A survey on security control and attack detection for industrial cyber-physical systems

Derui Ding, Qing-Long Han*, Yang Xiang, Xiaohua Ge, Xian-Ming Zhang

The School of Software and Electrical Engineering, Swinburne University of Technology, Melbourne, VIC 3122, Australia

ARTICLE INFO

Article history:
Received 22 August 2017
Revised 6 October 2017
Accepted 6 October 2017
Available online 12 October 2017
Communicated by Prof. Zidong Wang

Keywords:
Industrial cyber-physical systems
Cyber-attacks
Attack detection
Security control

ABSTRACT

Cyber-physical systems (CPSs), which are an integration of computation, networking, and physical processes, play an increasingly important role in critical infrastructure, government and everyday life. Due to physical constraints, embedded computers and networks may give rise to some additional security vulnerabilities, which results in losses of enormous economy benefits or disorder of social life. As a result, it is of significant to properly investigate the security issue of CPSs to ensure that such systems are operating in a safe manner. This paper, from a control theory perspective, presents an overview of recent advances on security control and attack detection of industrial CPSs. First, the typical system modeling on CPSs is summarized to cater for the requirement of the performance analysis. Then three typical types of cyber-attacks, i.e. denial-of-service attacks, replay attacks, and deception attacks, are disclosed from an engineering perspective. Moreover, robustness, security and resilience as well as stability are discussed to govern the capability of weakening various attacks. The development on attack detection for industrial CPSs is reviewed according to the categories on detection approaches. Furthermore, the security control and state estimation are discussed in detail. Finally, some challenge issues are raised for the future research.

© 2017 Elsevier B.V. All rights reserved.

1. Cyber-physical systems

Recent years have witnessed rapid developments of cyber-physical systems (CPSs) due to advances in computing, communication, and related hardware technologies. As a new research frontier, a CPS is an integration of physical processes, ubiquitous computation, efficient communication and effective control [12]. Its holistic framework is shown in Fig. 1. Various social and physical applications have been performed in light of CPSs. The application fields include, but are not limited to, transportation networks, smart grids, health care, and water/gas distribution networks [69]. Moreover, networked control systems, wireless sensor and actuator networks, and wireless industrial sensor networks can be referred to as a subgroup of CPSs in the published literature [32,39,46,51,52,58,83,84,86,108,119,129]. For the recent developments, see survey papers [39,58,59] and the references therein. CPSs have been regarded as a core ingredient in the so-called 4th industrial revolution, and lots of efforts have been made for establishing its important position, such as, Industry 4.0 in Germany [62], and Industrial Internet in the U.S. [8]. It is worth mentioning that, performance analysis and synthesis have been intensively investigated for networked systems with various network-induced phenomena or communication protocols that include, but are not limited to, missing measurements [41,42,123], fading channels [24], communication delays [40,44,65,85,87,95], sampled data [47,104,143], Round-Robin protocols [77], stochastic communication protocols [142], event-triggering protocols [25,33,54,96,97,130,135].

CPSs are large-scale, geographically dispersed, federated, heterogeneous, and life-critical systems in which embedded devices such as sensors and actuators are networked to sense, monitor and control the physical world. In a CPS operation, there is no doubt that the resource scheduling via various shared or own networks plays an important role. One of the essential tasks is to decide which actuators/sensors should be activated to perform a particular action or how to manage control/sampling actions properly. Due to physical constraints or technological limitations, data among sensors, actuators and other networked components may be transmitted over networks without proper security protections. On the one hand, the interconnection of large-scale networked components makes it complicated to protect against inherent physical vulnerabilities therein. On the other hand, however, cyber-integration usually sets up an underscore on security and resilience against unforeseen patterns or threats from cyberspace [45]. Therefore, some new challenges are posed to traditional control, communication, and software theory. This paper aims to provide a

* Corresponding author.
E-mail address: q.han@cqu.edu.au (Q.-L. Han).
survey on the state-of-the-art of cyber-attack schemes and defense strategies in industrial CPSs from the perspective of control theory and propose several open research issues from system modeling, performance requirements, attack types to attack detection and security control.

The remainder of this paper is organized as follows, which is also illustrated in Fig. 2. In Sections 2 and 3, several typical system models and cyber-attack types are summarized from the engineering point of view. The performance indices are presented in Section 4 in detail. An overview of recent developments on attack detection for industrial CPSs is summarized in Section 5 from the aspect of the categories on detection approaches. The security control and state estimation are discussed in Section 6. Some challenge issues are raised in Section 7 to guide the future research.

