is essential for the systems because, as described in section II, even minor delays can lead to fatal errors.
- Attacks against the OT systems. Direct manipulation of individual OT systems connected to the network can have different consequences. Depending on the component and the level at which the component operates. These attacks target the sensor and actuator layers.

The various components and the dependencies between them increase the attack surface of the VPPs enormously [46] (TE_C10). In addition, VPPs often pursue very different goals [2]. Only the physical layer of the architecture is not mentioned in the listing because they are not crucial for this work. The effects of such an attack reflect in the data. Attacks within the layer can cause cascading damage in the other layers and vice versa [13].

The extended attack surface affects a VPP in terms of an APT, see figure 2, in such a way that both the damage caused can be more severe and more entry points can be used to achieve the goal.

What an APT looks like for VPPs depends on the attackers, the target, and the defenses (TE_C11). Known APT campaigns are very different, and new ones will not be similar. Likewise, detection teams must expect that attackers have already prepared corresponding backdoors in the systems. The expanded threat landscape of VPPs poses a great danger, especially in this aspect. The general threats explanation in the introduction of this section, along with the description of the attack surface, answers the remaining half of SQ2 in detail.

In summary, VPPs are a critical and desired target, and all kinds of motivations and attacker types are conceivable. In addition, due to the unique architecture of these systems, different attack vectors are feasible, which have cascading effects on each other. A direct attack on one layer is often an indirect attack on the others. Since APTs are very dangerous and considered particularly difficult to detect [17], [47], this work focuses on detecting them.

IV. Defense Methods against APTs

Successfully defending against APTs is considered an unsolved problem [17], [48], [49]. Therefore the following section examines previous approaches and defense methods (DM). High situational awareness (SA) is needed to properly assess the situation and take appropriate action to protect against attacks successfully (DM_C1). Detecting attacks is an essential part of defense; only when operators know something is suspicious can they take action.

Conventional attack detection methods are not able to detect an APT. These are primarily based on human-generated cyber threat intelligence and aim to accurately detect and report specific indicators of compromise (IoC) (DM_C2). [17], [49]

Attack detection systems divide into three classes: signature-based intrusion detection systems (SIDS), anomaly-based (AIDS), and hybrid. The AIDS category classifies systems based on statistics, knowledge, or machine learning (ML). [50] SIDS can be ruled out as a single solution or the core building block because APTs do not have significant IoCs. Some approaches, such as HOLMES [51],use a hierarchy of small-step rules to map more significant issues, but there is a significant risk of missing individual activities. While anomaly-based techniques can detect new types of attacks, including new TTPs of APTs, they produce many false positives and unclassified alerts (DM_C3). In addition, they usually require a training phase and a normal behavior model (DM_C4). [50] Statistical approaches are also less suitable for APT detection since only a few attacks are known, and the attack techniques are also unpredictable, see the properties of an APT in section III-A. For the same reason, knowledge-based approaches are also inadequate for the task. Therefore, only the ML category of AIDS remains that meets the APT detection requirements (DM_C5). ML makes it possible to identify unknown deviations from normal (trustworthy) behavior. Thus, zero-day exploits

and internal employees trying to infiltrate the system can be exposed. However, if the operators do not determine normal behavior in advance, or this is impossible, the method cannot meet the detection performance. Further, the detection methods can be classified based on their position in the system. A typical distinction is between network-based IDS (NIDS) and host-based IDS (HIDS). With NIDS, one quickly achieves a general presumption but has little information about actual activities (DM_C6). The overall view is low, while HIDS provides more details [50]. Due to the particular requirements in our subject domain (section II), HIDS cannot achieve complete coverage since such systems rarely exist for OT, and operations often cannot be performed for technical reasons (DM_C7).

All sources of information are valuable, so an all-encompassing solution must be considered. Provenance tracking systems, for example, are increasingly being considered. Correlation between events from different systems is essential for selecting the detection method (DM_C8). The detection system must perform cognitive tasks, combining individual clues to construct incriminating events. While only AIDSs can detect new TTPs, distinguishing APTs, SIDS could provide valuable partial information. Furthermore, AIDS results can usually only serve as clues and do not provide evidence (DM_C9) [52].

In total, the APT defense must combine several concepts. The basis is an intensive monitoring strategy of the systems to have any chance to detect all steps of the attacker theoretically (DM_C10). Building on the monitoring, detection at several different locations is required, which relies on its core machine-learning-based anomaly detection. The research comprises different aspects so that both technical tools, such as algorithms and human-performed processes, interact (DM_C11).

As mentioned, it is vital to increase SA to make the system more resilient. SA means that the focus is on the current status of the overall system, and thus not only prevention but also the restoration of a safe state can be a consequential action. Nevertheless, both methods usually require a pre-designed detection of abuses. [17] Subsequently, mitigation methods are required to fend off the detected attack. Mitigation methods are not part of the work due to the complexity of the detection itself.

The essence of the challenges for defense is that success lies in robust detection. However, this isn't easy to implement because the characteristics of APTs do not work with the assumptions of previous detection systems. In addition, complete coverage of the systems is preferable, but this isn't easy to achieve due to the challenges within the VPP domain. Thus comes the move toward a resilient system, away from strict adherence to CIA security goals. Protecting APTs, especially in the area of VPPs, is complex and an unsolved problem. The section has explained many founders for this; thus, the SQ3 is answered.

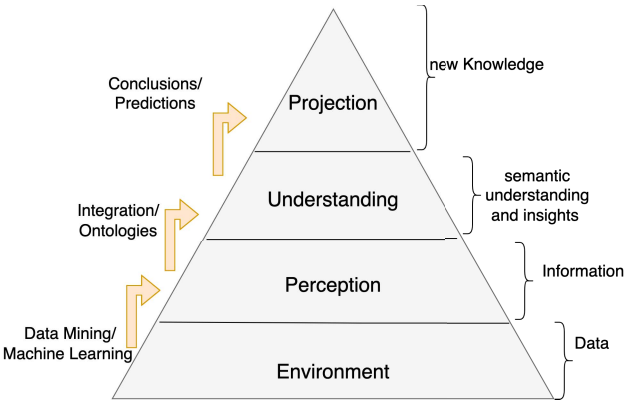

Fig. 3. Situation Awareness. Adopted from [56].

