Annual Reviews in Control 47 (2019) 394–411

Contents lists available at ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Review article

A systems and control perspective of CPS security

Seyed Mehran Dibaji[a,*], Mohammad Pirani[b], David Bezalel Flamholz[a], Anuradha M. Annaswamy[a], Karl Henrik Johansson[b], Aranya Chakrabortty[c]

[a] Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA
[b] Department of Automatic Control, KTH Royal Institute of Technology, Sweden
[c] Department of Electrical Engineering, North Carolina State University, Raleigh, NC, USA

ARTICLE INFO

Article history:
Received 16 January 2019
Revised 23 April 2019
Accepted 24 April 2019
Available online 23 May 2019

Keywords:
Cyber-physical systems
Cyber-security
Cyber-physical security
Secure control
Resilient control

ABSTRACT

The comprehensive integration of instrumentation, communication, and control into physical systems has led to the study of Cyber-Physical Systems (CPSs), a field that has recently garnered increased attention. A key concern that is ubiquitous in CPS is a need to ensure security in the face of cyber attacks. In this paper, we carry out a survey of systems and control methods that have been proposed for the security of CPS. We classify these methods into three categories based on the type of defense proposed against the cyberattacks: prevention, resilience, and detection & isolation. A unified threat assessment metric is proposed in order to evaluate how CPS security is achieved in each of these three cases. Also surveyed are the risk assessment tools and the effect of network topology on CPS security. Furthermore, an emphasis has been placed on power and transportation applications in the overall survey.

Contents

[*] Corresponding author.
E-mail addresses: dibaji@mit.edu (S.M. Dibaji), mpirani@uwaterloo.ca (M. Pirani), flamholz@mit.edu (D.B. Flamholz), aanna@mit.edu (A.M. Annaswamy), kallej@kth.se (K.H. Johansson), achakra2@ncsu.edu (A. Chakrabortty).

https://doi.org/10.1016/j.arcontrol.2019.04.011

1. Introduction

Motivated by concerns about sustainability, efficiency, and resiliency, several sectors including energy, transportation, water, and healthcare systems have witnessed significant advances in instrumentation, monitoring, and automation over the past decade. The resulting integration of information, communication, and computation with physically engineered systems demands a detailed investigation into the analysis and synthesis of Cyber-Physical Systems (CPS) as a means to realize the desired performance metrics of efficiency, sustainability, and safety. The extensive and intricate presence of cyber components also introduces concerns over unwanted access to these systems. The available communication technologies, referred to as SCADA (Supervisory Control and Data Acquisition), are witnessing significant advances, triggering a shift from protected, closed, and wired networks to open and wireless networks, which, as a side effect, are more vulnerable to outside interference. This, in turn, has led to a recent systematic investigation of security of CPS, various attack models, tools for analysis of CPS security, and most importantly, methods for ensuring resilience against cyber attacks. This paper surveys this emerging area and offers a systems and control-theoretic perspective to provide a snapshot of the current state of research in the field. For the purposes of this paper, we denote the term CPS security to include both security, which sometimes is used as a system property that corresponds to defense against attacks, and resiliency, a system property that corresponds to survival and recovery after occurrence of an attack.

The notion of security against unwanted intrusions and attacks can be traced back to the times of Caesar (Tranquillus, 1957) and early warfare strategies. A technological intersection with this topic, however, has its origins in the proliferation of computers in the commercial sector. Grouped under the rubric of InfoSec, information security breaches were recognized to be central to the satisfactory performance of a system. In particular, three security breaches were often considered to be important for the protection of information (Cherdantseva & Hilton, 2012; 2013; Saltzer & Schroeder, 1975): Confidentiality, Integrity, and Availability which denote as an unauthorized information release, an unauthorized information modification, and an unauthorized denial of use of the information, respectively[1].

Given the central role that information plays in a feedback control system, the approaches to achieving CPS security can be grouped using the same taxonomy (see Fig. 1). A confidentiality breach can be viewed as the monitoring of information that is used to control the system, integrity breach as the corruption of the sensor data sent to the network for processing, and availability breach as either blocking or delaying of the information between the computational block and the actuation node in a system (Cárdenas et al., 2011; Cárdenas, Amin, & Sastry, 2008; Cardenas, Amin, & Sastry, 2008; Sandberg, Johansson, & Amin, 2015).

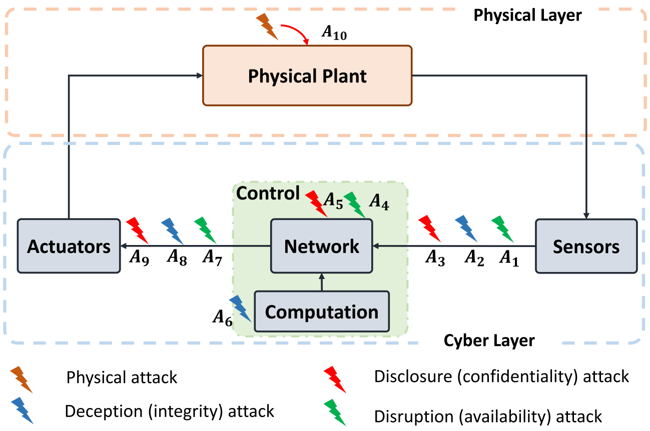

Fig. 1. A schematic of various attacks that can occur in a CPS grouped under the three categories, Disclosure attacks, Deception attacks, and Disruption attacks.

If protection against the security breaches above can be viewed from a defender's perspective, an attacker's perspective can be considered as well to address CPS security. Broadly speaking, cyber attacks have been grouped under three headings; disclosure attacks, deception attacks, and disruption attacks (Bishop, 2005; Teixeira, Sou, Sandberg, & Johansson, 2015) denoted as DDD attacks in what follows (Fig. 1). Disclosure attacks refer to any intrusions that include eavesdropping (Nozari, Tallapragada, & Cortés, 2017); deception attack corresponds to the corruption of signals (such as a spoofing attack (Jafarnia-Jahromi, Broumandan, Nielsen, & Lachapelle, 2012) or a false-data injection attack (Pasqualetti, Dörfler, & Bullo, 2013)), and a disruption attack corresponds to another active intrusion where the signal may either be blocked or delayed (e.g., denial of service (De Persis & Tesi, 2015)). These three attacks are not mutually exclusive–almost all deception attacks can be disruptive as well; disruption attacks need not necessarily coincide with a deception attack to achieve a more active action such as blocking or delaying. It is clear that there is a direct mapping between these three attack-models and the three security goals of confidentiality, integrity, and availability (Fig. 1). The disclosure attack is analogous to the confidentiality breach, the deception attack to the integrity breach, and the disruption attack to the availability breach. Both the CIA goals and the DDD attacks have been extensively analyzed in the literature for analysis and synthesis of CPS security over the past few years (Amin, Cárdenas, & Sastry, 2009; Bishop, 2005).

In a well-designed control system where performance goals of accuracy, speed, and robustness are met, allowing cyber attacks to have an impact, let alone a significant one, seems like an impossibility. To the contrary, the number of attacks, as well as their impact on the underlying infrastructure, has been quite compelling. We summarize some of the major attacks on control systems in power and transportation infrastructures in the following section. Each of the major attacks is classified using the security breaches and attack models described above. The specific set of components compromised in the underlying feedback control loops is indicated in Fig. 1.

[1] In the literature (Zeldovich (Fall 2014)), security goals are also defined in the same manner but with a positive voice: confidentiality is to maintain the secrecy of the important data, integrity is to guarantee the fidelity of the data, and availability is to ensure the accessibility of the data at the right time.

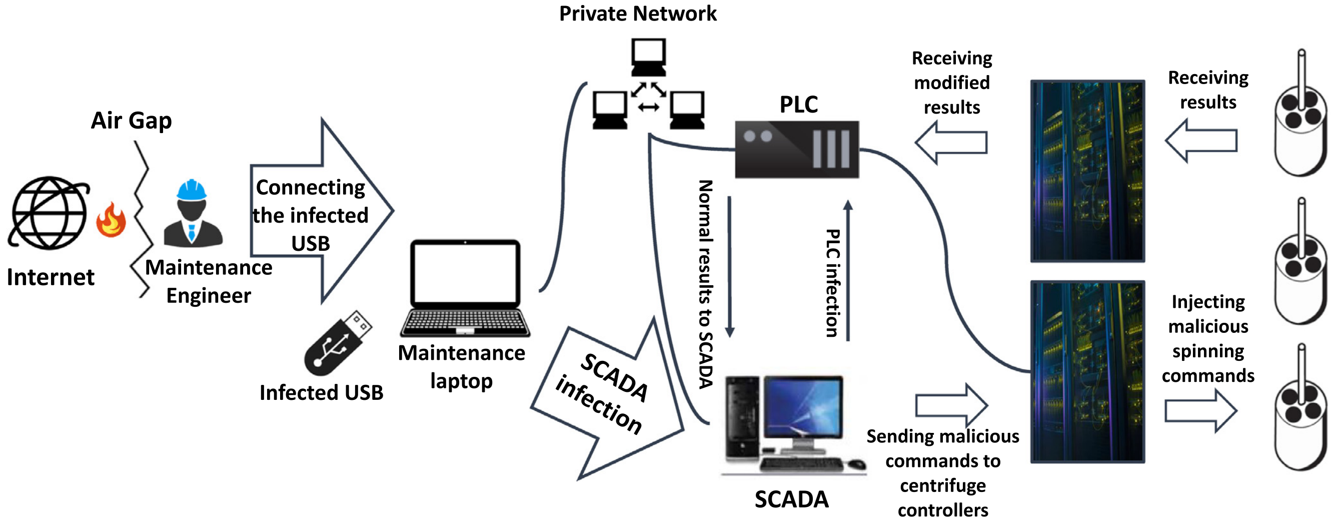

Fig. 2. A schematic of the Stuxnet attack that occurred in 2011.

1.1. Examples of CPS cyber-attacks

In this subsection, we name a few of the most consequential attack scenarios that have occurred in real cyber-physcal systems.

1.1.1. Stuxnet

Stuxnet was a cyber-physical attack on an Iranian uranium enrichment plant in late 2009. In targeting a commercially available Programmable Logic Controller, operating under a narrow set of conditions, the attackers were able to ensure the attack reached its intended recipient with limited fallout. They inserted a malware which would lie dormant in the system and go undetected (Falliere, Murchu, & Chien, 2018; Zero Days, 2016). With such a stealthy presence, observing critical and confidential system data, the attacker observed key outputs of the system under stable conditions, and replayed those measurements to other monitoring sites of the network. Simultaneously, malicious actuation signals were injected into other critical actuation sites, resulting in a significant damage to a number of centrifuges (Falliere et al., 2018). In general, many cyber-attacks can remain undetected after insertion, for a significantly long period up to a year (Mo, Weerakkody, & Sinopoli, 2015; Zero Days, 2016). Fig. 2 illustrates Stuxnet in a schematic form. One can view Stuxnet as a combination of deception and disclosure attacks.

1.1.2. RQ-170

In 2011, US operators lost control of an RQ-170 unmanned aerial vehicle (UAV) which subsequently landed in Iran. One speculation as to what caused this to occur is that Iranian forces jammed GPS communications followed by a spoof of GPS signals, thereby deceiving the drone into landing in the desired location (Hartmann & Steup, 2013). In addition to this attack on a UAV, a number of studies have been carried out to show the potential threat of GPS spoofing on vehicles (Harding et al., 2014). The RQ-170 attack can be viewed as a disruption attack followed by a deception attack.

1.1.3. Ukraine attack

Traditional practice in power grids is to institute safeguards against physical faults (Watts, 2003) using protective devices. A singular departure from such occurrences happened in the Ukraine attack. This consisted of a series of attacks on Ukrainian power distribution networks causing outages as well as lasting damage in 2015. The first was introduced via phishing emails containing the Black Energy malware. Once it infiltrated the system, it enabled the attacker to steal critical data and study the system environment. This, in turn, enabled access to a more critical control level and allowed the spoofing of control commands (Case, 2016). That is, first there was a confidentiality breach, followed by an integrity breach. Finally, by overwriting the firmware in a few substations, the attacker was able to ensure remote inoperability of breakers, leading to an availability breach. In 2016, yet another attack was launched on a transmission station using the Crash Override malware. This malware could communicate directly with grid control software and its modular design enabled it to be modified to work for US or European grid protocols as well (Greenberg, 2017).

1.1.4. Maroochy attack

In 2000, the Maroochy water services in Queensland, Australia, were attacked by a disgruntled employee. Motivated by revenge, he accomplished the attack by infiltrating the SCADA network of water services and altered the control signals. The attacker took control of 150 sewage pumping stations resulting in the evacuation of one million liters of untreated sewage, over a three-month period, into stormwater drains and on to local waterways (Slay & Miller, 2007). This is clearly a deception attack/integrity breach on actuators.

1.1.5. Jeep hack

Car hacking shows a large level of vulnerability that modern automotive systems seem to possess against adversarial actions. One of the examples was an (under control) attack on a Jeep which was driving in 70 mph on a highway in St. Louis, USA, where the car was hijacked remotely by attackers to show how various Electronic Control Units, from wiper to brake and engine systems, can be manipulated remotely through the cellular connection inside the vehicle (Greenberg, 2018). Although this attack was set to be under control, it is claimed that remote car hacking can have life-threatening consequences for passenger vehicles in the future (Koscher et al., 2010).

1.1.6. Other attacks

The attacks listed above are by no means comprehensive. They are meant to be an overview of some of the major cyber-attacks that have had a noticeable impact on power and transportation infrastructures. The earliest cyber attack on critical infrastructure is reported to have occurred in 1982 when the sale of intentionally damaged control software to the Soviet Union resulted in an explosion in Siberia (Onyeji, Bazilian, & Bronk, 2014). Over the past five years, there have been several other cyber attacks on ground transportation infrastructures (ENISA cybersecurity report, 2016), the service industry, and the manufacturing industry to name a few (for a list of cyber-attacks refer to Data Breach Investigations Report (2009, 2015); Hackmageddon (2018)). We have excluded physical attacks like the Metcalf sniper attack (Sniper attack CNN report, 2015) that have occurred on a PG&E transmission substation in California leading to a large financial loss and pilot intended crash of Airbus (Airbus A320-211 report, 2015) Based on Rus et al. (2018), two thirds of attacks have been initiated by phishing emails. A majority (70-80%) of attacks are abetted by insiders. 67% of the cyber threats are enabled by victim errors, 64% are directly introduced by hackers, and 38% by malware.

1.2. Research opportunities in CPS security

The FY 2019 US Presidents Budget includes $15 billion of budget authority for cyber-security-related activities, a $583.4 million (4.1 percent) increase above the FY 2018 Estimate (Cybersecurity funding, 2019) which indicates the level of attention being paid to this topic. As the problem of CPS security is of huge interest to the engineering community, it is not surprising that there is a large number of research investigations over the past decade. Earlier works such as Cárdenas et al. (2011, 2008); Cardenas et al. (2008) and Sandberg, Teixeira, and Johansson (2010) brought to attention the fact that the topic of cyber-attacks is not of interest just to the cybersecurity community, but out of significantly broader interest. These works also demonstrated that component-wise solutions may not suffice, and instead, these threats must be analyzed from a comprehensive system and infrastructure perspective. Also as mentioned above, DDD attacks have been discussed at length in Teixeira, Shames, Sandberg, and Johansson (2015) and Cardenas et al. (2008).