2. System models for industrial CPSs

As is done for almost all of real-world engineering systems, the modeling of a CPS plays a key role in understanding and analyzing its dynamic behaviors. In other words, it is of both theoretical significance and practical importance to construct a unified system model before any subsequence analysis and synthesis. On the one hand, due to tight coupling and high coordination between cyber and physical worlds, a CPS can be regarded as a dynamically reorganizing and reconfiguring control system with high degree of automation at multiple spatial and temporal dimensions [34]. Such a spatial-temporal characteristic can be described as distributed parameter systems (DPSs). On the other hand, since control and monitor tasks are typically implemented on digital platforms, physical components usually run in a continuous-time way while other cyber-components are operated in a discrete-time manner. As such, how to bridge these two operations is an inevitable challenging issue for modeling and verifying CPSs. Although some existing modeling techniques are no longer adequately valid, as an alternative approach, a hybrid system model provides a possible cornerstone to cater for this kind of temporal complexity of CPSs. These two modeling strategies are surveyed as follows.

Due to the pioneering work [27], DPSs are extensively adopted to characterize cyber-physical processes by researchers from various fields. The distributed model can reflect the complexity of modeling CPSs by taking into consideration various phenomena or performance requirements, such as time asynchronism in measurement and control, imperfect communication (packet dropouts or network-induced delays), communication protocols, consistency of system states, and distributed consensus [18]. For example, PTIDES [138], which is a distributed programming model, relies on time synchronization while recognizing the imperfections. In [57], a cyber-based dynamic model is developed to model the energy systems in which a resulting mathematical model depends closely on cyber technologies. Noting that fractional calculus can better capture the complex dynamics of natural and man-made systems, three general models for generalized DPSs based on fractional Laplacian operators, fractional powers of operators or fractional derivative are employed [31] to describe some natural, physical, and anomalous phenomena such as sub-diffusion processes or super-diffusion processes.

In the ever-increasing research of power plants, monitoring, operation and control can be done in the framework of distributed control. It is widely recognized that a distributed control approach is better than a centralized one for the implementation of monitoring/control in smart grids with robustness requirements [28]. For example, a distributed PI-control scheme is proposed and applied to the frequency control of micro-grids [113]. The performance of communication infrastructures is investigated in the wide area power systems for an IEEE-118 bus network with both centralized and decentralized topologies [121]. The distributed cooperative control of multi-agent systems [11] is used to design the secondary voltage and frequency control of microgrids via input-output feedback linearization for transforming the nonlinear heterogeneous dynamics into a linear one. For limited communication networks, network-induced phenomena, such as network-induced delays, packet-dropouts and quantizations, possibly cause the degradation of the system performance. Therefore, considerable research interest is stirred to investigate the effects of these network-induced phenomena, and some preliminary results are reported in the literature, see [53,82,110] and the references therein. In [53], a filtering algorithm is developed to deal with the state estimation on power systems by taking the PMU measurements as inequality constraints. An intelligent controller based on reinforcement learning is designed in [110] for the load frequency control (LFC) of smart grids, in which network-induced phenomena (time-delays and changes in communication topology) are taken into consideration to examine the system performance.

As discussed in the first paragraph, hybrid system theory provides an effective mathematical tool for analyzing, designing, and optimizing the system performance or some parameters. For example, instead of maximizing the throughput or minimizing the delay, the design of certain medium access control (MAC) layer in wireless communications is investigated in [70] by using a developed scheduling algorithm. A distributed switching control scheme is presented in [50] to assure the almost sure safety in smart transportation systems. Moreover, by considering the digital nature of CPSs, an event-based sampling strategy is introduced to cater for the resource limitations of networked devices or the software program paradigm. In the scenario of sampling data, the resulting event-triggered scheme causes jumps of networked system states at some instants, which leads to typical hybrid dynamics [98,122,126]. Note that the hybrid invariance principle deserves paying some extra attention in the context of distributed event-based control so as to analyze Zeno phenomenon, which means that the hardware cannot tolerate arbitrarily close-in-time updates [98]. Recently, a new mixed time/event-triggered architecture, called MixCPS, is proposed in [128]. This kind of model supports the joint design of both time-triggered and event-triggered tasks. There are also some remarkable results on combining discrete and continuous dynamics in smart grids. For example, reachability analysis of hybrid systems is investigated [114] so that some
specifications in a power grid can be satisfied. In [26], a primary proportional-integral (PI) controller in an event-triggered architecture is designed to perform the load frequency control via an adaptive dynamic programming.