V. Detection system architecture

All the previously identified particularities and challenges from the three perspectives (SQ1 - SQ3), subject domain II, threat environment III, and defense methods IV have led to the proposal of an architecture for a detection system (RQ).

The goal is to identify the attacker's TTP. The TTPs are considered "tough" on the pyramid of pain [53]. In the detection maturity level model (presented in [54], based on [55]), the proposed architecture would be to cover levels four to six.

A. Important Concepts

To achieve this goal, the concept of SA is at the forefront, as depicted in figure 3. The observation of the environment, also mentioned in section IV, is done through an intensive monitoring approach. A defense-in-breath[1] strategy is one way to accomplish this purpose. All necessary data is collected and processed into information by data mining and machine-learning methods.

Then, the information at the perceptual level is brought to the comprehension level by integrating and adding ontologies. The integration looks like linking different sources and using other sources to enrich and better understand the current information. Similarly, the information is brought to the understanding level by adding ontologies. For example, the classification into cyber kill chain phases could be used in our domain to add additional knowledge to events. Subsequently, this understanding allows for making deductions and predictions.

[1]Definition defense-in-breath: "A planned, systematic set of multi-disciplinary activities that seek to identify, manage, and reduce risk of exploitable vulnerabilities at every stage of the system, network, or subcomponent life cycle (system, network, or product design and development; manufacturing; packaging; assembly; system integration; distribution; operations; maintenance; and retirement). [48]"

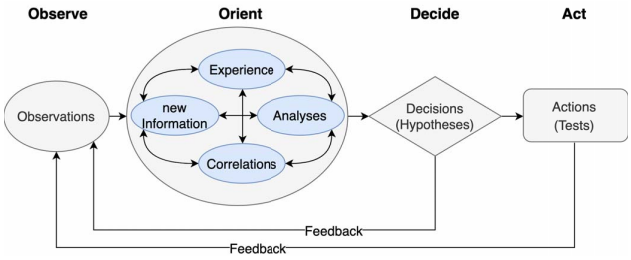

Fig. 4. Orient, Observe, Decide and Act Loop. Adopted from [57].

Considering a defense-in-depth strategy[2], we would like to extend the concept by the observe, orient, decide, act (OODA) loop, visualized in figure 4. The loop promotes the interaction of data, algorithms, and humans. Automation is desirable, but humans still play a significant and essential role, especially in the decision, act, and feedback loop. Automating the observe and orient phases is the decisive point because the better this is, the more successfully the human can complete the following steps. The combination of strategies leads to NIST's general recommendation on security risk management [48] recommendation.

Current research has pointed out two possible ways to address the APT "low and slow" property, once top-down and bottom-up approaches. Top-down approaches use attacks and cyber threat intelligence to map known knowledge to current events. Bottom-up approaches correlate low-level events to generate higher-level ones, primarily by provenance tracking [49]. Top-down examples are HOLMES [51], Conan [58], and the approach of Zou et al. [49]. Approaches based on human-generated cyber threat intelligence (CTI) are only helpful for applying already known knowledge, which is no protection against new sophisticated attacks like APTs. Examples of bottom-up approaches include Poirot [59], Anubis [60], and Unicorn [47].

Unicorn is one of the most promising detection methods so far. They can bypass the time factor by using its provenance graph. A provenance graph is a type of data management in which the system landscape is mapped using edges for relationships and nodes for entities. Within the graph, connections between the nodes can then be recognized, and the origin can be determined. Thus, despite a defense-in-breath approach, they can show hidden connections between activities otherwise lost in methods that consider, e.g., classical windowing or single events. Therefore, this proposed approach of this paper also wants to use nodes and edges as the data basis. A hierarchical provenance graph is preferred for the specifics of the VPPs. A mixture of HIDS and NIDS must be used to cover the system landscape. This graph goes into detail at the nodes, so more information is available. The top layer tries to represent all systems on a high abstraction level.

Furthermore, not only the correlations of the individual events are interesting, but also the individual properties of the events. In the sense of a defense-in-depth approach, the events are considered individually in parallel. This parallel view is based on the concepts of Anubis [60], where an autoencoder takes two inputs, one from a provenance graph and one event-based.

B. Architecture

To still use human knowledge and other methods well, the architecture includes a pre-analysis, which holds enrichment processes, plausibility checks, signature-based attack detection, and simple anomaly detection. The pre-analysis should have a plug-able architecture, where different methods can subscribe to the events under consideration and process them append-only. Figure 5 shows the architecture. The architecture proposes how a system can detect APTs in a complex domain such as VPPs. It combines the concept presented in figures 3 and 4 with the defense-in-breath and defense-in-depth strategies mentioned earlier.

The architecture consists of data sources, which are then passed through pre-processing and pre-analysis. At this point, enrichers add valuable state information to the collected data points, plausibility checks are performed on the data points, and various analyses of anomaly detections, rule-based detections, etc., are performed. Here, everything is optional and must be adapted to the data sources. This leads to many indications and also some evidence, which can also be reported directly. Subsequently, the data sources that have undergone pre-analysis and pre-processing are fed into the main analysis. Here, algorithms are used that specifically take into account the interaction of the different sources with their clues. Finally, the reported alarms and indications are reviewed by analysts and operators. Actions are decided, and, if necessary, feedback is sent to the previously involved systems to point out false alarms and positive hits.