Issues of CPS security arise in a range of applications. On a daily basis, there are reports of cyberattacks in almost every sector that includes a cybercomponent. To give the readers a better sense of the impact of cyberattacks and the general problem of CPS security, in what follows, we expand on the impact in the context of power systems (Ashok, Govindarasu, and Wang, 2017; Bobba et al., 2012; Gusrialdi and Qu, 2019; Huang, Satchidanandan, Kumar and Xie, 2018; Humayed, Lin, Li, and Luo, 2017; Li, Shahidehpour, and Aminifar, 2017; Liang, Zhao, Luo, Weller, and Dong, 2017; Liao and Chakrabortty, 2018; Liu and Li, 2017; Milani, Khan, Chakrabortty, and Husain, 2018; Nordell, 2012; Onyeji et al., 2014; Sandberg et al., 2010; Sanjab, Saad, Guvenc, Sarwat, and Biswas, 2016; Sridhar, Hahn, Govindarasu et al., 2012; Stoustrup, Annaswamy, Chakrabortty, and Qu (ed.), 2019; Wang and Lu, 2013; Xie, Mo, and Sinopoli, 2011; Yan, Qian, Sharif, and Tipper, 2013). In particular, see Nordell (2012) for definitions of security and Gusrialdi and Qu (2019) for a chapter on cybersecurity in smart grid control with comprehensive discussion on attacks by insiders and outsiders and countermeasures in a power system. The presence of a large number of subsystems in power systems implies that the impact of attacks vary significantly depending on where they occur. Broadly speaking, this impact can be summarized over the following three broad headings:

1. Transmission level: Attacks may happen in AGC control loops (generator governor control) (Huang et al., 2018), PSS, FACTS controller, and wide-area controllers (Ashok et al., 2017; Bobba et al., 2012; Liao & Chakrabortty, 2018). Each of these cases may involve denial-or-service type attacks, hardware failures, control software failure, replay attacks, and data tampering attacks. Wide-area control is especially susceptible to attacks as substantive long-distance sensitive communication is required for WAC, opening up many vulnerable points for attackers to intrude through. For example, if the computation of the wide-area control signals is happening in a cloud, then attackers may cause DoS, data tampering and other such attacks to either degrade the closed-loop CPS performance or completely destabilize the system. Moreover, GPS receivers of PMU measurements are prone to spoofing attacks (Jafarnia-Jahromi et al., 2012). Direct hardware attacks on the bare metal of the shared virtual computers used in the cloud for performing these computations, or software attacks on their data storage units and hypervisors, are also highly possible.
2. Distribution level: Attacks can happen in islanded microgrids, grid-connected microgrids, or networked microgrids (Li, Shahidehpour, et al., 2017). Similarly, in a distribution-level power grid, an attacker may change the current and voltage setpoints of the power converters that connect the renewable resources to the grid to wrong values such that the power flow equations no longer have any feasible solution, forcing the grid to enter into an unsafe zone (Milani et al., 2018). Another important point is that the majority of communication protocols used for microgrid operation and control are executed using wireless communication, where security may be a serious concern. If this communication is hacked, and, as a result, messages do not reach the microgrid controllers from the supervisory management layer, then severe issues of frequency stability and voltage stability can arise.
3. Market level: False data injection in electricity markets has been investigated in Xie et al. (2011), where a convex optimization problem is solved by the attacker to find which nodal price of Ex-Post market must be manipulated to maximize the financial profit of the attacker.

It has to be noted that in power systems and more generally in complex cyber-physical systems, fault-tolerance is the basic requirement to make sure a blackout would not occur even in the presence of unintentional faults. One example of power grids lacking sufficient fault-tolerance is the Venezuela power grid that encountered frequent balckouts in March 2019 (Venezuela et al., 2019).

Similar to power systems, transportation systems experience a wide range of impact depending on the specific subsystem that has been targeted. Unmanned aerial vehicles and their vulnerabilities are addressed in Rani, Modares, Sriram, Mikulski, and Lewis (2016). General transportation systems have been examined in Javed and Hamida (2017), Hoppe, Kiltz, and Dittmann (2008), Fagnant and Kockelman (2015), Sucasas, Mantas, Saghezchi, Radwan, and Rodriguez (2016), Xue, Wang, and Roy (2014), Sherif, Rabieh, Mahmoud, and Liang (2017), Alam, Ferreira, and Fonseca (2016), Parkinson, Ward, Wilson, and Miller (2017), Petit and Shladover (2015), Woo, Jo, and Lee (2015), Siegel, Erb, and Sarma (2018), Ahmed and Gharavi (2018), Amoozadeh et al. (2015), and industrial systems in Bradbury (2012), Fovino, Carcano, Masera, and Trombetta (2009), Huitsing, Chandia, Papa, and Shenoi (2008), Ding, Han, Xiang, Ge, and Zhang (2018). As evidenced by the Jeep-Hack and Stuxnet examples, the impact of cyber-attacks can be significant, as they represent safety-critical infrastructures.

In addition to the above applications, CPS security has been investigated in Cho and Woo (2017) for protection methods in nuclear power plants, motivated by the cyber attack in 2014. In Wang, Wang, Shen, Alsaadi, and Hayat (2016), a review on deception and disruption attacks in CPSs has been performed. The importance of security in SCADA has been discussed in Bradbury (2012), Fovino et al. (2009) and in Modbus control systems in Huitsing et al. (2008). A survey on information and communication-based security aspects of industrial control systems is done by McLaughlin et al. (2016). The repeated occurrence of the term *security* in the recent control systems survey (Lamnabhi-Lagarrigue et al., 2017) is another indicator of the importance of this topic.

1.3. Contributions of this paper

The purpose of this paper is to present an overview of the research activities in the area of CPS security in two critical infrastructures, *power* and *transportation*. We provide this overview with two objectives in mind; the first is to provide a broad system and control perspective within which most of the research contributions to-date can be viewed. The second objective is to map these contributions to a CIA-taxonomy of security breaches and DDD-taxonomy of attack models. We accomplish both of these objectives in Section 2, where all systems and control defense mechanisms are grouped under three headings of prevention, resilience, and detection & isolation, based on the underlying concept employed. In this section, a metric that quantifies the CPS security is also proposed. We discuss threat assessment tools and required topological conditions for underlying networks in Section 3. A summary and suggestions for future research are included in Section 4.

2. Systems and control methods for CPS security

With the overall importance of CPS security mentioned in the introduction, we now focus on the details of the intersection of this topic with closed-loop control systems. Given that control systems consist of key components such as sensors which measure key variables of interest and actuators that synthesize control inputs that help the system perform tasks of regulation and tracking, it is important that the CPS security focuses on sensors and actuators. In addition to these components, another important component of a control system is a communication network often used to relay crucial information to relevant places. Therefore CPS security needs to include its focus not only on sensors and actuators but also the underlying communication network. It is essential to model the adversary and its available resources. One such conceptual model is illustrated in Fig. 3. It should be noted that attacks on CPSs often are confused with the faults that may occur randomly and are oblivious of the system model. Attacks, unlike faults, through eavesdropping, publicly available knowledge, operator faults, and so on, may have access to the system model and make use of this knowledge to design smarter and more effective attacks (Teixeira, Sou, et al., 2015). This is one of the reasons that a system and control level of defense is required against such attacks. However, in order to develop a systematic set of methodologies that provides CPS security for control systems, one needs to begin with an analysis of the different type of attacks that are possible. Attack models that characterize the capabilities of the attacker, such as their computational power, the type of access they may have, the data they collect, and their collaborative capabilities are very much needed. One such model is developed in Teixeira, et al. (2015) which is based on the available resources to an attacker. These models need to be designed using questions such as the following: What are the points and signals that attackers have access to? What can they do precisely on the signals? What are their limitations? These must be stated clearly in order to understand the logic behind the associated defense mechanisms, to understand the level of conservatism, to compare with other security mechanisms, and to improve the defense mechanisms.[2] The DDD attack models mentioned in the introduction provide a starting point for such designs and form our primary focus in this paper.

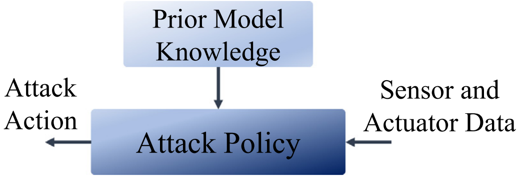

Fig. 3. A schematic of how attacks impact a CPS. It should be noted that an attack action differs from random faults in the sense that attacks may have access to the system model and suitably leverage to exacerbate the impact on the CPS.

In Fig. 1, we illustrate a feedback control system and various vulnerable points that can be attacked, indicating the corresponding attacks as $A_i, i = 1, \dots, 9,$ which are often referred to as attack surfaces (Manadhata & Wing, 2011). In what follows, we classify the existing literature on attacks into one of the three DDD categories, and proceed within each category to subclassify them on the basis of the specific point in the CPS system which is attacked. It should however be noted that this mapping is not necessarily unique. A typical attack may include both features of deception and disruption, and therefore could be grouped under either of these categories. For instance, the paper Gil, Kumar, Mazumder, Katabi, and Rus (2017) considers a masquerading attack in which a malicious node spoofs a large number of legitimate nodes, which could be viewed as either a deception or a disruption attack.

2.1. Attack models

As shown in Fig. 1, a typical cyber-physical system consists of inputs $u(k)$ at time step k and outputs $y[k]$ that are measured by sensors, communicated through a network, and with suitable computations, the control input is communicated and delivered to the physical system through actuators. The attacks $A_i, i = 1, 2, 3$ are on the sensors, A_5, A_6 are on the networks, A_6 is on the computational layer, $A_i, i = 7, 8, 9$ are on the actuators. In what follows, we group these attacks into the DDD categories, and wherever possible, we present the underlying attack model.

Disclosure attacks: Disclosure attacks aim to find access to informative signals or obtain some conclusive information about them. A successful disclosure attack may directly use or sell the obtained data or use them in order to extract other information about the system. The latter is called *inference attacks* or sometimes *known plaintext attacks* (Yuan & Mo, 2015) which are to infer the private information of a system, such as its transfer function, using the access to some potentially legitimate parts of the system (du Pin Calmon & Fawaz, 2012) such as sensory data and control inputs. Disclosed, or indirectly inferred data, can also be used to design smarter attacks in the future. One of the reasons that disclosure attacks are, despite their simple definitions, more vital in security is that the detection of disclosure attacks usually take a long time, i.e. the attacks are in the so-called *zero days* mode. It has been proved that in some systems, the attacks that use the system's information, potentially gained via an initial disclosure attack, can destabilize the closed-loop system. Disclosure attacks may take place on either sensor measurements, control computations, or actuation signals, which are indicated in Fig. 1, respectively, by A_3, A_5, and A_9.

Deception attacks: Deception attacks or false data injection (FDI) are accomplished when the signals are somehow different from

[2] There is a wide spectrum of assumptions that can be made on attacks. Based on *Shannon's Maxim*, the enemy knows the system. The Shannon's paradigm is in contrast to *Security by Obscurity* in which the security is guaranteed by assumptions on the secrecy of the system's data (Shannon, 1949). The best security solutions are those that with the assumptions on which data are shared or under a direct access of the public (or non-trusted insiders), the attacks are defined thoroughly. This is done in computer security via *Access Control* tables (Kern, Kesavan, & Daswani, 2007).

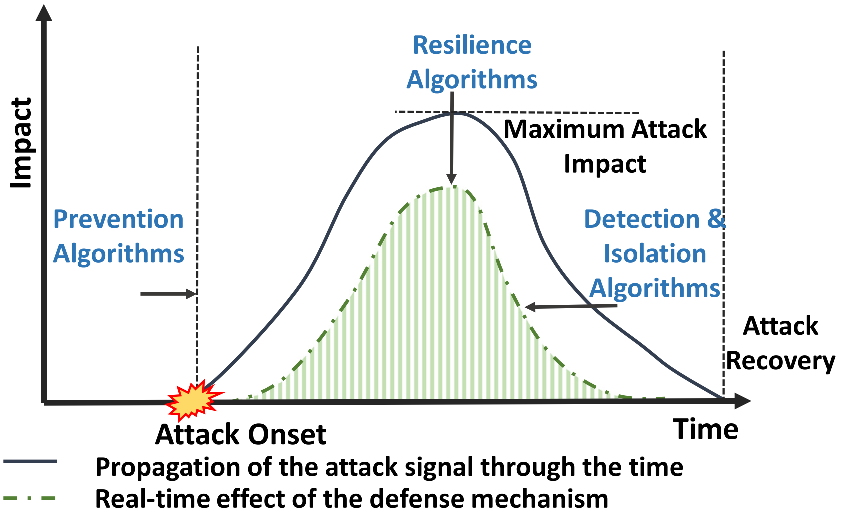

Fig. 4. Threat assessment of a CPS system and its reduction due to three defense mechanisms, based on prevention, resilience, and detection & isolation. The variable I denotes a quantity of interest that signifies the system vulnerability.

their true value. They can occur in three distinct locations in the closed-loop system (see Fig. 1):

(i) Sensor attacks, which change the operating conditions ruining the fidelity of the measurements (A_2 in Fig. 1), i.e.,

$$y^a[k] = y[k] + a_y[k], \tag{1}$$

where $y^a[k]$ is the corrupted measurement vector and $a_y[k]$ is the attack verctor, non-zero for some of the measurements (Cardenas et al., 2008).

(ii) Actuation attacks, which deviate the control signals from the values they have to be (A_8 in Fig. 1), and (iii) Computational attacks, in some CPSs, which alter the control law (A_6 in Fig. 1), i.e., in both cases the attacked control input $u^a[k]$ is given by

$$u^a[k] = u[k] + a_u[k],$$

where $a_u[k]$ is a non-zero value for some of the control inputs (Cardenas et al., 2008).

Deception attacks are the strongest attacks in terms of the level of damages they may create. For example, it is easy to imagine how a deception attack can quickly destabilize the closed-loop system. A good example of such attacks is introduced in Brown and Demarco (2018) for power systems.

Disruption attacks: Any intentional tampering of information comes under the category of disruption attacks, sometimes denoted as *denial of service* (DoS) attacks or *jamming* attacks. DoS attacks can be on the sensor data (A_1 in Fig. 1), in the underlying network for (A_4) or on the actuation signals (A_7) (Cardenas et al., 2008), all of which can be classified into the following attack models:

$$y^a[k] = \begin{cases} 0 & \text{if } A_1 \text{ occurs,} \\ y[k] & \text{otherwise.} \end{cases} \tag{2}$$

$$u^a[k] = \begin{cases} 0 & \text{if } A_4 \text{ or } A_7 \text{ occurs,} \\ u[k] & \text{otherwise.} \end{cases} \tag{3}$$

Often such DoS attacks are countered by using a Zero-Order-Hold approach. Gusrialdi and Qu (2019) provides a more detailed exposition of these models and addresses grid-specific attacks such as those on *load frequency control* and *interdiction attacks* and deception attacks such as *circuit breaker attacks* and *load altering attacks*.

It could be argued that the attacks as in (2) and (3) could have been easily grouped under deception attacks. A different example of disruption attacks, rather than (2), is an erasure, where the goal of the attacker is simply to prevent authorized entities from being made available the information that is required for their operation. Such an attack can be expressed as

$$y^a[k] = \begin{cases} \emptyset & \text{if } A_1 \text{ occurs,} \\ y[k] & \text{otherwise.} \end{cases} \tag{4}$$

where $\emptyset$ denotes the total lack of arrival of the data at the intended recipient.

In addition to the above perspective which is that of an attacker, a defender's perspective is important as well. One can argue that the focus of cyber-security (Rus et al., 2018) is from such a perspective and seeks to provide protection to a system by securing key components through firewall, encryption, etc. However, as the complexity of the overall system increases, it becomes difficult to ensure that a defense mechanism of the entire system can be guaranteed only through protection of every one of its individual components. Rather, a systems perspective is needed, which focuses on prevention of these attacks, and if attacks do occur, ensure that the system is resilient by containing the impact of these threats, and/or detect and isolate these threats and recover quickly. This is the focus of the next subsection.

2.2. Defense mechanisms

In this paper, we characterize three defense mechanisms, employed either prior to, or during the occurrence of the attack, to ensure CPS security. In order to present these three mechanisms in a unified manner, we consider an overall threat assessment metric illustrated in Fig. 4. Our thesis is that in order to develop a comprehensive defense mechanism for security, all three components of prevention (to postpone the onset of an attack), resilience (to contain the maximum impact of the attack and operate as closely to normal as possible), and detection and isolation (to identify the source of the attack, isolate the corrupted subsystems, and restore the normal mode as quickly as possible) are equally important and have to be layered in. If the defense strategy relies on detection alone, then the threat of the same attack recurring is not minimized. In addition, in the interval between the onset of the attack and detection, the system could experience a significant damage. A good example of such a scenario is the Stuxnet (Chen, 2010). Maroochy is also an outcome of the lack of detection and resilience mechanisms (Slay & Miller, 2007). The absence of resilience in

RQ-170 is apparent, as the control system was unable to defend against the spoofing attack. It could be viewed that preventive mechanisms are active prior to the attack whereas resilience and detection and isolation mechanisms are invoked during the attacks and until the system is restored to normal operation.

Each of the three defense mechanisms represents a certain point of view of ensuring security and therefore corresponds to a certain control methodology and related systems tools. The goals, the tools used, and the resulting performance are therefore intimately connected with the defense mechanism. In the sections that follow, the control methodology, the tools, and the results reported in the literature are provided in detail.

2.2.1. Prevention mechanisms

Methods in this category are to guard against disclosure attacks, which start from an infiltration stage to steal the vital information of the system and leverage them in future attacks. A simple example of this stage is through an insider (like the case in Maroochy attack) or *Advanced Persistent Threats* (APTs), an attack in which the access of the system is given to an unauthorized user in a stealthy fashion for an extensive period of time (Chen, Desmet, & Huygens, 2014). We group defense mechanisms in this category into two cases; Cryptography and Randomization. The former is a long-standing topic with its underpinnings in computer science and extensively studied (Katz, Menezes, Van Oorschot, & Vanstone, 1996). The latter, on the other hand, is grounded in control theory and has a rich history in robust control problems (Milanese, 2013).