Finally, the decision-making behavior may occur in CPSs due to the application of artificial intelligence or human intent prediction, such as machine-mediated human-human interactions in supervisory control and data acquisition (SCADA) systems. Moreover, the behaviors of attackers or defenders in the domain of security and defense can be described by a discrete Markovian decision process due to the capability of characterizing such kinds of complex behavior. For example, a hidden Markov model (HMM) is employed to track the attacker’s action in the attack scenario [141]. Inspired by this idea, a finite-state HMM is utilized in [105] to describe the stochastic dynamics of CPSs under attacks.

3. Cyber-attacks in industrial CPSs

Since the vulnerability of CPSs, attacks can be injected into systems in a stealthy and unpredictable way through the cyber-parts [22]. For example, an adversary may lead to a disruption of coordination packets in medium access control layers or a compromise of networked components by injecting some malware (e.g., viruses and worms). Besides, while obtaining the encryption key, an attacker can illegally get access to the monitoring centers to destroy the normal operation. In other words, the attacker can either arbitrarily disturb the system dynamics with certain types or induce any perturbation when without enough security protection of hardware or software strategies. It is no doubt that the cyber-attacks are regarded as one of the major threats of CPSs. Within the model-based analysis framework, it is critical to describe cyber-attacks mathematically. Up to date, attacks in the published literature can be roughly divided into three categories: denial of service (DoS) attacks, replay attacks, and deception attacks, whose models are briefly listed in Table 1.
3.1. DoS attacks

The DoS attack is a kind of attempt to make the system resources unavailable. From the technology point of view, attackers can fill buffers of user domains or kernel domains, jam the shared network medium to prevent devices from communicating or receiving, or change the routing protocol. Some vivid examples on DoS attacks from networked control systems are described in [76]. So far, a few mathematical models are employed to quantitatively analyze the performance degradation, such as queuing models [76,91,109], Bernoulli models [7] or Markov models [10]. Notice that the system with queuing modelled attacks can be transformed into a time-delayed system and traditional analysis approaches [88,89] can be employed to effectively solve the stability-based problem, such as the methods developed in [65,120,136]. For instance, the saddle point equilibrium control is designed by utilizing the well-known zero-sum games [38]. In [91], a recursive predictive control scheme dependent on round-trip time delays is proposed to compensate for the adverse impact from the weak DoS attacks as well as the communication constraints. Recently, the optimal schedule of DoS is discussed in [134] to maximize the expected average estimation error at a remote estimator with/without intrusion detection systems. In addition, although different physical mechanisms are shown between DoSs and missing measurements, their mathematical models are the same in the framework of Bernoulli. Thus, it is effective to analyze the system performance of CPSs subject to DoSs by employing typical approaches for missing measurements.

3.2. Replay attacks

A replay attack is a natural strategy, in which a valid data transmission is maliciously or fraudulently repeated or delayed. For instance, in Byzantine replay attacks, attackers can repeat the data recorded from the compromised sensors or actuators in certain time. Such an attack cannot be detected easily due to the capability of passing examination of cryptographic keys and the resulting attack violates the timing constraints of CPSs. For wormhole attacks commonly existed in wireless sensor networks (WSNs) [107], adversaries can create a communication link (a wormhole tunnel) between two end points to replay messages observed in different regions [68]. To be more concrete, an attack can disrupt the routing protocol to make a false representation of the distance between the two colluding nodes. It is clear to see that one feature of this kind of attack is that no any system information is needed, including information on the designed controllers or estimators. Usually, counters or time-stamps may be adopted in the transmitted messages against replay attacks. According to the discussion above, replay attacks can be modelled as time-varying delays, whose information on their upper bounds, change rates and so forth is unknown. However, from the scheduling point of view, the admissible maximum upper bounds can be calculated by applying the time-delayed system theory together with optimization approaches.

3.3. Deception attacks

A deception attack is a type of cyber-attack, in which the data integrity is modified for transmitted packets among different cyber-parts [19,20]. For instance, Stuxnet worms can reprogram the code running in programmable logic controllers (PLCs) in SCADA systems such that the systems deviate from their expected behaviors. In power grid transmission systems, adversaries may launch attacks through hacking remote terminal units (RTUs) such as sensors in substations. Moreover, a hierarchical attack in water SCADA systems is described, and various attacks with different goals in different cyber layers are shown in [5]. It is worth mentioning that deception attacks in different scenarios can also be called as false data-injection attacks, malicious attacks, to just name a few. In [66], an optimization framework-based formulation is proposed to construct sparse attack vectors, and then to be extended to a distributed case by Ozay [102]. Recently, it is demonstrated [43] that stealth attacks exist if the number of compromised measurements exceed a certain value. Moreover, two models on sparse stealth attacks are constructed, respectively, for two typical scenarios: random attacks in which arbitrary measurements can be compromised; and targeted attacks in which specified states are modified. From the attacker point of view, the paper [60] is concerned with the attack scheduling of deception attacks with $\chi^2$ detectors and obtains the maximum number in the case of consecutive attacks and the desired attack probability in the case of randomly launched deception.