Figure 5 shows the phases of SA, from Environment through Perception and Understanding to Projection. The SA phases also contain the OODA loop's logical steps. The Environment phase contains the Observations. The Perception phase represents the Orientation step. In addition, the step also reaches into the Understand phase, as more complex cognitive processes are triggered here to gain understanding. Decisions and actions are also part of two phases, Understanding and Projection. The understanding phase can also generate hypotheses, while the projection phase generates hypotheses and decisions. The action step is indirectly found in the projection phase since the conclusions and decisions also include actions. The feedback loop differs somewhat in architecture from the concept since all phases provide feedback on the previous phases. The data sources cover the defense-in-breath strategy

[2]Definition defense-in-depth: "Information security strategy integration people, technology, and operations capabilities to establish variable barriers across multiple layers and missions of the organization. [48]"

and the defense-in-depth strategy by further analysis and enrichment steps in the Perception phase, the cognitive processes in the Understanding phase, human decisions, and the resulting feedback.

The challenges of the VPPs are mainly addressed in the environment stage since it is essential to collect all data sources equally and to take them into account later in the understanding phase. Furthermore, the perception phase offers excellent potential for individualization. Here, the knowledge about the physical processes and the physical properties of the DERs can be of great advantage. Through the feedback loop and the streaming property, the system should map an increasingly well-founded normal behavior model overtime to make even more accurate predictions about potential intruders in the systems. The cornerstone for successful attack detection that can effectively combat APTs in the long term is in place with this architecture. Due to the pre-analysis and the basis of anomaly detection, the system will also be able to detect other attacks effectively.

The architecture does not automatically make the system more resilient but allows it to become more resilient. Because based on the messages of the detection performed here, it is possible to take measures that enable the protection of the security objectives. The proposal of this architecture in the section thus also answers the RQ stated in section I of the paper, see section I.

VI. Evaluation of the proposed architecture

The section reviews the architecture proposed in section V based on the challenges and specifics previously examined in sections II, III, and IV to determine whether it meets the requirements.

In summary, the subject domain II has produced the following challenges and features:

- SD_C1, (SD_C2), (SD_C4): VPPs are interconnected IT and OT systems of many heterogeneous systems.
- SD_C3: The EMS is the central control and monitoring unit of a VPP.
- SD_C5: A VPP has components that interact with the outside world.
- SD_C6: A VPP consists of several layers, with one component influencing the others.
- SD_C7: The OT systems involved have different goals and focuses and are therefore structured differently.
- SD_C8: The CIA triad is weighted differently for IT and OT systems. A VPP must therefore take both requirements into account. Fail safety, in particular, is becoming increasingly important.

The architecture addresses the challenges (SD_C1 - SD_C8) by adopting a defense-in-breath monitoring approach, adapting each system to its capabilities and needs. Thus, the central EMS, which can perform detailed logging, is configured accordingly. While the OT systems, which may also have to be perceived as a black box because the manufacturer does not allow them to add their logging, are monitored accordingly at the interfaces. Challenge SD_C6, due to the VPP system structure, requires a corresponding analysis, which refers to the relationships between the systems, such as graph-based recognition. Based on the different and sometimes critical runtime behavior requirements (SD_C7), the architecture follows passive logging and reporting without actively intervening in the systems. Actions are performed consciously by humans. The architecture does not directly cover SD_C8 because the countermeasures affect the CIA's protection goals. The architecture contributes by detecting the attacker as early as possible to allow the operators to comply with the protection goals.

For the second perspective, the demand landscape (section III), the challenges are summarized as follows:

- TE_C1: The VPPs are a desirable target.
- TE_C2: All attacker types and motivations are represented.
- TE_C3: The domain has implemented a few and sometimes wrong security measures for a long time.
- TE_C4: Physical attacks are possible.
- TE_C5: APTs are sophisticated and sustained, and he keeps persisting.
- TE_C6: An APT follows the "stay slow and low" pattern and is tailored to the target.
- TE_C7: An APT is a multi-stage attack.
- TE_C8: An APT has different goals, stealing information, influencing operations, or securing a good position for the future.
- TE_C9: Direct and indirect attacks on IT, as well as on the OT level, are possible.
- TE_C10: The heterogeneity between the logical components and the concrete system's versions increases the attack surface.
- TE_C11: No known attacks can be transferred to the target system to learn how to defend. The attack design is up to the attacker and is adapted to the specific target and its defenses.

The architecture addresses the challenges through defense-in-breath and depth since the defense cannot make assumptions about who is attacking and where the access is occurring. Defense-in-Breath is also a key component here, as it was previously with the challenges since this provides complete coverage of the system environment and is the only way to detect even tiny steps taken by the attacker. The defense-in-depth strategy for detection mainly addresses the challenges of APTs. As such, an attack can be detected only through multiple detection points and a combination of different methods. This is why the combination of graph and event analysis is so promising. Integrating conventional detection methods in the perception phase ensures effective detection for less sophisticated attacks or even missteps by an APT.

A look at challenge TE_C11 is necessary because of the

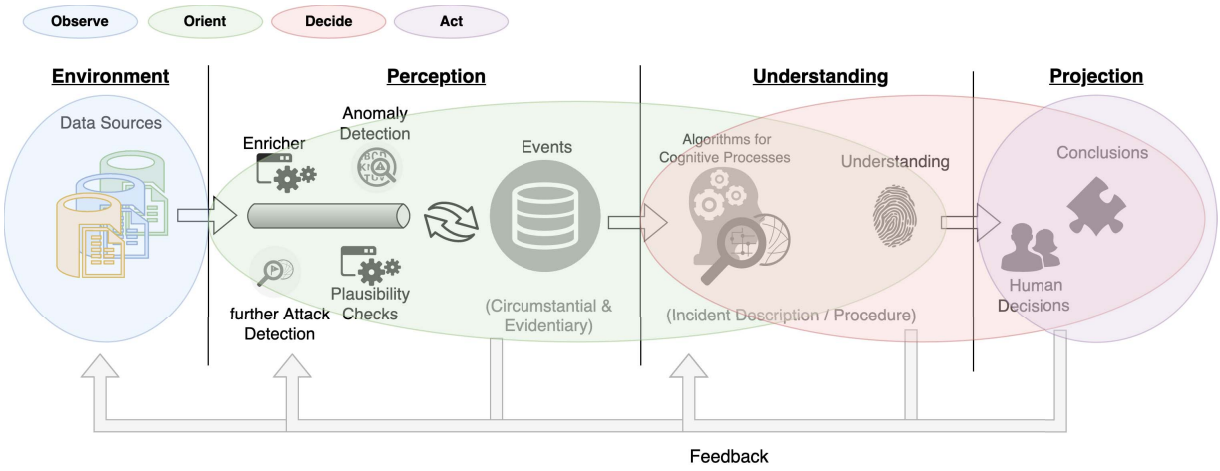

Fig. 5. Solution sketch for APT detection in VPPs with OODA loop steps.