(i) Cryptography:

Cryptography is the science of constructing and analyzing protocols that prevent third parties or the public from reading private messages. Modern cryptography started after World War II making use of the concept of public key (Diffie & Hellman, 1976), Fig. 5 (a). The idea behind cryptography is to make sure that the data between a sender and a receiver cannot be revealed via an unauthorized user. Authentication can be checked with sharing the secure acknowledge messages. Fig. 5(b) shows why making use of encryption and decryption is helpful in maintaining the confidentiality of data. However, if the eavesdropper has access to the points between decryptor and B, or encryptor and A, it can still read the message. As A and B can be any of three components, sensors, communication network, or actuators, shown in Fig. 1, this kind of attacks may take place in CPS. However, if a form of encryption that allows computation on ciphertexts is used, it can prevent the eavesdropper from accessing these messages. Farokhi, Shames, and Batterham (2017) and Darup, Redder, Shames, Farokhi, and Quevedo (2018) discuss a homomorphic cryptographic platform with closed-loop stability analysis to address. An application of this method to secure transportation systems is discussed in Farokhi, Shames, and Johansson (2017). A key management scheme for privacy issues in SCADA systems is also proposed in Rezai, Keshavarzi, and Moravej (2013). A polynomial-based scheme for a symmetric key generation in SCADA is discussed in Pramod and Sunitha (2015) and a cryptographic framework for the threats in cyber-physical systems is analyzed in Burmester, Magkos, and Chrissikopoulos (2012). Also Sherif et al. (2017) proposes a similarity technique between encrypted data to preserve the privacy of ride-sharing autonomous vehicles. Secure estimation with privacy assurance of the encoded data is discussed in Wiese et al. (2018).

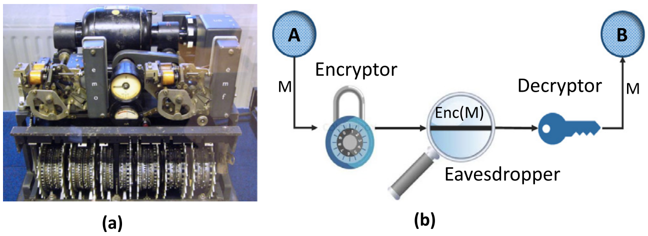

Fig. 5. (a) German Lorenz cipher machine, used in World War II to encrypt very-high-level messages, (b) Encryption and Decryption's roles in confidentiality.

(ii) Randomization:

Randomization as a defensive tool is utilized to confuse the potential attacker and has proved useful whenever the predictability of the deterministic rules may be leveraged by the attackers to obtain key information of the system, potentially for conducting much more advanced attacks. Randomized algorithms have proved useful in a wide range of mathematical and algorithmic problems (Motwani & Raghavan, 2010). Randomization as a robust control technique has been employed in the last decade (Frasca, Ishii, Ravazzi, & Tempo, 2015; Milanese, 2013). Most of the techniques which aim to provide a confidentiality service use randomization of data. An example of masking the private data in the presence of an adversarial agent is Mo and Murray (2017). The regular (non-adversarial) agents obtain the correct states and compute the average consensus using the masked data with a noise. A similar technique in a network of agents is proposed by Nozari et al. (2017), where the privacy of the states is preserved in an approximate manner. The latter method uses the differential privacy technique to tackle the problem (Corts et al., 2016; Dwork, 2011). The idea there is to use an alternative randomized data set to maintain the main data set from confidentiality breaches. The idea of randomization has been proposed also in *adversarial machine learning* (Huang, Joseph, Nelson, Rubinstein, and Tygar (2011)). In Gupta, Katz, and Chopra (2017), the idea of masking data to achieve the exact average consensus in the presence of an eavesdropper is proposed. Dibaji, Pirani, Annaswamy, Johansso, and Chakrabortty (2018) proposes a random gain selection method to secure the closed loop system against disclosure attacks on A_3 and A_9.

2.2.2. Resilience mechanisms

Resilience is a property defined as the ability to withstand and recover from severe stresses induced by natural stresses or deliberate attacks (Annaswamy, Malekpour, and Baros, 2016; Fawzi, Tabuada, and Diggavi, 2014; Khargonekar, 2015; Obama presidential policy; Rieger, Gertman, & McQueen, 2009). Resilience may not be an inherent property of the system and needs to be bestowed through a suitable design of the control system. A large number of the methods reported in the literature can be viewed as a resilience-increasing mechanism. In what follows, we group these methods into four types, which include (i) Game theory, (ii) Event-triggered Control, (iii) Mean Subsequence Reduced algorithms, and (iv) Trust-based approaches. While (i) and (ii) are based on state-space methods, (iii) and (iv) are graph-based.

(i) Game-theoretic methods:

A game-theoretic approach that provides resilience consists of trying to maximize the price of attacking a system or minimize the damage that an attacker can apply to the system. Game theory, in a nutshell, is an interaction between two or multiple players, where each player tries to optimize some objective function. The challenging part of games is that the objective function of a player depends on the choices of at least one other player in the game. Thus, each player cannot optimize its objective independent of choices of other players.

There is a vast literature on game-theoretic approaches to the security and resilience of control systems since the past decade. These approaches vary depending on the structure of the cyber-physical system or based on the specific type of malicious action acting on the cyber layer. Each of these two approaches is discussed briefly as follows:

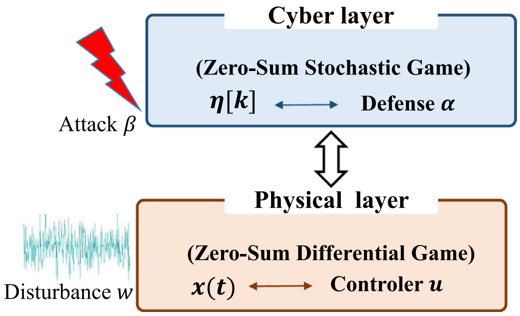

Fig. 6. Schematic figure of games in games in physical and cyber layers.

The first approach is to model the game for the security of CPS based on the structure of the cyber and the physical layers (Amin, Schwartz, & Sastry, 2013; Chen & Zhu, 2015; Clark, Zhu, Poovendran, & Başar, 2013; Ferdowsi, Saad, & Mandayam, 2017; La, 2017; Sanjab & Saad, 2016; Sanjab, Saad, & Başar, 2017; Zhu, Bushnell, & Başar, 2013; Zhu, Tembine, & Başar, 2010). One of the common approaches is to define games in both physical and cyber layers. More formally, considering that in the physical layer, the evolution of the system is modeled with the following general dynamics

$$\dot{x}(t) = g(t, x, u, w, \eta(t, \alpha, \beta)), \quad (5)$$

where $g(.)$ is a nonlinear function of the state x, the control action u, the disturbance effect w and $\eta(t, \alpha, \beta)$ which is a switching signal indicating the state of the cyber-layer. Here t is the time and α and β are the actions of the attacker and defender in the cyber layer, respectively. Parameter η evolves in discrete time, e.g., Markov jump model, in the cyber-layer which makes the overall hybrid system shown in Fig. 6. The concept of *games-in-games* reflects two interconnected games, one in the physical layer and the other in the cyber layer. At the physical layer control system, a zero-sum differential game between the robust controller and the disturbance is used to design an $\mathcal{H}_\infty$ controller for achieving robust performance for uncertain parameters or disturbances (Pan & Başar, 1999). At the cyber layer defense system, a zero-sum stochastic game between a defender and an attacker is used to design an optimal cyber policy for ensuring system security (Zhu & Başar, 2011).

Another approach is based on the type of the attack and malicious behaviour (Horák, Zhu, & Bošanskỳ, 2017; Khanafer, Touri, & Baar, 2013; Miao, Zhu, Pajic, & Pappas, 2018; Ugrinovskii & Langbort, 2017; Wu, Li, & Shi, 2017). More particularly, in this case, depending on the type of adversarial or malicious behavior that is active or passive, an appropriate game strategy, e.g., Nash or Stackelberg, has been discussed. More specifically, the interaction between a jammer and a passive defender can be reasonably captured by a Stackelberg game in that the jammer is an active player who sends signals at an intended level to interfere with communication channels while the legitimate user rationally defends itself from such an attack. On the other hand, in the case where the defending user behaves actively or either side has an information advantage, the Nash equilibrium becomes a reasonable solution concept (Felegyhazi & Hubaux, 2006; Gupta, Langbort, & Başar, 2010). Another example is eavesdropping action. As eavesdropping is a passive attack where an eavesdropper receives information that leaks from a communication channel, the behavior of an eavesdropper can be viewed as that of a follower in a Stackelberg game against a user who employs active defenses (Manshaei, Zhu, Alpcan, Basar, & Hubaux, 2013). Recently, an attacker-defender game framework on networks with unknown topology is proposed in which the defender injects control inputs to reach a synchronization while attenuating the (worst case) attack signal from adversarial agents (Vamvoudakis & Hespanha, 2018a; 2018b).

In addition to the above game-theoretic approaches, other approaches have been proposed as well. For instance, the evolution of network control systems has been modeled as cooperative games (Marden, Arslan, & Shamma, 2009) and the resilience of these cooperative games to the actions of adversarial agents or communication failures have been investigated. In Brown, Borowski, and Marden (2019), Brown and Marden (2017) and Amin, Schwartz, et al. (2013) the effect of adversarial agents and communication failures on a cooperative game was discussed. Moreover, in Vamvoudakis, Hespanha, Sinopoli, and Mo (2014) a zero-sum game for the problem of estimation under attacked sensors is suggested. In order to address the threats on cloud-based control systems, a signaling game is designed to model the trust between the defender and the threats (Chen & Zhu, 2017; Pawlick, Farhang, & Zhu, 2015).

(ii) Event-triggered control:

Based on how frequent the attacks occur, event-triggered control schemes instead of time-triggered schemes emerged as appropriate tools to increase the resilience of control systems (for an introduction to event-triggered control, refer to Heemels, Johansson, & Tabuada, 2012). Sensor disruption attacks (also called jamming or DoS), in some time intervals, on measurements (A_1 in Fig. 1), are among the threats whose effects can be mitigated via appropriate event-triggered control policies. Event-triggered control techniques have been used to design the sequence of control inputs $u(t_k)$ in order to preserve the input to state stability of the closed-loop system. The DoS attacks in these works are limited by the frequency and length. The application of event-triggered control to the resilience of cyber-physical systems has been studied in De Persis and Tesi (2014), De Persis and Tesi (2015), De Persis and Tesi (2018), Cetinkaya, Ishii, and Hayakawa (2017) and Sun, Peng, Zhang, Yang, and Wang (2018). In these works, the control input is sample-and-hold in the time sequence of $t_k - t_{k-1} > \delta$ instead of periodic sampled-data systems. The triggering function to generate a new control input is based on the errors of state variables $x(t_k) - x(t)$. For a comprehensive survey on DoS attacks and event-triggered control tools against them, the reader can refer to Cetinkaya, Ishii, and Hayakawa (2019) and the references therein. In addition to the case of disruption attacks, mitigating the effects of computational deception attacks (A_6 in Fig. 1) via event-triggered control techniques has been investigated (Lei, Yang, & Yang, 2016; Yang, Lei, & Yang, 2017).

(iii) Mean Subsequence Reduced (MSR) algorithms:

MSR is a resilient control approach in which at each time of the updates, the controller, in order to not get affected by the attacks, ignores the suspicious values and computes the control input. One of the well-known applications of MSR algorithms is against *Byzantine* threats. Byzantine nodes are the computational nodes that, in an adversarial manner, send inconsistent information to their neighbors (Dibaji, Ishii, & Tempo, 2018; LeBlanc & Koutsoukos, 2018; LeBlanc, Zhang, Koutsoukos, & Sundaram, 2013; Lynch, 1996; Usevitch & Panagou, 2018b; Zhang, Fata, & Sundaram, 2015). Byzantine attacks have been investigated in the '80s in computer science (e.g., Lynch, 1996). Recently, Byzantine consensus is getting revisited, again in the computer science community, to develop secure and reliable cryptocurrencies (see, e.g., Algorand). MSR algorithms have been applied to distributed computational problems, including consensus (Dibaji & Ishii, 2017; Dibaji, Ishii, et al., 2018; LeBlanc et al., 2013), distributed state estimation (Mitra & Sundaram, 2018), synchronization (LeBlanc & Koutsoukos, 2018), clock synchronization (Kikuya, Dibaji, & Ishii, 2018), and distributed optimization (Sundaram & Gharesifard, 2016). MSR algorithms act as local filters, in which, by assuming that the maximum number f of malicious agents in the network is known, every node disregards

f largest and f smallest values from its neighbors. Hence, there is no need to have a knowledge about the global topology.[3] In these studies, network-theoretic necessary and sufficient conditions for the convergence of MSR algorithms have been introduced. The critical property is called *graph robustness* which is a measure of connectivity within a graph and characterizes how well groups within the network are connected via multiple paths. Network robustness was first introduced by LeBlanc et al. (2013) for the resilient consensus of agents with first-order interaction dynamics. Graph robustness can be determined with linear programming (Usevitch & Panagou, 2018a) and in general was shown in Zhang et al. (2015) to be a computationally hard problem but can be obtained almost surely in random large networks. While a similar problem of multiple sensors being attacked simultaneously has been addressed in Fawzi et al. (2014) as well, the defense approach taken is different from the MSR-approach and is based on compressed sensing and error correction.

(iv) Trust-based approaches:

Trust-based methods have been investigated for not only cybersecurity but also general problems where some of the subsystems may be untrustworthy. Mikulski, Lewis, Gu, and Hudas (2011), Mikulski, Lewis, Gu, and Hudas (2014) and Haus et al. (2014) have used a multi-agent approach in order to improve overall resilience. This strategy is equivalent to *redundancy-based approaches* in graphs and is based on the assumption that if the number of attacks is sufficiently small, correct information can flow through the paths formed by trusted nodes. Trust-based approaches have been investigated in Jiang and Baras (2006) and Abbas, Laszka, and Koutsoukos (2018) to spread the information in a multi-agent system in the presence of adversarial nodes. An alternative way is to define a function of trust and update the trust value between the nodes as the system evolves. In such approaches, the reliance and effects of each healthy node on its neighbors is a function of the trust value. A survey on how to use trust models in different network domains is Momani and Challa (2010). Trust-based approaches have been used mainly for defense against deception attacks and more often in the context of sensor networks (Ahmed, Bakar, Channa, Haseeb, & Khan, 2015; Khan & Stanković, 2013) and in DC microgrid control (Abhinav, Modares, Lewis, & Davoudi, 2019).

(v) Other approaches:

In addition to the above four methods, resilience mechanisms have been proposed using a variety of other control methods. Sun, Peng, Yang, Zhang, and He (2017), for instance, suggests a resilient control assuming that the probability of the disruption attacks at each time is at least partially known. A sliding mode control for the resilience against DoS attacks in nonlinear and chaotic systems has been proposed in Zhao and Yang (2017). An acknowledge-based cheating scheme is proposed in Ding, Ren, and Shi (2016). Another technique is Liu, Xu, Li, and Liu (2017), where it proposes a decomposition of Kalman filters as a weighted sum of local state estimates under sparse sensor deception attacks (A_2) into a more secure estimation framework. With the help of compressed sensing methods and their relation in error corrections over the reals, Fawzi et al. (2014) proposes a decoding algorithm to recover true states despite the existence of attacks. Moreover, by using separation principle, it shows that if the system is controllable, one can enforce the number of correctable errors (attacks) to be maximum without loosing the performance of the system (Fawzi, Tabuada, & Diggavi, 2012). In Satchidanandan and Kumar (2018a,b) when the state space is subject to malicious actions, a decomposition of the state space into a securable and an unsecurable subspace is carried out, where the malicious nodes cannot degrade the state estimation performance in the former but only along the latter. Another recent work is Dibaji, Pirani, et al. (2018) where for defending against the deception attacks (A_6) on the cyber layer, an information retrieval approach is hired so that the state feedback, at each time step, makes use of healthy and unattacked data. Finally, in Lu, Chang, Zhang, Marinovici, and Conejo (2016), a Lyapunov stability method is employed for DoS attacks in wide-area control of power systems. In Dibaji, Safi and Ishii (2019), a resilient distributed retrieval algorithm based on secure broadcasting and accepting has been employed to compute averaging over strongly robust graphs. Yet another tool used for obtaining resilience, mainly against disclosure attacks, is the use of privacy loss as a penalty component in the underlying cost function (e.g., Tanaka, Skoglund, Sandberg, & Johansson, 2017). By constructing an information theoretic measure, I, between two data sets X and Y, given by

$$I(X;Y) = H(X) - H(X|Y), \tag{6}$$

where H is the entropy, to form the penalty component; the approach consists of the optimization of this cost function with and without the penalty component and evaluating the resulting tradeoff.

2.2.3. Detection & isolation mechanisms

We now direct our attention to the third component illustrated in Fig. 4, detection & isolation. As the name suggests, this corresponds to a quick detection & isolation of the attack. These mechanisms, similar to the resilience described in Section 2.2.2, get activated after the attack, and constitute the bulk of the research in CPS security from the controls community. It should be noted that methods such as patch and pray (Rus et al., 2018), stemming from the computer science community, can be grouped under this category as well. This is commonly used in cybersecurity, and has to do with responding to existing threats and hoping that the results will deter future attacks.