4. Performance analysis for industrial CPSs

In the context of dynamical systems, stability is an essential requirement, under which, other objectives, such as robustness, security and reliability, can be imposed on CPSs via various approaches. Compared with traditional industrial systems, performance analysis and synthesis for CPSs are still at an elementary stage especially in the theoretical level due to challenges from uncertainties in the environment, modeling errors of physical and software operations, undesired network induced phenomena, and various attacks and disturbances. Benefiting from the relatively mature modeling, stability issues on multi-agent systems and smart grids are discussed. For example, in [68], stability behavior is thoroughly analyzed for systems subject to both the in-band and out-of-band wormhole attacks via a developed passivity-based control-theoretic framework. Network-induced effects on stability from time delays, packet losses, bandwidths, quantizations, and changes in communication topologies, are examined in a load frequency control application of multi-area power systems [111].

Robustness, as one of the most useful properties, is the capability of how a system is insensitive to component variations. Recently, a new notion of robustness is introduced [115] to suppress the impact from both the bounded deviations of nominal behaviors and the sporadic disturbances in finite steps, and the satisfactory controller is designed under pseudo-polynomial time. A two-stage robust optimization model is presented in [16] to schedule jobs with time-of-use price information and a random local generation, where the probabilistic information is not available or not reliable. Furthermore, in a smart grid, the intermittent power generation from wind causes a large frequency fluctuation if the load-frequency control (LFC) capacity is not enough to compensate the unbalance of generation and load demand [117]. A proportional integral controller, whose parameters are optimized by the particle swarm optimization with mixed $H_2/H_\infty$ indices, is designed to enhance the robust performance against system uncertainties. In [49], a robust distributed control scheme is developed to achieve the desired power regulation by coordinating distributed generations and energy storage systems. More discussions on robustness of power systems can be found in [48,112] and the references therein.

In comparison with conventional networked control systems, CPSs usually run in a high risk of cyber-attacks due to the introduction of communication networks and heterogeneous IT elements, which gives rise to some additional security vulnerabilities. Therefore, the security of CPSs has been attracting recurring attention from a variety of research communities, and its significance has been exhibited in many works from several points of view [4,133]. Different from the robustness against uncertainties or disturbances, the security shows the ability of governing malicious behaviors or unanticipated events [139]. In this regard, a lot of engineering efforts are put forward to the security design.
and two types of techniques have been widely adopted to reduce the security risk, that is, information security tools and control-theoretic tools [4]. As summarized in [119], information security tools employed usually include authentication and access control mechanisms, network intrusion detection systems, patch management, security certification, and so forth. For instance, a wide-area monitoring and control (WAMC) is carried out in power systems for guaranteeing reliable operation and protection [13]. In addition, encryption is also adopted to protect data from destruction by unauthorized users and adversaries. For the development of the second tools, a detailed discussion is made in Section 6.

It is worth pointing out that the requirements of robustness and security are commonly predetermined and the system is designed off-line before attacks occur. Thus, it is impractical against all possible attacks and events for the designed system. Such a fact naturally gives rise to a new requirement, called resilience, which refers to the ability of recovery online after suffering from adversarial attacks [94]. The design of resilient control and estimation can be achieved by comparing the outputs with the ideal system behavior under an analytical model. It should be pointed out that this is a typical method employed in fault detection and diagnosis, see, e.g., [15]. A separation principle of estimation and control is revealed and a new characterization on the maximum resilience is exhibited in [30] for the system suffering from attacks. In [90], a resilient cruise controller is considered for an autonomous ground vehicle.

From the discussions above, the mathematical modeling and the performance indices of CPSs are surveyed in details and it is recognized that vulnerability comes mainly from the integration of cyber-based components via multiple networks. The following two sections focus on the detection of cyber-attacks and the security-based control under attacks.

5. Attack detection for industrial CPSs

Notice that the purpose of cyber-attacks is to destroy the desired performance of safety-critical CPSs. Thus, cyber-attacks are harmful to the enormous economy benefits or the loss of human lives. When they are detected and located in a timely fashion, the damage to overall systems can be controlled within a tolerable limit [105]. In light of such a perspective, attack detection plays a crucial role in maintaining the performance of CPSs. For many real-world systems, such as power systems and sensor networks, bad data detectors [99] are generally equipped to detect the deviation of the estimated state and provide an alarm operation. In this scenario, the change and the alarm caused by attacks will be limited to a tolerable level due to the presence of a specified threshold. In the published literature, there are two schemes to defend against cyber-attacks. One is to protect the important system components beforehand and the other is to identify the false data injected by attackers afterwards [43].