nature of an APT. It becomes clear that the architecture favors and sometimes enforces the integration of semi- to unsupervised learning methods. This is the only way that the defense-in-depth strategy can detect unknown and novel attacks or even existing backdoor exploits.

The last segment considers the remaining C extracted from the DM perspective.

- DM_C1: An effective defense requires the most situational awareness possible.
- DM_C2: APTs cannot be detected by systems based on the detection of specific IoCs.
- DM_C3: AIDSs often provide many false positives.
- DM_C4: AIDSs usually requires a training phase.
- DM_C5: Statistical and knowledge-based methods are hardly applicable due to the lack of data on known APTs.
- DM_C6: NIDS provides data on many systems but with low levels of detail.
- DM_C7: HIDS only provide data about their system, but with a high level of detail and influence on the system's performance.
- DM_C8: The correlation of events from different sources is essential.
- DM_C9: The results of AIDS are often only indicators, not evidence of an incident.
- DM_C10: Intensive monitoring is needed even to have the ability to detect all steps.
- DM_C11: Besides algorithms, the recognition process should also consider humans.

The proposed architecture is built on the concept of SA and was designed around SA (DM_C1). In addition, the OODA loop is integrated, promoting SA and involving people in the process (DM_C11). The "Defense-in-Depth" strategy focuses on challenges DM_C2, DM_C5, and DM_C8. The focus is on detection and correlation, achievable, for example, through graph-based detection. Above this, ML methods that exploit the basis of graph structure are envisioned. DM_C3 becomes notorious for not relying exclusively on AIDS, instead using a whole process. DM_C4 is an unsolved challenge so far. Here it is assumed that appropriate training in the normal behavior for the used AIDS is possible. The defense-in-breath monitoring takes DM_C6, DM_C7, and DM_C10 into account.

This section showed that the architecture proposed here suits the specifics and challenges studied. Open challenges are the training data and maintaining the CIA security goals. However, this can be accessed in the concrete implementation of the architecture and therefore does not question the usefulness.

VII. Future Work

Detecting anomalies is an essential first step in preventing APTs. The overall goal of the presented architecture is to make a VPP more resilient. For this, early signs must result in resilient actions. The goal is not to detect a complete APT. The goal is to use early indications to fend off attacks, especially APTs. Therefore, in the further course of this research, the entirety of data sources to actions and procedures that make the systems under consideration resilient to attacks will be considered. In addition, the architecture will consider incidents that are not necessarily the result of an attack.

The overall goal of the previous steps is to fill the architecture proposed in this paper with concrete algorithms and methods. For the time being, the main focus is evaluating cognitive processes, where Unicorn, e.g., offers a possible solution. In addition, we would like to investigate the combination of neighborhood algorithms with point and time-series methods to improve the quality of the results. This step offers the most potential for addressing the challenges of APTs.

Optimizations are achieved in an extended perception phase, supplemented by intelligent enricher and simple procedures, which are, e.g., directly adapted to the domain.

After implementation, field testing and evaluation of open-source data to practically investigate a concrete instantiation of the architecture. There is still room for improvement in processing the results, as anomaly detection is more indicative. For this purpose, in the future, a user interface (UI) that links analysis results with raw events and human CTI will presumably show relevant examples from the past to classify the incident in the best possible way.

VIII. Related Work

The section of the paper presents a subset of previous detection systems that claim to meet the requirements of APTs. The work of Skopik, Friedberg, and Fiedler [61] introduces the dangers of APTs in ICT networks and offers a possible detection. They use the assertion that each component produces log data somehow, giving valuable information for establishing SA. This information is to be used in their approach agnostically with a self-learning anomaly detection method. The research team around the authors has several publications on the topic, such as [62], where they switched from agnostic log consideration to small parsers and used multi-layer detection with different correlation points.

Another tool worth mentioning is Unicorn [47], which has already been mentioned several times in this paper, as it is particularly promising. The authors present an APT detection based on provenance graphs. They use a Thinking-Like-a-Vertex graph analysis tool, GraphChi [63]. A Weisfeil-Lehman graph kernel algorithm, added subsequently by the authors, extends GraphChi [64]. The algorithm is responsible for observing the neighborhood of a node. The neighborhood's histograms are generated at runtime and then used to create fixed-length sketches. These are used in subsequent steps to cluster and thus distinguish pre-learned normal behavior from abnormal behavior because if there are changes in the neighborhood, this becomes visible in the sketches. The distance from the current state increases the learned behavior.

Another tool that provides promising results and is built on provenance graphs is HOLMES [51]. The authors rely on rules to find patterns in the provenance graph. These patterns are then guided to a higher level of abstraction to find connections between individual TTPs. Found patterns are directed to a further abstraction, including an APT's steps. Then, thresholds are used to decide when an APT has been detected. This approach can provide promising results despite a set of rules serving as a basis. Since if they are agnostic enough and rules conspicuous flag behavior in the neighborhood, this combination of rules and provenance graph could meet the requirements of APTs, although rule-based systems as used so far do not. This tool is based on expert knowledge and is, therefore, irrelevant for our research for the time being, as we are looking for an approach that does not require it. The research group behind Holmes further suggests Poirot[3] [65], ProPatrol [66], and SLEUTH [67], which are also detection systems for APTs.