A detection mechanism usually uncovers the existence of an attack by monitoring its effects on the outputs of the system. In addition to detecting the existence of an attack, stronger strategies can be proposed to *identify* (or *localize*) the set of nodes/signals that are attacked (e.g., Pasqualetti, Dorfler, & Bullo, 2013; Pasqualetti, Dorfler, & Bullo, 2015a). If the effect of the attack signals cannot be traced by the outputs, they are called *covert* (Teixeira, Sou, et al., 2015) or *stealthy* attacks (Teixeira, Shames, Sandberg, & Johansson, 2012). The survey papers Ding et al., 2018; Giraldo et al., 2018 have reviewed some detection mechanisms for deception, as well as disruption attacks in cyber-physical systems. Detection tools stemming from the control-theoretic literature have been used primarily against deception attacks while in the computer-science literature have been employed for confidentiality attacks as well (Zeldovich, 2014).

In what follows, we classify all detection & isolation methods proposed in the controls literature into five categories which include Observer-based techniques, Analytical consistency, Watermarking, Baiting, and Learning-based anomaly detection.

(i) Observer-based techniques:

Observers in control systems are designed to estimate unmeasurable state variables. Detection can therefore be enabled using observers and a comparison between the resulting state estimates in the healthy and attacked cases, often termed *residues*. If the residues exceed a certain threshold, an alarm is activated (Teixeira, et al., 2015). A common method used for designing such observers is geometric control theory (De Persis & Isidori, 2001; Massoumnia, Verghese, & Willsky, 1989). Termed Unknown Input

[3] One reason that in such algorithms detection is not utilized is that detection-based approaches require global topology of the network and have a heavy computational burden on each node (Sundaram & Hadjicostis, 2011).

Observers, the approach consists of using this method in the presence of unknown input that here it refers to the attacked inputs which cannot be relied upon. Another example can be found in Pasqualetti et al. (2013), where deception and disruption attacks on both sensors and actuators are modeled as linear algebraic conditions for detection and identification of the attacked sets. Pasqualetti et al. (2013) also proposes centralized and distributed filters. Pasqualetti, Dörfler, and Bullo (2015b) proposes several algorithms for distributed and decentralized detection and identification of systems with some certain coupling features. The identification phase, in particular, is based on a combinatorial search on all potential sets of attacks. Same ideas have been used in multi-agent systems in the presence of misbehaving nodes (Chen, Kar, & Moura, 2017; Pasqualetti, Bicchi, & Bullo, 2012; Sundaram & Hadjicostis, 2011). Moreover, Murguia, van de Wouw, and Ruths (2017) uses the same technique on sensor attacks (A_2 in Fig. 1) and analyzes the reachable sets of attacks. However, in these works, different matrices for prediction of the outputs and detection have to be used which take significant amount of memory and computational complexities. A scalable version of these works is Shoukry et al. (2018) where attack-free sensors using a Satisfiability Modulo Theory (SMT) are identified with Luenberger observers.

A specific subcase of observers corresponds to the case when the underlying model is static. For example, if

$$z = Hx + e, \tag{7}$$

where H is the Jacobian matrix, z is the measurement, x is the state variables, and e is the measurement/modeling noise, the goal is to estimate x using z, in the presence of attacks, which may either be on the sensor z or on H. This problem is ubiquitous in power systems where measurements of either voltage or current are not possible everywhere in the network but have to be estimated (Gomez-Exposito & Abur, 2004). For example, if deception attacks on sensors (e.g., A_2 in Fig. 1) occur, Teixeira, Amin, Sandberg, Johansson, and Sastry (2010), Sandberg et al. (2010), Chakhchoukh and Ishii (2015), Chakhchoukh, Vittal, Heydt, and Ishii (2017) and Liu, Ning, and Reiter (2009) propose a solution based on robust signal processing techniques such as Least Trimmed Squares (LTS) to minimize the residue. In general, the underlying idea here is to treat the corrupted data and ignore them as outliers before doing the required analysis. Application of such works on Automatic Generation Control (AGC) and SCADA are studied in Andersson et al. (2012).

(ii) Analytical consistency:

Physical coupling and the correlation between state variables and control decisions across individual subsystems of a CPS can be an effective way for detecting attacks on the communication layer. They can enable us to partially or fully reconstruct a signal at one physical location using signals measured at other locations, forming redundancy relationships that can be used to determine if data have been manipulated during communication. These redundancy relationships, often referred to as *analytical consistency*, are quite common in spatially distributed CPSs. For example, in power systems, one approach to design controllers that both stabilize frequency and minimize dispatch cost is Zhao, Mallada, and Dörfler (2015)

$$C_j\dot{u}_j(t) = -\omega_j(t) - C_j\sum_{k\in\mathcal{N}_j}(C_ju_j(t) - C_ku_k(t)), \tag{8}$$

where u_j is the adjustable mechanical power input of generator j with an associated cost C_j, $\mathcal{N}_j$ is associated with the neighbors of j-th generator bus, and ω_j is the frequency deviation of generator j from the synchronous frequency, with its dynamics governed by the swing equations. The coupling between the control decisions of different generators in (8) leads to an approximate constraint of the form

$$\begin{aligned}C_j\dot{u}_j(t) - C_k\dot{u}_k(t) = -\frac{\dot{p}_{jk}}{B_{jk}} + C_k\sum_{i\in\mathcal{N}_k}(C_ku_k(t) - C_iu_i(t))\\ -C_j\sum_{i\in\mathcal{N}_j}(C_ju_j(t) - C_iu_i(t)),\end{aligned} \tag{9}$$

where p_{jk} denotes the power flow between generator buses j and k. Thus, if any feedback information is dropped or corrupted during communication, then physical truths such as (9) can be checked for consistency, and that too via local sensing and computation as the power flow p_{jk} can be measured locally at node j. Cross-checks between physics and computation for detecting anomalies in a CPS have been used in some recent papers such as Macwan et al. (2016) and Nicol, Sanders, and Trivedi (2004). An effective future research direction on this topic would be to extend this concept to control design, i.e., to develop control strategies such as (9) that enhance spatial redundancy, and equip us with additional consistency checks than what is simply available from the open-loop system. The controller should preferably be implemented in a distributed way as centralized computation during these types of attacks can be quite dangerous.

(iii) Watermarking:

The concept of Watermarking is often used to authenticate an entity. For example, a watermark on a piece of paper is effectively a signature that cannot be erased. This concept is used in the context of detection & isolation by constructing a suitable metric and a perturbation of the input signal such that the metric dropping below a certain threshold signals the presence of an attacker (Mo, Hespanha, & Sinopoli, 2014b). Particular success of the watermarking approach has been reported in the context of replay attacks and is discussed below.

A replay attack corresponds to one where the attacker hijacks the sensors or eavesdrops for a certain amount of time and replays the same data over and over again. In particular, a recorded horizon of data, in normal conditions, is sent to the monitors of the operators so that the alarms would not be triggered and the operators are tricked into thinking that the closed-loop system is operating normally (Chen, 2010; Mo et al., 2014b). Replay attacks are sometimes grouped under the category of disruption attacks as they cause the current data to become unavailable. Obviously they can also be viewed as deception attacks (Ding et al., 2018).

Replay attacks can be overcome using watermarking (Mo et al., 2015) by perturbing an optimal control input in a particular manner. In order to ensure that the controller does not become overly sub-optimal, Mo et al. (2015) discusses methods to maximize the likelihood of attack detection while constraining the effect of the watermark on ideal system operation. An application of watermarking in SCADA networks is introduced in Mo, Chabukswar, and Sinopoli (2014a). A watermarking-based approach to defend against replay attacks in multi-agent systems (A_6) is proposed in Khazraei, Kebriaei, and Salmasi (2017). In Ferrari and Teixeira (2017b), a multiplicative sensor based watermark is used to detect replay attacks on sensors. In Ferrari and Teixeira (2017a), the same multiplicative watermarking technique is used to detect *routing attacks* where the wires of the sensors are intentionally swapped (A_2). Motivated by the replay attacks, Lucia, Sinopoli, and Franze (2016); Mo et al. (2015) posit a form of replay attacks on linear stochastic systems and propose a χ^2-detection method to alert the system operator of the presence of an adversary conducting a replay attack.

The effects of watermarking on a more general set of sensor deception attacks is studied in Satchidanandan and Kumar (2017). It is also shown that a set of uncompromised actuators, each injecting its own added watermark signal, can be used to check the honesty of the sensors which should report back measurements that contain a history of the effect of watermarking. More specifically, Satchidanandan and Kumar (2017) analyzes a number of systems including SISO (Single Input, Single Output) and MIMO (Multi-Input, Multi-Output) linear systems with Gaussian noise models and shows that the asymptotic behavior of a system with the watermark constrains the damage a sensor spoofing attacker can do without being undetected. Dynamic watermarking has been validated in power systems in Huang et al. (2018) and in transportation systems (Ko, Satchidanandan, & Kumar, 2016). A combination of Gaussian and Bernoulli processes to generate a watermarking signal is suggested in Weerakkody, Ozel, and Sinopoli (2017) for general detection of deception attacks on sensors and actuators. The same idea is proposed for use against covert attacks (on both A_2 and A_8) in Hoehn and Zhang (2016) by inserting a modulation system between A_8 and the actuation to misguide and confuse the attacker.

(iv) Baiting:

Like the watermarking case, suppose we begin with a worst-case scenario where the attacker is assumed to have a complete access to the entire system dynamics, all of its sensors, and all of its actuators. The question is if one can design a method by which such an attacker cannot remain stealthy and can be revealed. These methods, termed *Baiting* (Flamholz, Annaswamy, & Lavretsky, 2018) and *Moving Target* (Weerakkody, Mo, & Sinopoli, 2014), then seek to design the system in a way such that this worst-case scenario can be detected. In Flamholz et al. (2018) and Teixeira et al. (2012), the method consists of baiting the attacker to reveal themselves by introducing an arbitrary offset in the system dynamics which guarantees that a worst-case attack proposed in Kwon, Liu, and Hwang (2014) can be quickly detected. By introducing virtual state variables in addition to the original system state variables, Weerakkody and Sinopoli (2015, 2016) offer a method to prevent an eavesdropping attacker from inferring system knowledge, A, from the system output and control signals. This lack of knowledge of A prevents the attacker from achieving worst-case stealthy attacks. Moving Target Defenses (MTD), in a general cyber-security context, are defense schemes in which the defender varies system attributes in order to introduce unpredictability into the attack surface (Jajodia, Ghosh, Swarup, Wang, & Wang, 2011). The moving target approach can also be utilized to detect the presence of attacks on both the control inputs and sensor measurements (Weerakkody & Sinopoli, 2015).

(v) Learning-based anomaly detection:

Anomaly detection is a technique in machine learning to detect the presence of suspicious data (Ng, 2018). For a review on applications of anomaly detection in computer networks, the reader is referred to Garcia-Teodoro, Diaz-Verdejo, Maciá-Fernández, and Vázquez (2009) and Tsai, Hsu, Lin, and Lin (2009). Anomaly detection techniques in power systems are introduced in Ten, Hong, and Liu (2011). Applications of Neural Networks (NNs) and Baysian learning are studied for anomaly detection in the context of security in the presence of attacks (He, Mendis, & Wei, 2016; Kailkhura, Han, Brahma, & Varshney, 2013; Lippmann & Cunningham, 2000; Reddy, 2013; Shitharth and Prince Winston, 2017). Particularly, in the latter, it is assumed that the Byzantine nodes are aware of the true hypothesis and they are compromised to degrade detection performance. The problem of distributed detection is formulated as a *Binary hypothesis test* by the sensors. A direct application of this work is in Pal, Sikdar, and Chow (2018) to detect deception attacks on Phasor Measurement Units (PMUs) data. Also, it is used in Barbosa, Sadre, and Pras (2013) for SCADA networks, where a *whitelist* is generated by learning the network legitimate traffic for a given period of time and is used for detection of other threats. An unsupervised detection method is employed in Almalawi, Yu, Tari, Fahad, and Khalil (2014) to detect anomalies.

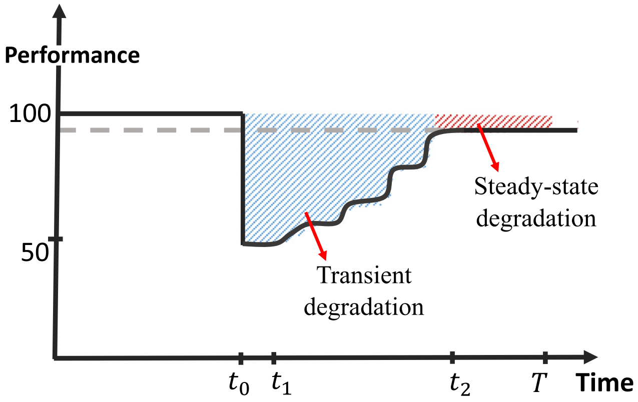

Fig. 7. A conceptual representation of performance degradation due to an attack on a CPS and its recovery over time.

There are several works that discuss designing anomaly detectors through other tools. In this direction, Liao and Chakrabortty (2018) uses a Round-Robin algorithm for localizing the deception attacks on power systems. Another localizing work in power systems is Nudell, Thomas, Nabavi and Chakrabortty (2015), where a graph-theoretic technique is used to localize where the effect of an attack exists in a wide-area control architecture. A recursive distributed Kalman filter in the presence of sensor attacks (A_2) is developed in Ding, Li, Quevedo, Dey, and Shi (2017) and Mishra, Shoukry, Karamchandani, Diggavi, and Tabuada (2017). Some works also combine the resiliency with detection methods and investigate the resilience of the detection tools under attacked conditions (e.g., Pajic, Lee, & Pappas, 2017). In Li, Lu, Wang, and Choo (2017), a majority voting is utilized for detetcion of deception attacks in smart grid. Application of Kalman filter in detection of replay attacks (A_2) in SCADA systems has been discussed in Do, Fillatre, and Nikiforov (2017). In Magiera and Katulski (2015) authors present an application of spatial processing methods for GPS spoofing detection and mitigation. The method is based on signal processing techniques rather than the topology of the network. Obtaining the direction of arrival or angle of arrival estimate of the receiving signals is the key technique utilized for applying the method.

2.3. Security metric

With various defense mechanisms and underlying systems and control tools described above, one may need to ascertain the suitability of one method over another for a given application. For this purpose, a security metric that quantifies the benefit obtainable from a given method is needed. There is a rich history in the literature on defining metrics for the security of systems (e.g., Annaswamy et al., 2016; Baros, Shiltz, Jaipuria, Hussain, & Annaswamy, 2017; Teixeira et al., 2015). One could use Fig. 4 for this purpose as well, and denote it as the ratio of the areas under the curves Γ_s and Γ.

A somewhat different representation of secure performance is often utilized in the literature (Bruneau et al., 2003). Rather than viewing security as the reduction of threat as in Fig. 4, one can view security and resiliency as an improvement in the steady-state performance following an attack (see Fig. 7). In Fig. 7, it is assumed that an attack occurs at t_0, that suitable defense mechanisms are invoked at t_1 which allows performance to gradually recover until t_2, leaving a net steady-state degradation. One can then propose a security metric based on this degradation as

$$\mathcal{S}(I) = \frac{\sum_{k=0}^{k=T} \mathcal{D}(I^*[k] - I[k])}{\sum_{k=0}^{k=T} \mathcal{D}(I^*[k])}, \tag{10}$$

where $\mathcal{D}(z)$ is a norm of z, the variable $I(t)$ denotes the vulnerability of the underlying system. I^* is the value of the variable I in a nominal and healthy condition and $[0, T]$ is the period of interest. That is $\mathcal{S}(I)$ can be viewed as the amount of degradation of the system performance when attacked compared to the nominal performance. In particular we can see that the transient degradation, denoted by the blue shaded area in Fig. 7, corresponds in some sense to the numerator of $\mathcal{S}(I)$ in (10). Alternatively, a security measure can be constructed based on steady-state degradation and simply correspond to the red shaded area.

Parameter I varies depending on the application and threat assessment. For example, I can be in the form of mutual information (Tanaka et al., 2017) or the difference of agents' values and their consensus value, if the goal is to reach the consensus amongst the agents (Sundaram & Hadjicostis, 2011). Some other works (e.g., Sandberg et al., 2010; Sandberg & Teixeira, 2016), define security metric as the capability of the attacks' undetectability.

In the following example, we apply security metric (10) to resilient consensus dynamics.