The first scheme can be reached via the deployment of either redundant components or redundant communication pathways. For instance, state information is directly monitored by the deployment of phasor measurement units (PMUs) absolutely free from attacks in power systems [66]. On the other hand, the scope of defense usually includes physical and cyber securities. As such, gates, guards, and authentication can be used against physical intrusions, and cryptography and fire-walls are exploited against cyber intrusions [3]. However, such an approach cannot provide adequately the protection against cyber adversaries. As an additional protection way, the second scheme can be applied to remove the contaminated data or correct them when attacks are detected. Up to date, four arguably representative detection strategies have been proposed in the published literature: (1) Bayesian detection with binary hypothesis; (2) weighted least square (WLS) approaches; (3) χ²-detector based on Kalman filters; and (4) quasi-FDI (fault detection and isolation) techniques.

Different from deception attacks and DoS attacks, detection of replay attacks needs to combine with some special ways. For this type of attack, one can insert an additional signal to the system inputs or disturb the system in non-regular time intervals. Because an attacker can not react on these changes, the behavior of the system outputs can not be subject to corresponding changes, which can be checked by utilizing above approaches. Therefore, when the system outputs are not the expected cases, replay attacks may occur. Furthermore, it is impossible to know a priori what type of attack may be inserted into the considered system in many scenarios and thus the identification of attack types is not always necessary, since the ultimate goal is to detect the existence of the attack (rather than its type) and then eliminate it to ensure safe and secure operation [106].

5.1. Bayesian detection with binary hypothesis

The hypothesis test with prior probabilities of two hypotheses is one of the most fashionable detection methods, and it is widely applied in the data fusion of sensor networks subject attacks [63,64,101]. For example, in [101], the performance limit of collaborative spectrum sensing is investigated under Byzantine attacks where malicious users send false sensing data to the fusion center leading to an increased probability of incorrect sensing results. Moreover, a closed form expression is further proposed for the optimal attacking strategies of the Byzantines in [63]. In order to overcome the limitation of 1-bit decision in sensor networks, a likelihood ratio detector based on binary hypothesis is introduced in [67] for smart grid security with limited number of meters compromised, and the trade-off between maximizing estimation error and minimizing detection probability is examined. When measurements are corrupted by colored Gaussian noises, an improved likelihood ratio detector in SCADA systems is designed to defense false data injection attacks in both observable and unobservable cases [116]. Similar results can be found in [35,124] and the references therein. Very recently, a novel Bayesian game-theoretic framework [14,103] has been developed to address the intrusion detection issues. Such a method caters for the nature that players do not have complete information on other players.

5.2. Weighted least square approaches

For measurement data, a weighted least-square (WLS) approach is a reliable and efficient scheme for the defense of attacks and therefore it has attracted ever-increasing research attention in power systems, see, e.g., [17,55,56,74]. In this framework, the measurement residual is usually constructed with the help of weighted least-square observers and then compared with a predetermined threshold in order to judge whether there exists a bad measurement. In light of the sparse nature of attacks, a detection mechanism is designed in [74] to detect the malicious attacks via the separation of nominal power grid states and anomalies. In the sense of weighted least-square estimation combined with two-player zero-sum games, a least-budget defense strategy is introduced [17] to protect power systems against false data injection attacks and meanwhile the meter selection is tackled by formulating it as a mixed integer nonlinear programming. Furthermore, statistics characteristics can be taken into account if a 2-norm on measurement residual obeying χ² distribution is introduced.

5.3. χ²-detector based on Kalman filters

The χ²-detector has the capability against both bad data and false data (such as DoS attacks, short-term/long-term random
attacks) by integrating a Kalman filter (instead of a WLS estimator) where it is implemented in the innovative mechanism. Due to the dependence of statistics characteristic, it is unable to detect false data-injection attacks derived statistically [78]. In [80], a feasibility condition of replay attacks is presented and the optimization countermeasure of the detection probability is obtained for SCADA systems. The cosine similarity matching and $\chi^2$-detector is simultaneously adopted in [100] to detect false data injection attacks in smart grids. It is worth noting that, compared with $\chi^2$-detector, the cosine similarity matching is more robust for detecting false data injection attacks. Recently, for discrete-time linear time-invariant processes with a residue-based detector following $\chi^2$ distribution, an optimized attack strategy is given in [37], under which an attacker remains undetectable to the exploited detector and the system renders maximum estimation errors.