In their work [68], authors Garrido, Dold, and Frank show how machine learning on knowledge graphs enables context-aware security monitoring. The previously considered provenance graphs are a particular type of knowledge graphs. The authors also work in a domain that links IT and OT. They use the linkage to represent observations from both domains together as a graph. Link prediction methods are used to detect anomalies, which is done by learning a normal behavior model beforehand. This work does not directly target APT detection but offers a high potential to address these threats.

Anubis [60] is a tool that also relies on provenance graphs, but not entirely. In addition to the provenance graph, nodes are considered individually to not only focus exclusively on an anomaly in the neighborhood but also to provide further consideration to conspicuous values. The ideas are up-and-coming; the implementation is based on a supervised learning approach, which makes the tool unattractive for our research, as it again only detects known and similar attacks.

The authors present in [69] an approach to automatically extract cyber threat intelligence (CTI) from disparate sources, put it into a graph structure, and then model relationships and determine severity. The method is based on machine-learning methods; this ensures that new entries can also be classified. The approach combines techniques from natural language processing, graph theory, and convolutional neural networks. The authors build on supervised learning, so they have high expectations for the dataset they create. This method is not suitable for the proposed architecture for two reasons: the supervised approach and also because it is based on expert knowledge, the CTI.

CONAN [58] is an APT detection system that relies on finite state automata (FSA). The automaton has state changes based on human CTI in the form of a set of rules. The authors link the FSA to a provenance graph to report alarms in a traceable way. This work is based on human knowledge and a ruleset. Therefore it does not qualify as a solution for primary detection. Still, it represents an exciting work due to its unconventional approach.

In addition to the papers, many papers could be considered as partial solutions or as solutions to specific instantiations of APTs, which were not mentioned in this section.

[3]Beware that there is another system with the same name [59], which also does APT detection.

IX. Conclusion

APTs are a particular threat to VPPs and other CIs, which must be given a high level of attention to ensure that the systems we rely on are not exploited. In summary, it can be stated that for SQ1, the combination of IT and OT and the operating processes of the VPPs place new and intense demands on an attack detection system. This must achieve the highest possible coverage of the systems and components involved. In addition, the real-time behavior of the monitored components must not be influenced, and the methods must be adapted to the landscape.

The answer to SQ2 is that the threat level for VPPs is exceptionally high due to the characteristics surveyed. APTs are challenging to detect due to their multi-layered structure, the means used, and their "low and slow" actions. Many pieces of the puzzle, the single TPP, which are not suspicious, have to be put together to form a larger picture.

Subsequently, the answer to SQ3 can be summarized. Active protection and mitigation of attacks are essential. Both methods require effective detection of grievances. Current detection methods often fail due to time or cannot detect novel attacks.

The architecture presented in the paper considers all the previously mentioned peculiarities and challenges and is ready for them. The focus on SA, in combination with the OODA loop, makes the system suitable for the defense-in-depth strategy and APT detection in the long run. And this answers the research question, "What does an APT detection system architecture for VPP look like?".

References

[1] K. E. Bakari and W. L. Kling, "Fitting distributed generation in future power markets through virtual power plants," in *2012 9th International Conference on the European Energy Market*, 2012, pp. 1–7. [Online]. Available: https://doi.org/10.1109/EEM.2012.6254692

[2] H. Saboori, M. Mohammadi, and R. Taghe, "Virtual Power Plant (VPP), Definition, Concept, Components and Types," in *2011 Asia-Pacific Power and Energy Engineering Conference*, 2011, pp. 1–4. [Online]. Available: https://doi.org/10.1109/APPEEC.2011.5749026

[3] S. K. Venkatachary, A. Alagappan, and L. J. B. Andrews, "Cybersecurity challenges in energy sector (virtual power plants) - can edge computing principles be applied to enhance security?" *Energy Informatics*, vol. 4, no. 1, Mar. 2021. [Online]. Available: https://doi.org/10.1186/s42162-021-00139-7

[4] BDEW, "Branchenspezifischer Sicherheitsstandard für Anlagen oder Systeme zur Steuerung / Bündelung elektrischer Leistung (B3S Aggregatoren)," *BSI-Gesetz*, 2021. [Online]. Available: https://www.bdew.de/media/documents/20210222

[5] K. Thakur, M. L. Ali, N. Jiang, and M. Qiu, "Impact of Cyber-Attacks on Critical Infrastructure," in *2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS)*, 2016, pp. 183–186. [Online]. Available: https://doi.org/https://doi.org/10.1109/BigDataSecurity-HPSC-IDS.2016.22

[6] S. K. Venkatachary, J. Prasad, R. Samikannu, A. Alagappan, and L. J. B. Andrews, "Cybersecurity infrastructure challenges in IoT based virtual power plants," *Journal of Statistics and Management Systems*, vol. 23, no. 2, pp. 263–276, 2020. [Online]. Available: https://doi.org/10.1080/09720510.2020.1724625

[7] T. Plėta, M. Tvaronavičienė, S. D. Casa, and K. Agafonov, "Cyber-attacks to critical energy infrastructure and management issues: overview of selected cases," *Insights into Regional Development 2(3)*, pp. 703–715, Sep. 2020. [Online]. Available: https://doi.org/10.9770/IRD.2020.2.3(7)

[8] J. Felker and M. Edwards, "ICS-CERT Year in Review 2016," *ICS-CERT*, 2016. [Online]. Available: https://www.cisa.gov/uscert/sites/default/files/Annual_Reports/Year_in_Review_FY2016_Final_S508C.pdf

[9] J. N. Alejandro Romero, "Why Europe's energy industry is vulnerable to cyber-attacks," Mar 2022. [Online]. Available: https://ecfr.eu/article/why-europes-energy-industry-is-vulnerable-to-cyber-attacks/

[10] B. Deighton, "Critical infrastructures under daily attack – ERNCIP head Georg Peter," *Horizon The EU Research & Innovation Magazine*, 2017. [Online]. Available: https://ec.europa.eu/research-and-innovation/en/horizon-magazine/critical-infrastructures-under-daily-attack-erncip-head-georg-peter

[11] C. C. for Cyber Security, "Cyber threat bulletin: The cyber threat to Canada's electricity sector, url=https://cyber.gc.ca/en/guidance/cyber-threat-bulletin-cyber-threat-canadas-electricity-sector," Nov 2020.