Example 1. Consider the case where agents synchronously update their states as

$$x_i[k+1] = w_{ii}[k]x_i[k] + \sum_{j\in\mathcal{N}_i} w_{ij}[k]x_j[k] + a_u[k], \tag{11}$$

where $\mathcal{N}_i$ is the neighbor set of agent i and w_{ij}s are some positive weights. Here, the update rule of agent i is affected by an attack signal $a_u[k]$. If we denote the consensus value $\bar{\boldsymbol{x}}$, we can define $I[k] \triangleq |\boldsymbol{x}[k] - \bar{\boldsymbol{x}}|$. Hence, $I^*[k]$ belongs to the cases when there is no attack signal. In an ideal case, $I^* = 0$. However, an attacker can inject the attack signal to make $I[k]$, and consequently the nominator of (10), arbitrarily large. In this direction, if we use one of the resilience mechanisms such as the MSR, we can make $I[k]$ small. We refer the reader to Section 5 of Dibaji and Ishii (2017) for details of numerical illustrations of this concept.

The security metric in (10) can be used for a resilient control design by incorporating an additional term to the standard cost function used in optimal control as

$$\mathcal{J} = \sum_{k=1}^{k=\infty} x^T[k]Qx[k] + u^T[k]Ru[k] + \mathcal{S}(I). \tag{12}$$

Such an addition obviously implies that a trade-off is introduced between resilient control and optimal control. It is noteworthy that the appropriate trade-off between resilience and optimality is labelled as an *intelligent* solution in transportation systems (see for example, Alam et al., 2016; Javed & Hamida, 2017; Woo et al., 2015).

In summary, in this section, we have presented attack models as well as three classes of defense mechanisms that have been proposed in the literature using systems and control methods. A threat assessment metric was proposed (in Fig. 4) in order to view these three methods in a unified manner, and quantified as in (10).

The threat assessment metric involves the identification of a suitable variable I that best denotes the most vulnerable quantity in a system under attack. It also includes the appropriate norm $\mathcal{D}$ that provides a suitable measure of the threat. In the next section, we discuss how the system threat and the corresponding variable I can be assessed. The role of the system network topology in this regard is also analyzed.

3. Threat assessment and network topologies

In this section, we review various threat assessment tools proposed in the literature. We also analyze the effect of network topologies on the three defense mechanisms discussed in Section 2.

3.1. Threat assessment

One can argue that even prior to assembling any defense mechanism and to understand which variables to be used for defining I in (10) for guarding against attacks, an overall analysis of the system vulnerability and a threat assessment needs to be carried out. In what follows, we survey various methods that have been proposed in the literature towards such a threat assessment.

One straightforward approach for threat assessment is to use a test bed to implement different attacks (Adhikari, Morris, & Pan, 2017). Another well-known technique is so-called *attack trees*, where the attacker's reachable points and possibilities are exhibited on an attack graph (Abdo, Kaouk, Flaus, & Masse, 2018; GhasemiGol, Ghaemi-Bafghi, & Takabi, 2016). A rich introduction on attack graphs can be found in GhasemiGol et al. (2016). Other quantitative vulnerability assessments in SCADA are discussed in Ten, Liu, and Manimaran (2008) and Yang, Cao, and Li (2015). For survey of threat assessment, the reader is referred to Cherdantseva et al. (2016).

The first step in threat assessment is the characterization of the worst case scenarios and the maximum damage that each attack can cause. Determining the conditions under which the maximum damage can be caused to the system is also part of the analysis. A good example of such analyses is Moghadam and Modares (2017) where it is shown that in a multi-agent system, attacks on the root nodes can destabilize the entire system with the access of at least one of its eigenvalues. Attack analysis mainly is carried out with the premise that the detection cannot be achieved by usual detection filters. This might be due to possibility of *zero dynamic attacks* or *stealthy attacks*, where in the former there are cases that the measured values are zero in the sampling times and the attack hides itself between the samples (Naghnaeian, Hirzallah, & Voulgaris, 2015) and in the latter the attack affects the measurement so that it becomes the same as at least one non-attacked case. The stealthiness of an attack is usually translated into certain algebraic conditions (Kwon et al., 2014) in the underlying system. Another example of attack analysis for introducing undetectable attack is discussed in Chen et al. (2017). In Guo, Shi, Johansson, and Shi (2018) and Bai, Pasqualetti, and Gupta (2015), Kullback-Leibler divergence is employed to develop a stealthiness and obtain worst-case attack policies as a trade-off between system performance degradation and attack stealthiness. In Miloševič, Tanaka, Sandberg, and Johansson (2017) and Bai, Pasqualetti, and Gupta (2017), worst case deception attacks are analyzed in stochastic systems and the number of sensors to secure the system using a Kalman filter approach is proposed. Attacks on Kalman filters have been also analyzed in Guerrero-Bonilla, Prorok, and Kumar (2017). Moreover, in Ntalampiras (2016) a Hidden Markov and a Neural Network model are used to classify the attacks. The detectability of the cyber-attacks on the state estimation problem (7) with graph-theoretical algorithms is analyzed in Bi and Zhang (2017). In Góes, Kang, Kwong, and Lafortune (2017), stealthy scenarios of sensor deception attacks in discrete-event systems are investigated through an *insertion-deletion* structure. Another example of such investigations is Hu, Wang, Han, and Liu (2018) where *insecurity* conditions that deception attacks remain undetected is studied. In Chen et al. (2017), *zero state inducing* attacks (on A_2) are introduced where the output of the system with such attacks would remain the same as the case when initial conditions are zero, with the assumption that the initial conditions cannot be altered by the attackers. Another type of attack arises when in sampled-data systems, the actuation hold is faster than sensing sampling times (Hirzallah & Voulgaris, 2018; Jafarnejadsani, Lee, Hovakimyan, & Voulgaris, 2017; Kim, Park, Shim, & Eun, 2016; Naghnaeian et al., 2015). In these works, a *zero-dynamic* output can be injected to the system so that while to the system operators it may appear that the output remains at zero, the actual system response is becoming unbounded. In Guo, Shi, Johansson, and Shi (2017) and Guo et al. (2018), the degradation of system performance under a linear attack policy on A_2 is analyzed. The timing of DoS attacks (A_1) is discussed in Krotofil, Cardenas, Larsen, and Gollmann (2014). The paper Anwar, Mahmood, and Pickering (2017) investigates the worst case attacks on state estimators (7). An attack analysis on SCADA water networks is discussed in Amin, Litrico, Sastry, and Bayen (2013). An optimization-based analysis on a single malicious agent in a network (A_6) is analyzed in Khanafer et al. (2013). Xie et al. (2011) is an example of threat assessment for electricity markets.

3.2. Effects of the network topology

Another important aspect that should be addressed when it comes to CPS security is the cost required to reduce I in (10). This cost may be in general networked systems dependent on the underlying topology. In this subsection, we explore the relation between the underlying network topology in a CPS and the three defense mechanisms discussed in Section 2. We discuss how prevention mechanisms may require a more dense graph, i.e. more number of edges, than what resilience mechanisms would require, and therefore relatively more expensive; resilience strategies in turn require more edges than prevention-based ones would. This also corresponds to the decreasing level of conservatism between these three defense mechanisms. Prevention mechanisms can be viewed as being most conservative, as they are designed assuming that very little is known about the system as well as the attack; in comparison, resilience mechanisms require more knowledge, and detection mechanisms require even more information. In addition, we can argue that as the information about the attack becomes less, the strength of the defense can be increased by making the corresponding graph topology denser. In the following section, we discuss the relation between network topology and the resulting properties of the three defense mechanisms. We begin with a few definitions before discussing the network topology properties.

Definition 1. (Some Graph Definitions): An undirected graph is denoted by $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ is a set of nodes (or vertices) and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of edges. The neighbors of node $v_i \in \mathcal{V}$ are given by the set $\mathcal{N}_i = \{v_j \in \mathcal{V} \mid (v_i, v_j) \in \mathcal{E}\}$ and the degree of node v_i is $d_i = |\mathcal{N}_i|$. A graph is called connected if there exists a path between any couple of nodes in $\mathcal{G}$. The *edge-boundary* of a set of nodes $S \subset V$ is given by $\partial S = \{(v_i, v_j) \in E \mid v_i \in S, v_j \in V \setminus S\}$. The *isoperimetric constant* of G is defined as (Chung, 1997)

$$i(G) \triangleq \min_{S \subset V, |S| \leq \frac{n}{2}} \frac{|\partial S|}{|S|}. \tag{13}$$

By choosing S as the vertex with the smallest degree we obtain $i(G) \leq \min_i d_i \triangleq d_{\min}$.

3.2.1. On prevention mechanisms

The effect of the structure of the underlying network has been studied in the context of prevention mechanisms using the concept of the *network expansion*. Network expansion, which is characterized by the isoperimetric or Cheeger constant $i(\mathcal{G})$, introduces how many edges connect any subset of nodes to the rest of the network. Network expansion has various applications in secure network coding and information diffusion in the networks (Hoory, Linial, & Wigderson, 2006; Pirani & Sundaram, 2016). This network connectivity measure, as we will see later, is the strongest compared to other well-known measures. This means that having a secure network to be preventive to attacks demands a high level of connectivity.

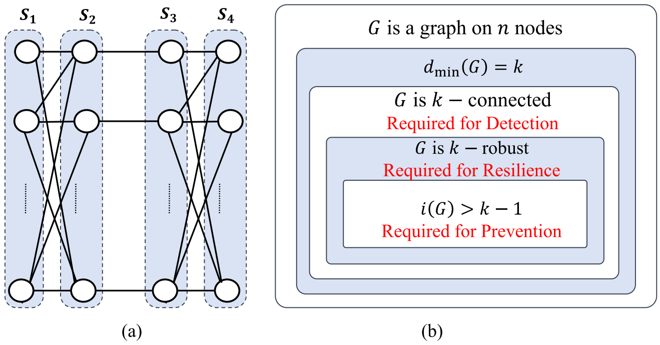

Fig. 8. (a) A graph showing the strength of different connectivity measures, (b) Venn diagram of various network connectivity indices.

3.2.2. On resilience mechanisms

A network connectivity measure used in the literature for ensuring the system resilience is called *network robustness*. In particular, network $\mathcal{G}$ is called k-robust, if for any disjoint and nonempty subsets of nodes in the network at least one of them has a node that is connected to k nodes outside of itself (LeBlanc et al., 2013). Network robustness has implications for the resilience of certain dynamics: if the network is $(2k+1)$-robust (for some non-negative integer k), then there are certain dynamics that allow the nodes in the network to reach consensus even when there are up to k malicious nodes in the neighborhood of every correctly behaving node (Shahrivar, Pirani, & Sundaram, 2017; Zhang et al., 2015). The network robustness is a weaker notion compared to network expansion, mentioned before. More particularly, if the network is k-robust, then we have $i(\mathcal{G}) \geq k$.

3.2.3. On detection mechanisms

A network is called k-vertex connected (or k-edge connected) if it is connected after removing up to any k nodes (or edges) from the graph (Gross & Yellen, 2005). The concept of node connectivity also has implications for the robustness of certain dynamics on networks. For instance, if the network is $(2F+1)$-connected (for some non-negative integer F), then there are certain information diffusion dynamics (or algorithms) that allow information to spread reliably in the network, even when there are up to F malicious nodes (in total) that deviate from the prescribed dynamics in arbitrary ways (Sundaram & Hadjicostis, 2011; Pasqualetti et al., 2012). Some works applied such a connectivity measure to ensure the privacy of a distributed online learning task against adversarial agents which try to reconstruct the updating rule of autonomous agents (Yan, Sundaram, Vishwanathan, & Qi, 2013). Network vertex and edge connectivity are the weakest notions of network connectivity compared to other connectivity measures such as network robustness and network expansion which were discussed earlier.

An example showing the relative strength of the above network connectivity measures is shown in Fig. 8. The gap between the robustness and vertex connectivity (and minimum degree) parameters can be arbitrarily large, as illustrated by the graph $\mathcal{G}$ in Fig. 8(a). While the minimum degree and node connectivity of the graph $\mathcal{G}$ are both equal to $\frac{n}{4}$, it is only 1-robust (consider subsets $S_1 \cup S_2$ and $S_3 \times S_4$). Moreover, this graph has isoperimetric constant of at most 0.5 (since the edge boundary of $S_1 \cup S_2$ has size $\frac{n}{4}$), but is 1-robust. The relationships between these different graph-theoretic measures of robustness are summarized in the Venn diagram in Fig. 8(b). We should note that although increasing network connectivity will result in increasing the tolerance of the control system, making extra communication links is costly in various applications. With this in mind, a few studies have looked at reaching a specific level of network connectivity with minimum number of links (Weerakkody, Liu, Son, & Sinopoli, 2017).

4. Summary and future works

A high integration of instrumentation, communication, and control into physical systems has led to the late study of CPS with increased attention. A key feature that is ubiquitous in CPS is a need to ensure their security in the face of cyberattacks. In this paper, we carried out a survey of systems and control methods that have been proposed for the security of CPS. We classified these methods into three categories based on the type of defense proposed against the cyberattacks, which include prevention, resilience, and detection & isolation. Prevention mechanisms are proposed to postpone and/or avoid disclosure attacks (Mo & Murray, 2017). Resilience, the property by which the maximum damage inflicted by the attack is contained, has been demonstrated in many papers using methods such as De Persis and Tesi (2015). Detection & isolation methods, as the names suggest, seek to restore the system to normalcy as quickly as possible by detecting and isolating the attack from the system (Pasqualetti et al., 2013). A unified threat assessment metric is proposed in order to evaluate how CPS security is achieved in each of these three cases. Also surveyed are risk assessment tools and the effect of network topology on CPS security. An emphasis has been placed on power and transportation applications in the overall survey.

The varied, impactful, and malicious nature of the actual cyberattacks underscores the huge importance of the study of CPS security. Given the scope of the systems and control methodology for achieving robustness, optimality, and efficiency in the presence of various perturbations, it is not surprising that the over two hundred papers cited in the survey correspond to systems and control methods by which such security in CPS can be achieved.

The papers referenced and the methods reported therein represent the first step towards achieving security in CPS. Unlike exogenous disturbances, cyberattacks correspond to a customized, system-specific, malicious and active inputs that can continuously increase in complexity as the system evolves. As a result, it is imperative that the defense mechanisms that are proposed continue to advance the state of the art, to not just keep in pace with the complexity of the attack, at least a few steps ahead.

The above clearly indicates that a lot more remains to be done to ensure CPS-security. We mention below a few specific directions of interest, in no particular order:

1. *CPS-security and machine learning*: The explosive interest in Machine Learning (ML), buoyed by its success in image and speech recognition, begs the obvious question of its role in ensuring CPS security. An important challenge here is to understand how anomaly detection with analytical guarantees can be carried out in the presence of dynamic inputs.
   Despite increasing role of ML in CPS applications, another important question that needs to be addressed is the inherent vulnerability that is associated with any ML tool. It is shown in Hein and Andriushchenko (2017) how non-resilient a classifier ML is against input data deception attacks. Likewise, several other researches show weakness of some reinforcement learning tools against attacks (Huang, Papernot, Goodfellow, Duan, & Abbeel, 2017). Adversarial machine learning (Barreno, Nelson, Sears, Joseph, & Tygar, 2006; Huang et al., 2011) must be developed considering ML's role in CPS and high impacts of CPS-security.
2. *Real-time threat-assessment*: An important ingredient associated any CPS-security tool is an accurate threat assessment. Determining the key indicators for this purpose, in real-time, is one of the key challenges that should be addressed. In this regard, network-theoretic tools may be highly relevant. Associated follow-on steps for detection, prevention, and resilience are of important interest as well.
3. *Scalability*: Extending all case studies reported to realistic systems with thousands of nodes and links are essential in order to make an impact on critical infrastructures. A central question here is the development of numerically scalable algorithms be developed for rapid detection, localization, and mitigation of attacks for such large scale networks?
4. *Resilience metric*: How can we define better resilience metrics for control systems? The relation between such a metric and standard control tools such as stability margins, controllability Gramians, $\mathcal{H}_2$ or $\mathcal{H}_\infty$ norms, need to be examined. The effect of network topology in resilience, especially in the context of cascading failures, needs to be explored.

We invite the entire systems and control community to engage in this important research topic.

Acknowledgement

We would like to thank Eugene Lavretsky for several useful discussions. The work reported here was supported in part by the National Science Foundation grant ECS 1054394 and in part by the Boeing Strategic University Initiative.

References

Abbas, W., Laszka, A., & Koutsoukos, X. (2018). Improving network connectivity and robustness using trusted nodes with application to resilient consensus. *IEEE Transactions on Control of Network Systems, 5*(4), 2036–2048.

Abdo, H., Kaouk, M., Flaus, J. M., & Masse, F. (2018). A safety/security risk analysis approach of industrial control systems: a cyber bowtie–combining new version of attack tree with bowtie analysis. *Computers & Security, 72*, 175–195.

Abhinav, S., Modares, H., Lewis, F. L., & Davoudi, A. (2019). Resilient Cooperative Control of DC Microgrids. *IEEE Transactions on Smart Grid, 10*(1), 1083–1085.

Accident to the Airbus A320-211, (2015) Registered D-AIPX and operated by Germanwings, flight GWI18G, on 03/24/15 at Prads-Haute-Blone, Available online at https://www.bea.aero/uploads/tx_elydbrapports/d-px150324.en.pdf (Accessed: 2018-08-3).