5.4. Quasi-FDI techniques

Fault detection and isolation (FDI) is an effective approach to monitor a system, identify when a fault has occurred, and pinpoint the type of fault and its location. The FDI is widely applied in networked control systems. Inspired by its mature approaches, an analysis of vulnerabilities of cyber-physical systems in the face of unforeseen failures and external attacks has received increasing attention in the recent years and some preliminary results have been published in literature, see, for instance, [9,94]. In [94], undetectable and unidentifiable attacks are characterized from system-theoretic and graph-theoretic perspectives and a Luenberger-type detection filter is designed. Similarly, detectability of attacks is explored in [9] and two adaptive sliding mode observers are designed to estimate state attacks and sensor attacks with ultimately uniformly bounded errors. Different from FDI approaches with residual constructing, a co-estimation of system states and attacks inspiration from fault-tolerant state reconstruction, as an alternative scheme, is investigated in [6,109]. For instance, a scheme based on an unknown input observer is developed in [6] for detecting faults, which may come from simultaneous water pilfering through offtakes and sensor compromise. In [6], an unknown input observer is used to estimate the states of SCADA systems subject to stealthy deception attacks. Recent research in [1] provides a model-free approach to detect faults. The main characteristic of this approach lies in the capability of learning both the nominal conditions of the system under inspection and the fault dictionary from acquired data. It should be pointed out that the FDI approach executed in engineering systems need to be carefully reexamined when dealing with CPS security due to the different physical mechanisms between cyber-attacks and component failures of physical systems [79].

6. Security control and estimation for industrial CPSs

State estimation plays an important role in better understanding system dynamical behaviors and executing some specific control tasks. In networked scenarios, it is an inevitable challenge how to mitigate the impact of cyber-attacks as well as various probabilistic communication failures. For steady states, a decentralized estimation scheme with adaptive weighted matrices is proposed [71] to alleviate the influence of bad data in SCADA systems. Recently, the distributed counterpart is formulated as a weighted least square (WLS) optimal issue in which an adaptive weight assignment is adopted to dynamically adjust the measurement weight in [137]. From the perspective of attackers, the effect on state estimation from attacks is investigated [17,118,125]. On the other hand, the estimation on a deterministic mean-shift parameter in the presence of Byzantine attacks is made and the proportion relationship between two different groups on distinct attacks is discussed specifically in [2,132].

For linear time-invariant systems with a prior constraint on the number of jamming attacks, a Nash equilibrium is provided between attackers and defenders [72]. Furthermore, the maximum tolerant number of attacks (similar to above prior constraints) is particularly discussed in [30]. It is highly related to a concept of strong observability under which the system states can be accurately reconstructed via attack-resilient estimators. It is worth noting that a similar notion, named as s-spars observability can be found in [109]. According to a similar framework in [30], an $l_0$-norm-based state estimator is adopted in [90] to guarantee the boundedness of estimation errors caused by noises, modeling errors as well as cyber-attacks. Besides, noting that the deployment of bad data detectors results in a bound constrain of attacks, some issues about state estimation on stochastic time-varying nonlinear systems subject to randomly occurring deception attacks are discussed [81] in the framework of Kalman filtering. Moreover, this idea is extended to the distributed case [21] and the stability of the proposed algorithm is analyzed in depth. Very recently, the identifying probability of bad data detectors has been integrated into an iterative process of the error covariance and an optimal linear estimation scheme has been developed against DoS attacks.

Besides the resilient state estimation above, CPSs also need to mitigate the threat from secret attackers via various control strategies. Compared with other control applications, security control techniques for CPSs remain at an infant stage and few results can be found in the published literature on this topic. Because of the core of critical infrastructures, the resulting successful attack on control networks is generally more serious in comparison with attacks on others. With respect to networked control systems subject to various cyber-attacks, some preliminary and interesting results can be found in [23,76,133] for DoS attacks, in [5,19,20,22,61,82,93] for deception attacks, and in [68,140] for replay attacks. To just mention a few, an event-triggered controller is designed in [23] to tolerate DoS attacks characterized by given frequency and duration properties. An optimal schedule of jamming attacks is proposed in [133] to maximize the linear quadratic Gaussian cost under energy constraints. An event-triggering consensus resilient-control with a state-independent threshold is discussed in [20] for discrete-time multi-agent systems with both lossy sensors and cyber-attacks.