[12] M. Blog, "Honda and Enel impacted by cyber attack suspected to be ransomware." Jun 2020. [Online]. Available: https://blog.malwarebytes.com/threat-analysis/2020/06/honda-and-enel-impacted-by-cyber-attack-suspected-to-be-ransomware/

[13] E. J. M. Colbert and A. Kott, *Cyber-Security of SCADA and Other Industrial Control Systems*, 1st ed. Springer Cham, 2016.

[14] J.-M. Ferré, "Stuxnet, Schmitt Analysis, and the Cyber "Use-of-Force" Debate." [Online]. Available: https://apps.dtic.mil/sti/citations/ADA618715

[15] proofpoint, "TA410: The Group Behind LookBack Attacks Against U.S. Utilities Sector Returns with New Malware," Jun 2020. [Online]. Available: https://www.proofpoint.com/us/blog/threat-insight/ta410-group-behind-lookback-attacks-against-us-utilities-sector-returns-new

[16] C. . I. S. Agency, "Russian Government Cyber Activity Targeting Energy and Other Critical Infrastructure Sectors," Mar 2018. [Online]. Available: https://www.cisa.gov/uscert/ncas/alerts/TA18-074A

[17] A. Alshamrani, S. Myneni, A. Chowdhary, and D. Huang, "A Survey on Advanced Persistent Threats: Techniques, Solutions, Challenges, and Research Opportunities," *IEEE Communications Surveys Tutorials*, vol. 21, no. 2, pp. 1851–1877, 2019. [Online]. Available: https://doi.org/10.1109/COMST.2019.2891891

[18] M. Rekik, Z. Chtourou, C. Gransart, and A. Atieh, "A Cyber-Physical Threat Analysis for Microgrids," in *2018 15th International Multi-Conference on Systems, Signals & Devices (SSD)*, 2018, pp. 731–737. [Online]. Available: https://doi.org/10.1109/SSD.2018.8570411

[19] D. Kushner, "The real story of stuxnet," *IEEE Spectrum*, vol. 50, no. 3, pp. 48–53, 2013. [Online]. Available: https://doi.org/10.1109/MSPEC.2013.6471059

[20] S. Al-Rabiaah, "The "Stuxnet" Virus of 2010 As an Example of A "APT" and Its "Recent" Variances," in *2018 21st Saudi Computer Society National Computer Conference (NCC)*, 2018, pp. 1–5. [Online]. Available: https://doi.org/10.1109/NCG.2018.8593143

[21] S. Lukovic, I. Kaitovic, M. Mura, and U. Bondi, "Virtual Power Plant As a Bridge between Distributed Energy Resources and Smart Grid," in *2010 43rd Hawaii International Conference on System Sciences*, 2010, pp. 1–8. [Online]. Available: https://doi.org/10.1109/HICSS.2010.437

[22] P. M. Naina, H.-S. Rajamani, and K. S. Swarup, "Modeling and simulation of virtual power plant in energy management system applications," in *2017 7th International Conference on Power Systems (ICPS)*, 2017, pp. 392–397. [Online]. Available: https://doi.org/https://doi.org/10.1109/ICPES.2017.8387326

[23] P. Lombardi, T. Sokolnikova, Z. Styczynski, and N. Voropai, "Virtual power plant management considering energy storage systems," *IFAC Proceedings Volumes*, vol. 45, no. 21, pp. 132–137, 2012, 8th Power Plant and Power

System Control Symposium. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1474667016319577

[24] S. K. Khaitan and J. D. McCalley, "Cyber physical system approach for design of power grids: A survey," in *2013 IEEE Power & Energy Society General Meeting*, 2013, pp. 1–5. [Online]. Available: https://doi.org/https://doi.org/10.1109/PESMG.2013.6672537

[25] S. Han, M. Xie, H.-H. Chen, and Y. Ling, "Intrusion Detection in Cyber-Physical Systems: Techniques and Challenges," *IEEE Systems Journal*, vol. 8, no. 4, pp. 1052–1062, 2014. [Online]. Available: https://doi.org/10.1109/JSYST.2013.2257594

[26] W. A. Conklin, "IT vs. OT Security: A Time to Consider a Change in CIA to Include Resilienc," in *2016 49th Hawaii International Conference on System Sciences (HICSS)*, 2016, pp. 2642–2647. [Online]. Available: https://doi.org/10.1109/HICSS.2016.331

[27] G. Murino, M. Ribaudo, S. P. Romano, and A. Tacchella, "OT Cyber Security Frameworks Comparison Tool (CSFCTool)," in *Proceedings of the Italian Conference on Cybersecurity, ITASEC 2021*, A. Armando and M. Colajanni, Eds. CEUR, 2021, pp. 9–22. [Online]. Available: http://ceur-ws.org/Vol-2940/paper2.pdf

[28] J. Marron, A. Gopstein, N. Bartol, and L. Feldman, "Cybersecurity Framework Smart Grid Profile," 2019-07-09 2019.