Adhikari, U., Morris, T., & Pan, S. (2017). WAMS cyber-physical test bed for power system, cybersecurity study, and data mining. *IEEE Transactions on Smart Grid, 8*(6), 2744–2753.

Ahmed, A., Bakar, K. A., Channa, M. I., Haseeb, K., & Khan, A. W. (2015). A survey on trust based detection and isolation of malicious nodes in ad-hoc and sensor networks. *Frontiers of Computer Science, 9*(2), 280–296.

Ahmed, E., & Gharavi, H. (2018). Cooperative vehicular networking: a survey. *IEEE Transactions on Intelligent Transportation Systems, 19*(3), 996–1014.

Alam, M., Ferreira, J., & Fonseca, J. (2016). Introduction to intelligent transportation systems. In M. Alam, J. Ferreira, & J. Fonseca (Eds.), *Intelligent transportation systems: Dependable vehicular communications for improved road safety* (pp. 1–17). Springer.

Algorand. (2018). Available online at https://www.algorand.com/ (Accessed: 2018-07-18).

Almalawi, A., Yu, X., Tari, Z., Fahad, A., & Khalil, I. (2014). An unsupervised anomaly-based detection approach for integrity attacks on SCADA systems. *Computers & Security, 46*, 94–110.

Amin, S., Cárdenas, A. A., & Sastry, S. S. (2009). Safe and secure networked control systems under denial-of-service attacks. In *Proceedings of international workshop on hybrid systems: Computation and control* (pp. 31–45).

Amin, S., Litrico, X., Sastry, S., & Bayen, A. M. (2013). Cyber security of water SCADA systems-part I: analysis and experimentation of stealthy deception attacks. *IEEE Transactions on Control Systems Technology, 21*(5), 1963–1970.

Amin, S., Schwartz, G. A., & Sastry, S. S. (2013). Security of interdependent and identical networked control systems. *Automatica, 49*(1), 186–192.

Amoozadeh, M., Raghuramu, A., Chuah, C.-N., Ghosal, D., Zhang, H. M., Rowe, J., & Levitt, K. (2015). Security vulnerabilities of connected vehicle streams and their impact on cooperative driving. *IEEE Communications Magazine, 53*(6), 126–132.

Andersson, G., Esfahani, P. M., Vrakopoulou, M., Margellos, K., Lygeros, J., Teixeira, A., et al. (2012). Cyber-security of SCADA systems. *IEEE PES Innovative Smart Grid Technologies*, 1–2.

Annaswamy, A. M., Malekpour, A. R., & Baros, S. (2016). Emerging research topics in control for smart infrastructures. *Annual Reviews in Control, 42*, 259–270.

Anwar, A., Mahmood, A. N., & Pickering, M. (2017). Modeling and performance evaluation of stealthy false data injection attacks on smart grid in the presence of corrupted measurements. *Journal of Computer and System Sciences, 83*(1), 58–72.

Ashok, A., Govindarasu, M., & Wang, J. (2017). Cyber-physical attack-resilient wide-area monitoring, protection, and control for the power grid. *Proceedings of the IEEE, 105*(7), 1389–1407.

Bai, C.-Z., Pasqualetti, F., & Gupta, V. (2015). Security in stochastic control systems: Fundamental limitations and performance bounds. In *Proceedings of American control conference* (pp. 195–200).

Bai, C.-z., Pasqualetti, F., & Gupta, V. (2017). Data-injection attacks in stochastic control systems: Detectability and performance tradeoffs. *Automatica, 82*, 251–260.

Barbosa, R. R. R., Sadre, R., & Pras, A. (2013). Flow whitelisting in SCADA networks. *International Journal of Critical Infrastructure Protection, 6*(3–4), 150–158.

Baros, S., Shiltz, D., Jaipuria, P., Hussain, A., & Annaswamy, A. M. (2017). *Towards resilient cyber-physical energy systems*. Available online at http://hdl.handle.net/1721.1/107408 (Accessed: 2019-05-21).

Barreno, M., Nelson, B., Sears, R., Joseph, A. D., & Tygar, J. D. (2006). Can machine learning be secure? In *Proceedings of ACM symposium on information, computer and communications security* (pp. 16–25).

Bi, S., & Zhang, Y. J. A. (2017). Graph-based cyber security analysis of state estimation in smart power grid. *IEEE Communications Magazine*, (99), 2–9.

Bishop, M. (2005). *Introduction to computer security*. Addison-Wesley.

Bobba, R. B., Dagle, J., Heine, E., Khurana, H., Sanders, W. H., Sauer, P., & Yardley, T. (2012). Enhancing grid measurements: Wide area measurement systems, NASPInet, and security. *IEEE Power Energy Magazine, 10*(1), 67–73.

Bradbury, D. (2012). SCADA: a critical vulnerability. *Computer Fraud & Security*, (4), 11–14.

Brown, H. E., & Demarco, C. L. (2018). Risk of cyber-physical attack via load with emulated inertia control. *IEEE Transactions on Smart Grid, 9*(6), 5854–5866.

Brown, P. N., Borowski, H., & Marden, J. R. (2019). Security against impersonation attacks in distributed systems. *IEEE Transactions on Control of Network Systems, 6*(1), 440–450.

Brown, P. N., & Marden, H. B. N. J. R. (2017). Are multiagent systems resilient to communication failures? arXiv:1711.00609.

Bruneau, M., Chang, S. E., Eguchi, R. T., Lee, G. C., O'Rourke, T. D., Reinhorn, A. M., et al. (2003). A framework to quantitatively assess and enhance the seismic resilience of communities. *Earthquake Spectra, 19*(4), 733–752.

Burmester, M., Magkos, E., & Chrissikopoulos, V. (2012). Modeling security in cyber—physical systems. *International Journal of Critical Infrastructure Protection, 5*(3–4), 118–126.

Cárdenas, A. A., Amin, S., Lin, Z.-S., Huang, Y.-L., Huang, C.-Y., & Sastry, S. (2011). Attacks against process control systems: risk assessment, detection, and response. In *Proceedings of ACM symposium on information, computer and communications security* (pp. 355–366).

Cárdenas, A. A., Amin, S., & Sastry, S. (2008). Research challenges for the security of control systems. In *Proceedings of USENIX workshop on hot topics in security (hotsec)*. Available online at https://people.eecs.berkeley.edu/~sastry/pubs/Pdfs%20of%202008/CardenasResearch2008.pdf (Accessed: 2019-05-21)

Cardenas, A. A., Amin, S., & Sastry, S. (2008). Secure control: Towards survivable cyber-physical systems. In *Proceedings of international conference on distributed computing systems workshops* (pp. 495–500).

Case, D. U. (2016). Analysis of the cyber attack on the Ukrainian power grid. *Electricity Information Sharing and Analysis Center (E-ISAC)*. Available onine at https://ics.sans.org/media/E-ISAC_SANS_Ukraine_DUC_5.pdf (Accessed: 2019-05-21).

Cetinkaya, A., Ishii, H., & Hayakawa, T. (2017). Networked control under random and malicious packet losses. *IEEE Transactions on Automatic Control, 62*(5), 2434–2449.

Cetinkaya, A., Ishii, H., & Hayakawa, T. (2019). An overview on Denial-of-Service attacks in control systems: Attack models and security analyses. *Entropy, 21*(210).

Chakhchoukh, Y., & Ishii, H. (2015). Coordinated cyber-attacks on the measurement function in hybrid state estimation. *IEEE Transactions on Power Systems, 30*(5), 2487–2497.

Chakhchoukh, Y., Vittal, V., Heydt, G. T., & Ishii, H. (2017). LTS-based robust hybrid SE integrating correlation. *IEEE Transactions on Power Systems, 32*(4), 3127–3135.

Chen, J., & Zhu, Q. (2015). *Resilient and decentralized control of multi-level cooperative mobile networks to maintain connectivity under adversarial environment*.

Chen, J., & Zhu, Q. (2017). Security as a service for cloud-enabled internet of controlled things under advanced persistent threats: a contract design approach. *IEEE Transactions on Information Forensics and Security, 12*(11), 2736–2750.

Chen, P., Desmet, L., & Huygens, C. (2014). A study on Advanced Persistent Threats. In B. De Decker, & A. Zúquete (Eds.), *Communications and multimedia security* (pp. 63–72). Springer.

Chen, T. (2010). Stuxnet, the real start of cyber warfare?[Editor's note]. *IEEE Network, 24*(6), 2–3.

Chen, Y., Kar, S., & Moura, J. M. F. (2017). Dynamic attack detection in cyber-physical systems with side initial state information. *IEEE Transactions on Automatic Control, 62*(9), 4618–4624.

Cherdantseva, Y., Burnap, P., Blyth, A., Eden, P., Jones, K., Soulsby, H., & Stoddart, K. (2016). A review of cyber security risk assessment methods for SCADA systems. *Computers & Security, 56*, 1–27.

Cherdantseva, Y., & Hilton, J. (2012). *The evolution of information security goals from the 1960s to today*. Available online at https://users.cs.cf.ac.uk/Y.V.Cherdantseva/LectureEvolutionInfoSecGOALS.pdf (Accessed: 2019-05-21).

Cherdantseva, Y., & Hilton, J. (2013). A reference model of information assurance & security. In *Proceedings of international conference on availability, reliability and security* (pp. 546–555).

Cho, H. S., & Woo, T. H. (2017). Cyber security in nuclear industry–Analytic study from the terror incident in nuclear power plants (NPPs). *Annals of Nuclear Energy, 99*, 47–53.

Chung, F. (1997). *Spectral graph theory*. American Mathematical Society.

Clark, A., Zhu, Q., Poovendran, R., & Başar, T. (2013). An impact-aware defense against Stuxnet. In *Proceedings of American control conference* (pp. 4140–4147).

Corts, J., Dullerud, G. E., Han, S., Ny, J. L., Mitra, S., & Pappas, G. J. (2016). Differential privacy in control and network systems. In *Proceedings of IEEE conference on decision and control* (pp. 4252–4272).

Cyber security and resilience of intelligent public transport: Good practices and recommendations. Technical Report. (2016) European Union Agency For Network And Information Security. Available online at https://www.enisa.europa.eu/publications/good-practices-recommendations/at_download/fullReport (Accessed: 2019-01-15).

Cybersecurity Funding 21. Cybersecurity Funding. (2019) Retrieved from https://www.whitehouse.gov/ (Accessed: 2019-01-15).

Darup, M. S., Redder, A., Shames, I., Farokhi, F., & Quevedo, D. (2018). Towards encrypted MPC for linear constrained systems. *IEEE Control Systems Letters, 2*(2), 195–200.

Data Breach Investigations Report (2009). Technical Report, Retrieved from https://www.verizonenterprise.com/ (Accessed: 2019-01-15).

Data Breach Investigations Report. (2015). Technical Report, Retrieved from https://www.verizonenterprise.com/ (Accessed: 2019-01-15).

De Persis, C., & Isidori, A. (2001). A geometric approach to nonlinear fault detection and isolation. *IEEE Transactions on Automatic Control, 46*(6), 853–865.

De Persis, C., & Tesi, P. (2014). On resilient control of nonlinear systems under denial-of-service. In *Proceedings of IEEE conference on decision and control* (pp. 5254–5259).

De Persis, C., & Tesi, P. (2015). Input-to-state stabilizing control under Denial-of-Service. *IEEE Transactions on Automatic Control, 60*(11), 2930–2944.

De Persis, C., & Tesi, P. (2018). A comparison among deterministic packet–dropouts models in networked control systems. *IEEE Control Systems Letters, 2*(1), 109–114.

Dibaji, S. M., & Ishii, H. (2017). Resilient consensus of second-order agent networks: Asynchronous update rules with delays. *Automatica, 81*, 123–132.

Dibaji, S. M., Ishii, H., & Tempo, R. (2018). Resilient randomized quantized consensus. *IEEE Transactions on Automatic Control, 63*(8), 2508–2522.

Dibaji, S. M., Pirani, M., Annaswamy, A. M., Johansso, K. H., & Chakrabortty, A. (2018). Secure control of power systems: Confidentiality and integrity threats. In *Proceedings of IEEE conference on decision and control* (pp. 7269–7274).

Dibaji, S. M., Safi, M., & Ishii, H. (2019). Resilient distributed averaging. In *Proceedings of American control conference.* in press.

Diffie, W., & Hellman, M. (1976). New directions in cryptography. *IEEE Transactions on Information Theory, 22*(6), 644–654.

Ding, D., Han, Q.-L., Xiang, Y., Ge, X., & Zhang, X.-M. (2018). A survey on security control and attack detection for industrial cyber-physical systems. *Neurocomputing, 275*, 1674–1683.

Ding, K., Li, Y., Quevedo, D. E., Dey, S., & Shi, L. (2017). A multi-channel transmission schedule for remote state estimation under DoS attacks. *Automatica, 78*, 194–201.

Ding, K., Ren, X., & Shi, L. (2016). Deception-based sensor scheduling for remote estimation under DoS attacks. In *Proceedings of IFAC workshop on distributed estimation and control in networked systems* (pp. 169–174).

Do, V. L., Fillatre, L., & Nikiforov, I. (2017). *Proceedings of IFAC symposium on fault detection, supervision and safety for technical processes* (21, pp. 746–753).

du Pin Calmon, F., & Fawaz, N. (2012). Privacy against statistical inference. In *Proceedings of annual allerton conference on communication, control, and computing* (pp. 1401–1408).

Dwork, C. (2011). Differential privacy. In *Encyclopedia of cryptography and security* (pp. 338–340). Springer.

Fagnant, D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice, 77*, 167–181.

Falliere, N., Murchu, L. O., & Chien, E. (2018). W32. Stuxnet dossier: Symantec security response. Technical Report, Symantec, Available online at https://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf (Accessed: 2018-08-05).

Farokhi, F., Shames, I., & Batterham, N. (2017). Secure and private control using semi-homomorphic encryption. *Control Engineering Practice, 67*, 13–20.

Farokhi, F., Shames, I., & Johansson, K. H. (2017). Private and secure coordination of match-making for heavy-duty vehicle platooning. In *Proceedings of IFAC world congress* (pp. 7345–7350).

Fawzi, H., Tabuada, P., & Diggavi, S. (2012). Security for control systems under sensor and actuator attacks. *Proceedings of IEEE conference on decision and control*, 3412–3417.

Fawzi, H., Tabuada, P., & Diggavi, S. (2014). Secure estimation and control for cyber-physical systems under adversarial attacks. *IEEE Transactions on Automatic Control, 59*(6), 1454–1467.

Felegyhazi, M., & Hubaux, J.-P. (2006). Game theory in wireless networks: A tutorial. *Technical Report*. Available online at https://www.152.66.249.135/~mfelegyhazi/publications/FelegyhaziH06tutorial.pdf (Accessed: 2019-05-21).

Ferdowsi, A., Saad, W., & Mandayam, N. B. (2017). Colonel Blotto game for secure state estimation in interdependent critical infrastructure. arXiv:1709.09768.

Ferrari, R. M., & Teixeira, A. M. (2017a). Detection and isolation of routing attacks through sensor watermarking. In *Proceedings of American control conference* (pp. 5436–5442).

Ferrari, R. M. G., & Teixeira, A. M. H. (2017b). Detection and isolation of replay attacks through sensor watermarking. In *Proceedings of IFAC workshop on distributed estimation and control in networked systems* (pp. 7363–7368).

Flamholz, D. B., Annaswamy, A. M., & Lavretsky, E. (2019). Baiting for defense against stealthy attacks on cyber-physical systems. In *Proceedings of AIAA Scitech Forum* (p. 2338).

Fovino, I. N., Carcano, A., Masera, M., & Trombetta, A. (2009). An experimental investigation of malware attacks on scada systems. *International Journal of Critical Infrastructure Protection, 2*(4), 139–145.

Frasca, P., Ishii, H., Ravazzi, C., & Tempo, R. (2015). Distributed randomized algorithms for opinion formation, centrality computation and power systems estimation: A tutorial overview. *European Journal of Control, 24*, 2–13.

Garcia-Teodoro, P., Diaz-Verdejo, J., Maciá-Fernández, G., & Vázquez, E. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. *Computers & Security, 28*(1–2), 18–28.

GhasemiGol, M., Ghaemi-Bafghi, A., & Takabi, H. (2016). A comprehensive approach for network attack forecasting. *Computers & Security, 58*, 83–105.

Gil, S., Kumar, S., Mazumder, M., Katabi, D., & Rus, D. (2017). Guaranteeing spoof-resilient multi-robot networks. *Autonomous Robots, 41*(6), 1383–1400.

Giraldo, J., Urbina, D., Cardenas, A., Valente, J., Faisal, M., Ruths, J., et al. (2018). A survey of physics-based attack detection in cyber-physical systems. *ACM Computing Surveys, 51*(4), Artilce76.

Góes, R. M., Kang, E., Kwong, R., & Lafortune, S. (2017). Stealthy deception attacks for cyber-physical systems. In *Proceedings of IEEE conference on decision and control* (pp. 4224–4230).