In the context of CPS security, a tradeoff between security and stability has also received growing attention. To name a few, some attacker-defender games are constructed in [29,73] to analyze the interactions of two parties and the impact on the stability. Moreover, a theoretical result on scheduling algorithms is developed [131] to achieve the balance among robustness, schedulability, and power consumption. Furthermore, schemes of attack-resilient cooperative control for the CPSs consisting of distributed agents are investigated in [75,127,128] via identification and isolation of the misbehaving cyber elements or the cascade protection design of lossless systems. For instance, a resilient control strategy is introduced [75] to regulate the active power from a cluster of distributed generators at a certain ratio of its maximal available power. It should be pointed out that almost all results on security in the framework of control theory are based on an assumption that system dynamics need to be simple, which leads to a gap between theoretical results and practical engineering applications.

7. Conclusions and challenging issues

Recent advances on security control and attack detection for industrial CPSs have been surveyed in the framework of control and estimation theory. First, the typical modeling on CPSs and cyber-attacks has been presented from the engineering point of view.
Then, robustness, security and resilience as well as stability have been discussed to govern the capability of weakening various attacks. Furthermore, developments on attack detection for industrial CPSs have been reviewed based on different detection approaches. Finally, some main results on security control and state estimation have been discussed in detail.

In what follows, we pay more attention to the limitations of some existing results and propose several challenging issues on this topic, which sheds insightful light on the further research.

(1) System modeling and methodologies:

- In the published literature, several existing tools based on system models are far from meeting CPS design requirements. Computing and decision-making behaviors combined with both communications and physical dynamics should be further abstracted and modeled at different levels.
- In the framework of control theory, the current analysis approach with the goals of mitigation attacks cannot completely handle the complex system dynamics, not to mention the case where there exist the time-varying node behaviors, time-varying topologies or nonlinearities.
- Note that security usually is a hard constraint. For the stochastic nature of CPSs, the methodologies in mean-square sense expose some considerable conservatism and how to make an effective scheme in almost-certain sense is still challenging.
- In the published literature, some results are proposed in the light of an assumption that system information is known for attack scheduling or an attack scheme is open for attack defenses. There is no doubt that such an assumption is a stumbling block for real applications.

(2) Attack detection and compensation:

- CPSs may be subject to multiple attacks at the same time. An adaptive strategy compensating different types of attacks has not received adequate attention yet for industrial CPSs and the impact on the system performance should be profoundly discussed.
- A large perturbation on system states could be caused by injected false data while a detection rate may suffer from a slightly increase, and hence sensitivity and reliability of attack detection should be further investigated.
- The system complexity is inevitably increased when communication protocols, network induced phenomena and cyber-attacks are simultaneously taken into account. Application conditions of typical detection schemes and technologies (such as, binary hypothesis, game theories, $\chi^2$ distribution) may not be guaranteed. Hence, developing some novel detection approaches and compensation strategies overcoming the challenging issues above is of significance.

(3) The system performance and the quality of service:

- It is nontrivial to fuse attack detection and resilience control in a uniform framework. The successful implementation of such an idea may lead to some significant improvements of the security performance of CPSs.
- In practical engineering, security requirements and resource constraints (communication bandwidth, limited energy, etc.) usually need to be taken into consideration simultaneously. How to co-design the system parameters of CPSs by considering both the security and the quality of service is of importance.
- The fusion of physical systems and cyber systems gives rise to higher performance requirements, such as robustness, stability, security and reliability. Thus, it may lead to a particularly attractive topic on multi-objective optimization to satisfy the requirement of real-world CPSs.

Acknowledgments

This work was supported in part by the Australian Research Council Discovery Project under Grant DP160103567, the National Natural Science Foundation of China under Grant 61573246, the Shanghai Rising-Star Program of China under Grant 16QJA1403000, and the Program for Capability Construction of Shanghai Provincial Universities under Grant 15550502500.

References


...
D. Ding et al. / Neurocomputing 275 (2018) 1674–1683

Derui Ding received both the B.Sc. degree in Industry Engineering in 2004 and the M.Sc. degree in Detection Technology and Automation Equipment in 2007 from Anhui Polytechnic University, Wuhu, China, and the Ph.D. degree in Control Theory and Control Engineering in 2014 from Donghua University, Shanghai, China. From July 2007 to December 2014, he was a teaching assistant and then a lecturer in the Department of Mathematics, Anhui Polytechnic University, Wuhu, China. He is currently a senior research fellow with the School of Software and Electrical Engineering, Swinburne University of Technology, Melbourne, Australia. From June 2012 to September 2012, he was a research assistant in the Department of Mechanical Engineering, the University of Hong Kong, Hong Kong. From March 2013 to March 2014, he was a visiting scholar in the Department of Information Systems and Computing, Brunel University London, UK. His research interests include nonlinear stochastic control and filtering, as well as multi-agent systems and sensor networks. He has published around 40 papers in refereed international journals. He is a very active reviewer for many international journals.