[29] R. A. Caralli, J. H. Allen, D. W. White, L. R. Young, N. Mehravari, and P. D. Curtis, "CERT Resilience Management Model, Version 1.2," 2016. [Online]. Available: https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=508084

[30] K. Park, B. Ahn, J. Kim, D. Won, Y. Noh, J. Choi, and T. Kim, "An Advanced Persistent Threat (APT)-Style Cyberattack Testbed for Distributed Energy Resources (DER)," in *2021 IEEE Design Methodologies Conference (DMC)*, 2021, pp. 1–5. [Online]. Available: https://doi.org/10.1109/DMC51747.2021.9529953

[31] J. Lewis, "Cyberwar Thresholds and Effects," *IEEE Security & Privacy*, vol. 9, no. 5, pp. 23–29, 2011. [Online]. Available: https://doi.org/10.1109/MSP.2011.25

[32] C. K. Keerthi, M. Jabbar, and B. Seetharamulu, "Cyber Physical Systems(CPS):Security Issues, Challenges and Solutions," in *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, 2017, pp. 1–4. [Online]. Available: https://doi.org/10.1109/ICCIC.2017.8524312

[33] S. Han, M. Xie, H.-H. Chen, and Y. Ling, "Intrusion Detection in Cyber-Physical Systems: Techniques and Challenges," *IEEE Systems Journal*, vol. 8, no. 4, pp. 1052–1062, 2014. [Online]. Available: https://doi.org/10.1109/JSYST.2013.2257594

[34] A. Lemay, J. Calvet, F. Menet, and J. M. Fernandez, "Survey of Publicly Available Reports on Advanced Persistent Threat Actors," *Comput. Secur.*, vol. 72, no. C, p. 26–59, jan 2018. [Online]. Available: https://doi.org/10.1016/j.cose.2017.08.005

[35] Y. Meijaard, P.-P. Meiler, and L. Allodi, "Modelling Disruptive APTs targeting Critical Infrastructure using Military Theory," in *2021 IEEE European Symposium on Security and Privacy Workshops (EuroS PW)*, 2021, pp. 178–190. [Online]. Available: https://doi.org/10.1109/EuroSPW54576.2021.00026

[36] R. S. Ross, "Managing Information Security Risk: Organization, Mission, and Information System View," 2011. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-39.pdf

[37] C. Paulsen and R. Byers, "Glossary of key information security terms," Jul. 2019. [Online]. Available: https://doi.org/10.6028/nist.ir.7298r3

[38] M. chieh Pan and S. ting Tsai, "Weapons of Targeted Attack Modern Document Exploit Techniques," 2011. [Online]. Available: https://paper.bobylive.com/Meeting_Papers/BlackHat/USA-2012/BH_US_12_Tsai_Pan_Exploiting_Windows8_Slides.pdf

[39] P. Hu, H. Li, H. Fu, D. Cansever, and P. Mohapatra, "Dynamic defense strategy against advanced persistent threat with insiders," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, 2015, pp. 747–755. [Online]. Available: https://doi.org/10.1109/INFOCOM.2015.7218444

[40] P. Chen, L. Desmet, and C. Huygens, "A Study on Advanced Persistent Threats," in *Communications and Multimedia Security*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 63–72. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-662-44885-4_5

[41] P. Bhatt, E. T. Yano, and P. Gustavsson, "Towards a Framework to Detect Multi-stage Advanced Persistent Threats Attacks," in *2014 IEEE 8th International Symposium on Service Oriented System Engineering*, 2014, pp. 390–395. [Online]. Available: https://doi.org/10.1109/SOSE.2014.53

[42] H. P. Sanghvi and M. S. Dahiya, "Cyber Reconnaissance: An Alarm before Cyber Attack," *International Journal of Computer Applications*, vol. 63, no. 6, pp. 36–38, 2013. [Online]. Available: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.278.5965&rep=rep1&type=pdf

[43] W. Mazurczyk and L. Caviglione, "Cyber Reconnaissance Techniques," *Commun. ACM*, vol. 64, no. 3, p. 86–95, feb 2021. [Online]. Available: https://doi.org/10.1145/3418293

[44] M. Lehto, "APT Cyber-attack Modelling: Building a General Model," *International Conference on Cyber Warfare and Security*, vol. 17, no. 1, pp. 121–129, Mar. 2022. [Online]. Available: https://doi.org/10.34190/iccws.17.1.36

[45] J. Straub, "Modeling Attack, Defense and Threat Trees and the Cyber Kill Chain, ATT&CK and STRIDE Frameworks as Blackboard Architecture Networks," in *2020 IEEE International Conference on Smart Cloud (SmartCloud)*, 2020, pp. 148–153. [Online]. Available: https://doi.org/10.1109/SmartCloud49737.2020.00035

[46] D. van der Velde, M. Henze, P. Kathmann, E. Wassermann, M. Andres, D. Bracht, R. Ernst, G. Hallak, B. Klaer, P. Linnartz, B. Meyer, S. Ofner, T. Pletzer, and R. Sethmann, "Methods for Actors in the Electric Power System to Prevent, Detect and React to ICT Attacks and Failures," in *2020 6th IEEE International Energy Conference (ENERGYCon)*, 2020, pp. 17–22. [Online]. Available: https://doi.org/10.1109/ENERGYCon48941.2020.9236523

[47] X. Han, T. Pasquier, A. Bates, J. Mickens, and M. Seltzer, "UNICORN: Runtime Provenance-Based Detector for Advanced Persistent Threats." Proceedings 2020 Network and Distributed System Security Symposium, 2020.