Gomez-Exposito, A., & Abur, A. (2004). *Power system state estimation: Theory and implementation*. CRC press.

Greenberg, A. (2018). *Hackers remotely kill a jeep on the highway with me in it*. Accessed: 2018-12-16. https://www.wired.com/2015/07/hackers-remotely-kill-jeep-highway/.

Greenberg, A. (2017). *How an entire nation became Russia's test lab for cyberwar*. Accessed: 2019-05-02. https://www.wired.com/story/russian-hackers-attack-ukraine/.

Gross, J. L., & Yellen, J. (2005). *Graph theory and its applications*. CRC press.

Guerrero-Bonilla, L., Prorok, A., & Kumar, V. (2017). Formations for resilient robot teams. *IEEE Robotics and Automation Letters, 2*(2), 841–848.

Guo, Z., Shi, D., Johansson, K. H., & Shi, L. (2017). Optimal linear cyber-attack on remote state estimation. *IEEE Transactions on Control of Network Systems, 4*(1), 4–13.

Guo, Z., Shi, D., Johansson, K. H., & Shi, L. (2018). Worst-case stealthy innovation-based linear attack on remote state estimation. *Automatica, 89*, 117–124.

Gupta, A., Langbort, C., & Başar, T. (2010). Optimal control in the presence of an intelligent jammer with limited actions. In *Proceedings of IEEE conference on decision and control* (pp. 1096–1101).

Gupta, N., Katz, J., & Chopra, N. (2017). Privacy in distributed average consensus. In *Proceedings of world congress of the international federation of automatic control* (pp. 9515–9520).

Gusrialdi, A., & Qu, Z. (2019). Smart grid security: Attacks and defenses. In J. Stoustrup, A. M. Annaswamy, A. Chakrabortty, & Z. Qu (Eds.), *Smart grid control* (pp. 199–223). Springer.

Hackmageddon: Information security timelines and statistics. Accessed: 2018-08-3. https://www.hackmageddon.com/.

Harding, J., Powell, G., Yoon, R., Fikentscher, J., Doyle, C., Sade, D., et al. (2014). Vehicle-to-vehicle communications: Readiness of V2V technology for application. *Technical report*. United States. National Highway Traffic Safety Administration, https://rosap.ntl.bts.gov/view/dot/27999, Accessed time (05-03-2019).

Hartmann, K., & Steup, C. (2013). The vulnerability of UAVs to cyber attacks-An approach to the risk assessment. In *Proceedings of international conference on cyber conflict* (pp. 1–23).

Haus, T., Palunko, I., Tolić, D., Bogdan, S., Lewis, F. L., & Mikulski, D. G. (2014). Trust-based self-organising network control. *IET Control Theory & Applications, 8*(18), 2126–2135.

He, Y., Mendis, G. J., & Wei, J. (2016). Real-time detection of false data injection attacks in smart grids: A deep learning-based intelligent mechanism. *IEEE Transactions on Smart Grid, 8*(5), 1–12.

Heemels, W., Johansson, K. H., & Tabuada, P. (2012). An introduction to event-triggered and self-triggered control. In *Proceedings of IEEE conference on decision and control* (pp. 3270–3285).

Hein, M., & Andriushchenko, M. (2017). Formal guarantees on the robustness of a classifier against adversarial manipulation. In *Advances in neural information processing systems* (pp. 2266–2276).

Hirzallah, N. H., & Voulgaris, P. G. (2018). On the computation of worst attacks: a LP framework. In *Proceedings of American control conference* (pp. 4527–4532).

Hoehn, A., & Zhang, P. (2016). Detection of covert attacks and zero dynamics attacks in cyber-physical systems. In *Proceedings of American control conference* (pp. 302–307).

Hoory, S., Linial, N., & Wigderson, A. (2006). Expander graphs and their applications. *Bulletin of the American Mathematical Society, 43*, 439–561.

Hoppe, T., Kiltz, S., & Dittmann, J. (2008). Security threats to automotive can networks–practical examples and selected short-term countermeasures. In *Proceedings of international conference on computer safety, reliability, and security* (pp. 235–248). Springer.

Horák, K., Zhu, Q., & Bošanskỳ, B. (2017). Manipulating adversarys belief: A dynamic game approach to deception by design for proactive network security. In *Proceedings of international conference on decision and game theory for security* (pp. 273–294).

Hu, L., Wang, Z., Han, Q.-L., & Liu, X. (2018). State estimation under false data injection attacks: Security analysis and system protection. *Automatica, 87*, 176–183.

Huang, L., Joseph, A. D., Nelson, B., Rubinstein, B. I., & Tygar, J. (2011). Adversarial machine learning. In *Proceedings of ACM workshop on security and artificial intelligence* (pp. 43–58).

Huang, S., Papernot, N., Goodfellow, I., Duan, Y., & Abbeel, P. (2017). Adversarial attacks on neural network policies. arXiv:1702.02284.

Huang, T., Satchidanandan, B., Kumar, P. R, & Xie, L. (2018). An Online Detection Framework for Cyber Attacks on Automatic Generation Control. *IEEE Transactions on Power Systems, 33*(6), 6816–6827.

Huitsing, P., Chandia, R., Papa, M., & Shenoi, S. (2008). Attack taxonomies for the modbus protocols. *International Journal of Critical Infrastructure Protection, 1*, 37–44.

Humayed, A., Lin, J., Li, F., & Luo, B. (2017). Cyber-physical systems securitya survey. *IEEE Internet of Things Journal, 4*(6), 1802–1831.

Jafarnejadsani, H., Lee, H., Hovakimyan, N., & Voulgaris, P. (2017). Dual-rate ℓ_1 adaptive controller for cyber-physical sampled-data systems. In *Proceedings of IEEE conference on decision and control* (pp. 6259–6264).

Jafarnia-Jahromi, A., Broumandan, A., Nielsen, J., & Lachapelle, G. (2012). GPS vulnerability to spoofing threats and a review of antispoofing techniques. *International Journal of Navigation and Observation*, 1–16.

Jajodia, S., Ghosh, A. K., Swarup, V., Wang, C., & Wang, X. S. (2011). *Moving target defense: Creating asymmetric uncertainty for cyber threats*: Vol. 54. Springer.

Javed, M. A., & Hamida, E. B. (2017). On the interrelation of security, QoS, and safety in cooperative ITS. *IEEE Transactions on Intelligent Transportation Systems, 18*(7), 1943–1957.

Jiang, T., & Baras, J. S. (2006). Trust evaluation in anarchy: A case study on autonomous networks. In *Proceedings of IEEE international conference on computer communications* (pp. 1–12).

Kailkhura, B., Han, Y. S., Brahma, S., & Varshney, P. K. (2013). Distributed Bayesian detection with Byzantine data. arXiv:1307.3544.

Katz, J., Menezes, A. J., Van Oorschot, P. C., & Vanstone, S. A. (1996). *Handbook of applied cryptography*. CRC Press.

Kern, C., Kesavan, A., & Daswani, N. (2007). *Foundations of Security: What every programmer needs to know*. Springer.

Khan, U. A., & Stanković, A. M. (2013). Secure distributed estimation in cyber-physical systems. In *Proceedings of IEEE international conference on acoustics, speech and signal processing* (pp. 5209–5213).

Khanafer, A., Touri, B., & Baar, T. (2013). *Robust distributed averaging in networks* (pp. 1–23).

Khargonekar, P. P. (August 2015). Enabling research for infrastructure resilience: An NSF perspective. *National science foundation-resilience week*. Available online at https://www.faculty.sites.uci.edu/khargonekar/files/2017/06/Resilient_Infrastructures_2015.pdf.

Khazraei, A., Kebriaei, H., & Salmasi, F. R. (2017). Replay attack detection in a multi agent system using stability analysis and loss effective watermarking. In *Proceedings of American control conference* (pp. 4778–4783).

Kikuya, Y., Dibaji, S. M., & Ishii, H. (2018). Fault-tolerant clock synchronization over unreliable channels in wireless sensor networks. *IEEE Transactions on Control of Network Systems, 5*(4), 1551–1562.

Kim, J., Park, G., Shim, H., & Eun, Y. (2016). Zero-stealthy attack for sampled-data control systems: The case of faster actuation than sensing. In *Proceedings of IEEE conference on decision and control* (pp. 5956–5961).

Ko, W.-H., Satchidanandan, B., & Kumar, P. (2016). Theory and implementation of dynamic watermarking for cybersecurity of advanced transportation systems. In *Proceedings of IEEE conference on communications and network security* (pp. 416–420).

Koscher, K., Czeskis, A., Roesner, F., Patel, S., Kohno, T., Checkoway, S., et al. (2010). Experimental security analysis of a modern automobile. *Proceedings of IEEE symposium on security and privacy*, 447–462.

Krotofil, M., Cardenas, A., Larsen, J., & Gollmann, D. (2014). Vulnerabilities of cyber–physical systems to stale data determining the optimal time to launch attacks. *International Journal of Critical Infrastructure Protection, 7*(4), 213–232.

Kurmanaev, A., & Herrera, I. (2019). *No end in sight to Venezuelas Blackout, experts warn*. Available online at https://www.nytimes.com/2019/03/11/world/americas/venzuela-blackout-maduro.html (Accessed: 2019-04-05).

Kwon, C., Liu, W., & Hwang, I. (2014). Analysis and design of stealthy cyber attacks on unmanned aerial systems. *Journal of Aerospace Information Systems, 11*(8), 525–539.

La, R. J. (2017). Estimation of externalities in interdependent security: A case study of large systems. In *Proceedings of IEEE conference on decision and control* (pp. 3961–3966).

Lamnabhi-Lagarrigue, F., Annaswamy, A., Engell, S., Isaksson, A., Khargonekar, P., Murray, R. M., et al. (2017). Systems & control for the future of humanity, research agenda: Current and future roles, impact and grand challenges. *Annual Reviews in Control, 43*, 1–64.

LeBlanc, H. J., & Koutsoukos, X. (2018). Resilient first-order consensus and weakly stable, higher order synchronization of continuous-time networked multiagent systems. *IEEE Transactions on Control of Network Systems, 5*(3), 1219–1231.

LeBlanc, H. J., Zhang, H., Koutsoukos, X., & Sundaram, S. (2013). Resilient asymptotic consensus in robust networks. *IEEE Journal on Selected Areas in Communications, 31*, 766–781.

Lei, L., Yang, W., & Yang, C. (2016). Event-based distributed state estimation over a WSN with false data injection attack. In *Proceedings of IFAC workshop on distributed estimation and control in networked systems* (pp. 286–290).

Li, B., Lu, R., Wang, W., & Choo, K.-K. R. (2017). Distributed host-based collaborative detection for false data injection attacks in smart grid cyber-physical system. *Journal of Parallel and Distributed Computing, 103*, 32–41.

Li, Z., Shahidehpour, M., & Aminifar, F. (2017). Cybersecurity in distributed power systems. *Proceedings of the IEEE, 105*(7), 1367–1388.

Liang, G., Zhao, J., Luo, F., Weller, S., & Dong, Z. (2017). A review of false data injection attacks against modern power systems. *IEEE Transactions on Smart Grid, 8*(4), 1630–1638.

Liao, M., & Chakrabortty, A. (2018). Optimization algorithms for catching data manipulators in power system estimation loops. *IEEE Transactions on Control Systems Technology*, (99), 1–16.

Lippmann, R. P., & Cunningham, R. K. (2000). Improving intrusion detection performance using keyword selection and neural networks. *Computer Networks, 34*(4), 597–603.

Liu, S., Xu, B., Li, S., & Liu, Y. (2017). Resilient control strategy of cyber-physical system under DoS attacks. In *Proceedings of Chinese control conference* (pp. 7760–7765).

Liu, X., & Li, Z. (2017). False data attack models, impact analyses and defense strategies in the electricity grid. *The Electricity Journal, 30*(4), 35–42.

Liu, Y., Ning, P., & Reiter, M. K. (2009). *False data injection attacks against state estimation in electric power grids* (14, pp. 21–32).

Lu, Y., Chang, C.-Y., Zhang, W., Marinovici, L. D., & Conejo, A. J. (2016). On resilience analysis and quantification for wide-area control of power systems. In *Proceedings of IEEE conference on decision and control* (pp. 5799–5804).

Lucia, W., Sinopoli, B., & Franze, G. (2016). Networked constrained cyber-physical systems subject to malicious attacks: A resilient set-theoretic control approach. arXiv:1603.07984.

Lynch, N. A. (1996). *Distributed algorithms*. Morgan Kaufmann.

Macwan, R., Drew, C., Panumpabi, P., Valdes, A., Vaidya, N., Sauer, P., & Ishchenko, D. (2016). Collaborative defense against data injection attack in iec61850 based smart substations. In *Proceedings of IEEE power and energy society general meeting* (pp. 1–5).

Magiera, J., & Katulski, R. (2015). Detection and mitigation of gps spoofing based on antenna array processing. *Journal of Applied Research and Technology, 13*(1), 45–57.

Manadhata, P. K., & Wing, J. M. (2011). An attack surface metric. *IEEE Transactions on Software Engineering, 37*(3), 371–386.

Manshaei, M., Zhu, Q., Alpcan, T., Basar, T., & Hubaux, J. P. (2013). Game theory meets network security and privacy. *ACM Computing Surveys, 45*, 53–73.

Marden, J. R., Arslan, G., & Shamma, J. S. (2009). Cooperative control and potential games. *IEEE Transactions on Systems, Man, and Cybernetics, 39*(6), 1393–1407.

Massoumnia, M.-A., Verghese, G. C., & Willsky, A. S. (1989). Failure detection and identification. *IEEE Transactions on Automatic Control, 34*(3), 316–321.

McLaughlin, S., Konstantinou, C., Wang, X., Davi, L., Sadeghi, A. R., Maniatakos, M., & Karri, R. (2016). The cybersecurity landscape in industrial control systems. *Proceedings of the IEEE, 104*(5), 1039–1057.

Miao, F., Zhu, Q., Pajic, M., & Pappas, G. J. (2018). A hybrid stochastic game for secure control of cyber-physical systems. *Automatica, 93*, 55–63.

Mikulski, D. G., Lewis, F. L., Gu, E. Y., & Hudas, G. R. (2011). Trust dynamics in multi-agent coalition formation. In *Proceedings of unmanned systems technology*.

Mikulski, D. G., Lewis, F. L., Gu, E. Y., & Hudas, G. R. (2014). Trust-based coalition formation in multi-agent systems. *The Journal of Defense Modeling and Simulation, 11*(1), 19–32.

Milanese, M. (2013). *Robustness in identification and control*. Springer.

Milani, A. A., Khan, M. T. A., Chakrabortty, A., & Husain, I. (2018). Equilibrium point analysis and power sharing methods for distribution systems driven by solid-state transformers. *IEEE Transactions on Power Systems, 33*(2), 1473–1483.

Miloševič, J., Tanaka, T., Sandberg, H., & Johansson, K. H. (2017). Analysis and mitigation of bias injection attacks against a Kalman filter. In *Proceedings of IFAC world congress* (pp. 8393–8398).

Mishra, S., Shoukry, Y., Karamchandani, N., Diggavi, S. N., & Tabuada, P. (2017). Secure state estimation against sensor attacks in the presence of noise. *IEEE Transactions on Control of Network Systems, 4*(1), 49–59.

Mitra, A., & Sundaram, S. (2018). Distributed observers for LTI systems. *IEEE Transactions on Automatic Control, 63*(11), 3689–3704.

Mo, Y., Chabukswar, R., & Sinopoli, B. (2014a). Detecting integrity attacks on scada systems. *IEEE Transactions on Control Systems Technology, 22*(4), 1396–1407.

Mo, Y., Hespanha, J. P., & Sinopoli, B. (2014b). Resilient detection in the presence of integrity attacks. *IEEE Transactions on Signal Processing, 62*(1), 31–43.

Mo, Y., & Murray, R. M. (2017). Privacy preserving average consensus. *IEEE Transactions on Automatic Control, 62*(2), 753–765.

Mo, Y., Weerakkody, S., & Sinopoli, B. (2015). Physical authentication of control systems: Designing watermarked control inputs to detect counterfeit sensor outputs. *IEEE Control Systems Magazine, 35*(1), 93–109.

Moghadam, R., & Modares, H. (2017). Attack analysis for distributed control systems: An internal model principle approach. arXiv:1710.03856.

Momani, M., & Challa, S. (2010). Survey of trust models in different network domains. arXiv:1010.0168.

Motwani, R., & Raghavan, P. (2010). *Randomized algorithms*. Chapman & Hall/CRC.

Murguia, C., van de Wouw, N., & Ruths, J. (2017). Reachable sets of hidden CPS sensor attacks: Analysis and synthesis tools. In *Proceedings of IFAC world congress*.

Naghnaeian, M., Hirzallah, N., & Voulgaris, P. G. (2015). Dual rate control for security in cyber-physical systems. arXiv:1504.07586.