Xiaohua Ge received the B.Eng. degree in electronic and information engineering from Nanchang Hangkong University, Nanchang, China, in 2008, the M.Eng. degree in control theory and control engineering from Hangzhou Dianzi University, Hangzhou, China, in 2011, and the Ph.D. degree in computer engineering from Central Queensland University, Rockhampton, Qld, Australia, in 2014. He was a Research Assistant with the Centre for Intelligent and Networked Systems, Central Queensland University, from 2011 to 2013. In 2014, he was a Research Fellow with the Centre for Intelligent and Networked Systems, Central Queensland University, Rockhampton, Australia. From 2015 to 2016, he was a Research Fellow with the Griffith School of Engineering, Griffith University, Gold Coast, Australia. He is currently a Lecturer with the School of Software and Electrical Engineering, Swinburne University of Technology, Melbourne, Australia. His current research interests include networked control and filtering, distributed networked control systems, multi-agent systems and sensor networks.

Qing-Long Han received the B.Sc. degree in Mathematics from Shandong Normal University, Jinan, China, in 1983, and the M.Sc. and Ph.D. degrees in Control Engineering and Electrical Engineering from East China University of Science and Technology, Shanghai, China, in 1992 and 1997, respectively. From September 1997 to December 1998, he was a Post-doctoral Research Fellow with the Laboratoire d’Automatique et d’Informatique Industrielle (LAI) (currently, Laboratoire d’Informatique et d’Automatique pour les Systemes, LIAS), cole Supérieure ding nieurs de Poitiers (ESIP) (currently, cole Nationale Supérieure ding nieurs de Poitiers (ENSIP)), Université de Poitiers, France. From January 1999 to August 2001, he was a Research Assistant Professor with the Department of Mechanical and Industrial Engineering at Southern Illinois University at Edwardsville, USA. From September 2001 to December 2014, he was Laureate Professor, an Associate Dean (Research and Innovation) with the Higher Education Division, and the Founding Director of the Centre for Intelligent and Networked Systems at Central Queensland University, Australia. From December 2014 to May 2016, he was Deputy Dean (Research), with the Griffith Schools, and a Professor with the Griffith School of Engineering, Griffith University, Australia. In May 2016, he joined Swinburne University of Technology, Australia, where he is currently Pro Vice-Chancellor (Research Quality) and a Distinguished Professor. In March 2010, he was appointed Chang Jiang (Yangtze River) Scholar Chair Professor by Ministry of Education, China. Prof. Han is one of The World’s Most Influential Scientific Minds: 2014–2016 and is a Highly Cited Researcher in Engineering according to Thomson Reuters. He is an Associate Editor of a number of international journals including IEEE Transactions on Industrial Electronics, IEEE Transactions on Industrial Informatics, IEEE Transactions on Cybernetics, and Information Sciences. His research interests include networked control systems, neural networks, time-delay systems, multi-agent systems and complex dynamical systems.

Xian-Ming Zhang received the M.Sc. degree in applied mathematics and the Ph.D. degree in control theory and engineering from Central South University, Changsha, China, in 1992 and 2006, respectively. In 1992, he joined Central South University, where he was an Associate Professor with the School of Mathematics and Statistics. From 2007 to 2014, he was a Post-Doctoral Research Fellow and a Lecturer with the School of Engineering and Technology, Central Queensland University, Rockhampton, Qld, Australia. From 2014 to 2016, he was a Lecturer with the Griffith School of Engineering, Griffith University, Gold Coast, Qld, Australia. In 2016, he joined the Swinburne University of Technology, Melbourne, VIC, Australia, where he is currently a Senior Lecturer with the School of Software and Electrical Engineering. His current research interests include H-infty filtering, event-triggered control systems, networked control systems, neural networks, distributed systems, and time-delay systems. Dr. Zhang was a recipient of the National Natural Science Award (Level 2) in China in 2013, and the Hunan Provincial Natural Science Award (Level 1) in Hunan Province in China in 2011, both jointly with Prof. M. Wu and Prof. Y. He. He received the IET Premium Award in 2016, jointly with Prof. Q.-L. Han. He serves as an Associate Editor for the Journal of the Franklin Institute and a member of the editorial board of Neural Computing and Applications.