[48] Nist and E. Aroms, *NIST Special Publication 800-39 Managing Information Security Risk*. Scotts Valley, CA: CreateSpace, 2012. [Online]. Available: https://dl.acm.org/doi/book/10.5555/2331278

[49] Q. Zou, X. Sun, P. Liu, and A. Singhal, "An Approach for Detection of Advanced Persistent Threat Attacks," *Computer*, vol. 53, no. 12, pp. 92–96, 2020. [Online]. Available: https://doi.org/10.1109/MC.2020.3021548

[50] A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intrusion detection systems: techniques, datasets and challenges," *Cybersecurity 2 20*, 2019. [Online]. Available: https://doi.org/10.1186/s42400-019-0038-7

[51] S. M. Milajerdi, R. Gjomemo, B. Eshete, R. Sekar, and V. Venkatakrishnan, "HOLMES: Real-Time APT Detection through Correlation of Suspicious Information Flows," in *2019 IEEE Symposium on Security and Privacy (SP)*, 2019, pp. 1137–1152. [Online]. Available: https://doi.org/10.1109/SP.2019.00026

[52] M. Landauer, F. Skopik, M. Wurzenberger, W. Hotwagner, and A. Rauber, "A Framework for Cyber Threat Intelligence Extraction from Raw Log Data," in *2019 IEEE International Conference on Big Data (Big Data)*, 2019, pp. 3200–3209. [Online]. Available: https://doi.org/10.1109/BigData47090.2019.9006328

[53] D. Binaco, "A framework for cyber threat hunting part 1: The pyramid of pain," 2015. [Online]. Available: http://blog.sqrrl.com/a-framework-for-threat-hunting-part-1-the-pyramid-of-pain

[54] S. Bromander, A. Jøsang, and M. Eian, "Semantic Cyberthreat Modelling," *STIDS 2016 Proceedings*, 2016. [Online]. Available: http://ceur-ws.org/Vol-1788/STIDS_2016_A03_Bromander_etal.pdf

[55] R. Stillions, "The DML Model," 2014. [Online]. Available: http://ryanstillions.blogspot.com/2014/04/the-dml-model21.html

[56] I. Paik, "Situation awareness based on big data analysis," in *2016 International Conference on Machine Learning and Cybernetics (ICMLC)*, vol. 2, 2016, pp. 911–916. [Online]. Available: https://doi.org/10.1109/ICMLC.2016.7873008

[57] M. Révay and M. Líška, "OODA loop in command & control systems," in *2017 Communication and Information Technologies (KIT)*, 2017, pp. 1–4. [Online]. Available: https://doi.org/10.23919/KIT.2017.8109463

[58] C. Xiong, T. Zhu, W. Dong, L. Ruan, R. Yang, Y. Cheng, Y. Chen, S. Cheng, and X. Chen, "Conan: A Practical Real-Time APT Detection System With High Accuracy and Efficiency," *IEEE Transactions on Dependable and Secure Computing*, vol. 19, no. 1, pp. 551–565, 2022. [Online]. Available: https://doi.org/10.1109/TDSC.2020.2971484

[59] J. Yang, Q. Zhang, X. Jiang, S. Chen, and F. Yang, "Poirot: Causal Correlation Aided Semantic Analysis for Advanced Persistent Threat Detection," *IEEE Transactions on Dependable and Secure Computing*, pp. 1–1, 2021. [Online]. Available: https://doi.org/10.1109/TDSC.2021.3101649

[60] M. M. Anjum, S. Iqbal, and B. Hamelin, "ANUBIS: A Provenance Graph-Based Framework for Advanced Persistent Threat Detection," in *Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing*, ser. SAC '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 1684–1693. [Online]. Available: https://doi.org/10.1145/3477314.3507097

[61] F. Skopik, I. Friedberg, and R. Fiedler, "Dealing with advanced persistent threats in smart grid ICT networks," in *ISGT 2014*, 2014, pp. 1–5. [Online]. Available: https://doi.org/10.1109/ISGT.2014.6816388

[62] M. Wurzenberger., F. Skopik., G. Settanni., and R. Fiedler., "AECID: A Self-learning Anomaly Detection Approach based on Light-weight Log Parser Models," pp. 386–397, 2018. [Online]. Available: https://doi.org/10.5220/0006643003860397

[63] A. Kyrola, G. Blelloch, and C. Guestrin, "GraphChi: Large-Scale Graph Computation on Just a PC," in *Proceedings of the 10th USENIX Conference on Operating Systems Design and Implementation*, ser. OSDI'12. USA: USENIX Association, 2012, p. 31–46. [Online]. Available: https://www.usenix.org/conference/osdi12/technical-sessions/presentation/kyrola

[64] N. Shervashidze, P. Schweitzer, E. J. van Leeuwen, K. Mehlhorn, and K. M. Borgwardt, "Weisfeiler-Lehman Graph Kernels," *J. Mach. Learn. Res.*, vol. 12, no. null, p. 2539–2561, nov 2011. [Online]. Available: https://www.jmlr.org/papers/volume12/shervashidze11a/shervashidze11a.pdf

[65] S. M. Milajerdi, B. Eshete, R. Gjomemo, and V. Venkatakrishnan, "POIROT: Aligning Attack Behavior with Kernel Audit Records for Cyber Threat Hunting," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1795–1812. [Online]. Available: https://doi.org/10.1145/3319535.3363217

[66] S. Milajerdi, B. Eshete, R. Gjomemo, and V. Venkatakrishnan, "ProPatrol: Attack Investigation via Extracted High-Level Tasks," 10 2018. [Online]. Available: https://doi.org/10.48550/arXiv.1810.05711

[67] M. N. Hossain, S. M. Milajerdi, J. Wang, B. Eshete, R. Gjomemo, R. Sekar, S. Stoller, and V. Venkatakrishnan, "SLEUTH: Real-time Attack Scenario Reconstruction from COTS Audit Data," in *26th USENIX Security Symposium (USENIX Security 17)*. Vancouver, BC: USENIX Association, Aug. 2017, pp. 487–504. [Online]. Available: https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/hossain

[68] J. S. Garrido, D. Dold, and J. Frank, "Machine learning on knowledge graphs for context-aware security monitoring," *CoRR*, 2021. [Online]. Available: https://doi.org/10.48550/arXiv.2105.08741

[69] J. Zhao, Q. Yan, X. Liu, B. Li, and G. Zuo, "Cyber Threat Intelligence Modeling Based on Heterogeneous Graph Convolutional Network," in *23rd International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2020)*. San Sebastian: USENIX Association, Oct. 2020, pp. 241–256. [Online]. Available: https://www.usenix.org/conference/raid2020/presentation/zhao