Ng, A. (2018). *Machine Learning*. Stanford University lecture notes

Nicol, D. M., Sanders, W. H., & Trivedi, K. S. (2004). Model-based evaluation: from dependability to security. *IEEE Transactions on Dependable and Secure Computing, 1*(1), 48–65.

Nordell, D. E. (2012). Terms of protection: The many faces of smart grid security. *IEEE Power and Energy Magazine, 10*(1), 18–23.

Nozari, E., Tallapragada, P., & Cortés, J. (2017). Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design. *Automatica, 81*, 221–231.

Ntalampiras, S. (2016). Automatic identification of integrity attacks in cyber-physical systems. *Expert Systems with Applications, 58*, 164–173.

Nudell, S., Thomas, R., Nabavi, & Chakrabortty, A. (2015). A real-time attack localization algorithm for large power system networks using graph-theoretic techniques. *IEEE Transactions on Smart Grid, 6*(5), 2551–2559.

Onyeji, I., Bazilian, M., & Bronk, C. (2014). Cyber security and critical energy infrastructure. *The Electricity Journal, 27*(2), 52–60.

Pajic, M., Lee, I., & Pappas, G. J. (2017). Attack-resilient state estimation for noisy dynamical systems. *IEEE Transactions on Control of Network Systems, 4*(1), 82–92.

Pal, S., Sikdar, B., & Chow, J. H. (2018). Classification and detection of PMU data manipulation attacks using transmission line parameters. *IEEE Transactions on Smart Grid, 9*(5), 5057–5066.

Pan, Z., & Başar, T. (1999). $\mathcal{H}$–infinity control of large scale jump linear systems via averaging and aggregation. *International Journal of Control, 72*, 866–881.

Parkinson, S., Ward, P., Wilson, K., & Miller, J. (2017). Cyber threats facing autonomous and connected vehicles: Future challenges. *IEEE Transactions on Intelligent Transportation Systems, 18*(11), 2898–2915.

Pasqualetti, F., Bicchi, A., & Bullo, F. (2012). Consensus computation in unreliable networks: A system theoretic approach. *IEEE Transactions on Automatic Control, 57*(1), 90–104.

Pasqualetti, F., Dörfler, F., & Bullo, F. (2013). Attack detection and identification in cyber-physical systems. In *IEEE Transactions on Automatic Control: 58* (pp. 2715–2729).

Pasqualetti, F., Dorfler, F., & Bullo, F. (2015a). Control-theoretic methods for cyber-physical security: Geometric principles for optimal cross-layer resilient control systems. *IEEE Control Systems Magazine, 35*(1), 110–127.

Pasqualetti, F., Dörfler, F., & Bullo, F. (2015b). A divide-and-conquer approach to distributed attack identification. In *Proceedings of IEEE conference on decision and control* (pp. 5801–5807).

Pawlick, J., Farhang, S., & Zhu, Q. (2015). Flip the cloud: Cyber-physical signaling games in the presence of advanced persistent threats. In *Proceedings of international conference on decision and game theory for security* (pp. 289–308).

Petit, J., & Shladover, S. E. (2015). Potential cyberattacks on automated vehicles. *IEEE Transactions on Intelligent Transportation Systems, 16*(2), 546–556.

Pirani, M., & Sundaram, S. (2016). On the smallest eigenvalue of grounded Laplacian matrices. *IEEE Transaction on Automatic Control, 61*(2), 509–514.

Pramod, T., & Sunitha, N. (2015). Polynomial based scheme for secure SCADA operations. *Procedia Technology, 21*, 474–481.

Presidential Policy Directive 21: Critical Infrastructure Security and Resilience. Retrieved from https://obamawhitehouse.archives.gov.

Rani, C., Modares, H., Sriram, R., Mikulski, D., & Lewis, F. L. (2016). Security of unmanned aerial vehicle systems against cyber-physical attacks. *The Journal of Defense Modeling and Simulation, 13*(3), 331–342.

Reddy, E. K. (2013). Neural networks for intrusion detection and its applications. In *Proceedings of the world congress on engineering: Vol. 2* (pp. 3–5).

Rezai, A., Keshavarzi, P., & Moravej, Z. (2013). Secure SCADA communication by using a modified key management scheme. *ISA Transactions, 52*(4), 517–524.

Rieger, C. G., Gertman, D. I., & McQueen, M. A. (2009). Resilient control systems: Next generation design research. In *Proceedings of conference on human system interactions* (pp. 632–636).

Rus, D., Shrobe, H., Chlipala, A., Clark, D., Devadas, S., Goldwasser, S., et al. (2018). *Cybersecurity: Technology, application and policy*. lecture notes, MIT xPRO

Shitharth, S., & Prince Winston, D. (2017). An enhanced optimization based algorithm for intrusion detection in SCADA network. *Computers & Security, 70*, 16–26.

Saltzer, J. H., & Schroeder, M. D. (1975). The protection of information in computer systems. *Proceedings of the IEEE, 63*(9), 1278–1308.

Sandberg, H., Johansson, K. H., Amin, S., & (guest editors) (2015). Special issue on cyberphysical security in networked control systems. *IEEE Control Systems Magazine, 35*(1).

Sandberg, H., Teixeira, A., & Johansson, K. H. (2010). On security indices for state estimators in power networks. In *Proceedings of the first workshop on secure control systems*.

Sandberg, H., & Teixeira, A. M. (2016). From control system security indices to attack identifiability. In *Proceedings of science of security for cyber-physical systems workshop* (pp. 1–6).

Sanjab, A., & Saad, W. (2016). On bounded rationality in cyber-physical systems security: Game-theoretic analysis with application to smart grid protection. *Proceedings of Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids*.

Sanjab, A., Saad, W., & Başar, T. (2017). Prospect theory for enhanced cyber-physical security of drone delivery systems: A network interdiction game. arXiv:1702.04240.

Sanjab, A., Saad, W., Guvenc, I., Sarwat, A., & Biswas, S. (2016). *Smart grid security: Threats, challenges, and solutions*. arXiv preprint arXiv: 1606.06992.

Satchidanandan, B., & Kumar, P. R. (2017). Dynamic watermarking: Active defense of networked cyber–physical systems. *Proceedings of the IEEE, 105*(2), 219–240.

Satchidanandan, B., & Kumar, P. R. (2018a). Control systems under attack: The securable and unsecurable subspaces of a linear stochastic system. *Emerging Applications of Control and Systems Theory*, 217–228.

Satchidanandan, B., & Kumar, P. R. (2018b). On the operational significance of the securable subspace for partially observed linear stochastic systems. *Proceedings of IEEE conference on decision and control*, 2068–2073.

Shahrivar, E. M., Pirani, M., & Sundaram, S. (2017). Spectral and structural properties of random interdependent networks. *Automatica, 83*, 234–242.

Shannon, C. E. (1949). Communication theory of secrecy systems. *Bell System Technical Journal, 28*(4), 656–715.

Sherif, A. B., Rabieh, K., Mahmoud, M. M., & Liang, X. (2017). Privacy-preserving ride sharing scheme for autonomous vehicles in big data era. *IEEE Internet of Things Journal, 4*(2), 611–618.

Shoukry, Y., Chong, M., Wakaiki, M., Nuzzo, P., Sangiovanni-Vincentelli, A., Seshia, S. A., et al. (2018). SMT-based observer design for cyber-physical systems under sensor attacks. *ACM Transactions on Cyber-Physical Systems, 2*(1), 5.

Siegel, J. E., Erb, D. C., & Sarma, S. E. (2018). A survey of the connected vehicle landscape architectures, enabling technologies, applications, and development areas. *IEEE Transactions on Intelligent Transportation Systems, 19*(8), 2391–2406.

Slay, J., & Miller, M. (2007). Lessons learned from the Maroochy water breach. In *Proceedings of international conference on critical infrastructure protection* (pp. 73–82).

Sniper attack on California power grid may have been ‘an insider,’ DHS says. (2015), Availble online at https://money.cnn.com/2015/10/16/technology/sniper-power-grid/index.html (Accessed: 2018-08-05).

Sridhar, S., Hahn, A., & Govindarasu, M. (2012). Cyber-physical system security for the electric power grid. *Proceedings of the IEEE, 100*(1), 210–224.

Stoustrup, J., Annaswamy, A., Chakrabortty, A., & Qu (ed.), Z. (2019). *Smart grid control: An overview and research opportunities*. Springer.

Sucasas, V., Mantas, G., Saghezchi, F. B., Radwan, A., & Rodriguez, J. (2016). An autonomous privacy-preserving authentication scheme for intelligent transportation systems. *Computers & Security, 60*, 193–205.

Sun, H., Peng, C., Yang, T., Zhang, H., & He, W. (2017). Resilient control of networked control systems with stochastic denial of service attacks. *Neurocomputing, 270*, 170–177.

Sun, H., Peng, C., Zhang, W., Yang, T., & Wang, Z. (2018). Security-based resilient event-triggered control of networked control systems under denial of service attacks. *Journal of the Franklin Institute*. In press.

Sundaram, S., & Gharesifard, B. (2016). Secure local filtering algorithms for distributed optimization. In *Proceedings of IEEE conference on decision and control* (pp. 1871–1876).

Sundaram, S., & Hadjicostis, C. N. (2011). Distributed function calculation via linear iterative strategies in the presence of malicious agents. *IEEE Transactions on Automatic Control, 56*(7), 1495–1508.

Tanaka, T., Skoglund, M., Sandberg, H., & Johansson, K. H. (2017). Directed information and privacy loss in cloud-based control. In *Proceedings of American Control Conference* (pp. 1666–1672).

Teixeira, A., Amin, S., Sandberg, H., Johansson, K. H., & Sastry, S. S. (2010). Cyber security analysis of state estimators in electric power systems. In *Proceedings of IEEE conference on decision and control* (pp. 5991–5998).

Teixeira, A., Shames, I., Sandberg, H., & Johansson, K. H. (2012). Revealing stealthy attacks in control systems. In *Proceedings of the annual allerton conference on communication, control, and computing* (pp. 1806–1813).

Teixeira, A., Shames, I., Sandberg, H., & Johansson, K. H. (2015). A secure control framework for resource-limited adversaries. *Automatica, 51*, 135–148.

Teixeira, A., Sou, K. C., Sandberg, H., & Johansson, K. H. (2015). Secure control systems: A quantitative risk management approach. *IEEE Control Systems Magazine, 35*(1), 24–45.

Ten, C.-W., Hong, J., & Liu, C.-C. (2011). Anomaly detection for cybersecurity of the substations. *IEEE Transactions on Smart Grid, 2*(4), 865–873.

Ten, C. W., Liu, C. C., & Manimaran, G. (2008). Vulnerability assessment of cybersecurity for SCADA systems. *IEEE Transactions on Power Systems, 23*(4), 1836–1846.

Tranquillus, G. S. (1957). *The twelve caesars*. Penguin Books London.

Tsai, C.-F., Hsu, Y.-F., Lin, C.-Y., & Lin, W.-Y. (2009). Intrusion detection by machine learning: A review. *Expert Systems with Applications, 36*(10), 11994–12000.

Ugrinovskii, V., & Langbort, C. (2017). Controller–jammer game models of denial of service in control systems operating over packet-dropping links. *Automatica, 84*, 128–141.

Usevitch, J., & Panagou, D. (2018a). Determining r-robustness of arbitrary digraphs using zero-one linear integer programming. arXiv:1810.01784.

Usevitch, J., & Panagou, D. (2018b). Resilient leader-follower consensus to arbitrary reference values. arXiv:1802.09654.

Vamvoudakis, K. G., & Hespanha, J. P. (2018a). Cooperative Q-learning for rejection of persistent adversarial inputs in unknown networked systems. *IEEE Transactions on Automatic Control, 63*(4), 1018–1031.

Vamvoudakis, K. G., & Hespanha, J. P. (2018b). Game-theory-based consensus learning of double-integrator agents in the presence of worst-case adversaries. *Journal of Optimization Theory and Applications, 177*(1), 222–253.

Vamvoudakis, K. G., Hespanha, J. P., Sinopoli, B., & Mo, Y. (2014). Detection in adversarial environments. *IEEE Transactions on Automatic Control, 59*(12), 3209–3223.

Wang, D., Wang, Z., Shen, B., Alsaadi, F. E., & Hayat, T. (2016). Recent advances on filtering and control for cyber-physical systems under security and resource constraints. *Journal of the Franklin Institute, 353*(11), 2451–2466.

Wang, W., & Lu, Z. (2013). Cyber security in the smart grid: Survey and challenges. *Computer Networks, 57*(5), 1344–1371.

Watts, D. (2003). Security and vulnerability in electric power systems. In *Proceedings of the north American power symposium: 2* (pp. 559–566).

Weerakkody, S., Liu, X., Son, S. H., & Sinopoli, B. (2017). A graph-theoretic characterization of perfect attackability for secure design of distributed control systems. *IEEE Transactions on Control of Nerwork Systems, 4*, 60–70.

Weerakkody, S., Mo, Y., & Sinopoli, B. (2014). *Detecting integrity attacks on control systems using robust physical watermarking* (pp. 3757–3764).

Weerakkody, S., Ozel, O., & Sinopoli, B. (2017). A Bernoulli-Gaussian physical watermark for detecting integrity attacks in control systems. In *Proceedings of annual allerton conference on communication, control, and computing* (pp. 966–973).

Weerakkody, S., & Sinopoli, B. (2015). Detecting integrity attacks on control systems using a moving target approach. In *Proceedings of IEEE conference on decision and control* (pp. 5820–5826).

Weerakkody, S., & Sinopoli, B. (2016). A moving target approach for identifying malicious sensors in control systems. arXiv:1609.09043.

Wiese, M., Oechtering, T. J., Johansson, K. H., Papadimitratos, P., Sandberg, H., & Skoglund, M. (2018). Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design. *IEEE Transactions on Automatic Control*.

Woo, S., Jo, H. J., & Lee, D. H. (2015). A practical wireless attack on the connected car and security protocol for in-vehicle CAN. *IEEE Transactions on Intelligent Transportation Systems, 16*(2), 993–1006.

Wu, Y., Li, Y., & Shi, L. (2017). A game-theoretic approach to remote state estimation in presence of a DoS attacker. In *Proceedings of IFAC world congress* (pp. 2595–2600).

Xie, L., Mo, Y., & Sinopoli, B. (2011). Integrity data attacks in power market operations. *IEEE Transactions on Smart Grid, 2*(4), 659–666.

Xue, M., Wang, W., & Roy, S. (2014). Security concepts for the dynamics of autonomous vehicle networks. *Automatica, 50*(3), 852–857.

Yan, F., Sundaram, S., Vishwanathan, S. V. N., & Qi, Y. (2013). Distributed autonomous online learning: Regrets and intrinsic privacy-preserving properties. *IEEE Transactions on Knowledge and Data Engineering, 25*(11), 2483–2493.

Yan, Y., Qian, Y., Sharif, H., & Tipper, D. (2013). A survey on smart grid communication infrastructures: Motivations, requirements and challenges. *IEEE Communication Surveys & Tutorials, 15*(1), 5–20.

Yang, L., Cao, X., & Li, J. (2015). A new cyber security risk evaluation method for oil and gas SCADA based on factor state space. *Chaos, Solitons and Fractals, 89*, 203–209.

Yang, W., Lei, L., & Yang, C. (2017). Event-based distributed state estimation under deception attack. *Neurocomputing, 270*, 145–151.

Yuan, Y., & Mo, Y. (2015). Security in cyber-physical systems: Controller design against known-plaintext attack. In *Proceedings of IEEE conference on decision and control* (pp. 5814–5819).

Zeldovich, N. (Fall 2014). *6.858 Computer systems security*. MIT Lecture Notes.

Zero Days (2016). *Producer: Gibney, M. and Shmuger, M./Director: M. Gibney*. United States: Showtime

Zhang, H., Fata, E., & Sundaram, S. (2015). A notion of robustness in complex networks. *IEEE Transactions on Control of Network Systems, 2*(3), 310–320.

Zhao, C., Mallada, E., & Dörfler, F. (2015). Distributed frequency control for stability and economic dispatch in power networks. In *Proceedings of American control conference* (pp. 2359–2364).

Zhao, L., & Yang, G.-H. (2017). Adaptive sliding mode fault tolerant control for nonlinearly chaotic systems against DoS attack and network faults. *Journal of the Franklin Institute, 354*(15), 6520–6535.

Zhu, Q., & Başar, T. (2011). Robust and resilient control design for cyber-physical systems with an application to power systems. In *Proceedings of IEEE conference on decision and control* (pp. 4066–4071).

Zhu, Q., Bushnell, L., & Başar, T. (2013). Resilient distributed control of multi-agent cyber-physical systems. In *Control of cyber-physical systems* (pp. 301–316). Springer.

Zhu, Q., Tembine, H., & Başar, T. (2010). Network security configurations: A nonzero-sum stochastic game approach. In *Proceedings of American control conference* (pp. 1059–1064).