Cybersecurity in process control, operations, and supply chain

Sandra Parker, Zhe Wu, Panagiotis D. Christofides

A R T I C L E   I N F O

Keywords:
Cybersecurity
Machine learning
Process control
Process operation
Process safety
Supply chain management

A B S T R A C T

With the integration of computation, networking, and physical process components to seamlessly combine hardware and software resources to improve process efficiency, cybersecurity has become increasingly important for reliable process control, process operation, and supply chain management in the chemical process industries. This paper presents an overview of recent works on cybersecurity issues in the area of process control, process operation and supply chain. We start with an overview of recent cyber-attack detection and mitigation works via machine learning (ML) and model predictive control (MPC) to detect and handle intelligent cyber-attacks. Several most common intelligent cyber-attacks in industrial control systems are first presented, followed by machine learning detection methods and resilient control strategies with encryption–decryption tools to achieve secure communication in the sensor–controller and controller–actuator links. Novel control architectures with inherent robustness to prevent cyber-attacks are then presented. We continue with an overview of cybersecurity issues in process operations and supply chains as well as the interface between information technology and operational technology. Finally, we discuss recent efforts on the interface of cybersecurity and process safety and conclude with a discussion of open issues in this emerging research field.

1. Introduction

Over the last two decades, internet communication and wireless networks have been starting to replace or complement existing wired point-to-point communications in traditionally large-scale process operations (e.g., Christofides et al., 2007; Venkatasubramanian, 2009; Li, 2016; Daozitis et al., 2018; Venkatasubramanian, 2019; Shah et al., 2020). As these new developments bring improved efficiency to the existing system, the heightened concern for unestablished, industrial cybersecurity at all levels has also been rising following cyber-attacks that disrupt standard operations. Due to the connectivity and interaction between the cyber and physical components in chemical processes, operational cybersecurity requires a different strategy from the traditional information technology (IT) approach. This is a consequence of key differences between IT and OT (operational technology): (a) OT employs purpose-built technologies and protocols, (b) OT systems are typically kept much longer than IT systems where most companies cannot easily perform upgrades or implement changes to the technology, (c) upgrades or changes in the OT space generally require plant shutdowns which are costly, and as a result, may lead to equipment running for years, making it difficult for its support, and (d) OT is very much concerned with reliability and intellectual property.

 Despite these differences that raise challenges in implementing cybersecurity solutions in the OT space, recent cyber-attacks have driven the need for developing and implementing novel cybersecurity solutions in the OT space. Most companies and organizations recognize today the need to deploy a combination of traditional IT cybersecurity products and services with tailored operational technology (OT)-specific cybersecurity solutions. The failure to ensure cybersecurity in OT can lead to unsafe and potentially catastrophic consequences in a chemical process operation, causing critical asset damage and human injuries. During the past two decades with the facilitation of technology and processes, the industry has exposed the vulnerabilities of unestablished cybersecurity systems following the rise of cyber-attacks. From 2000 to 2019, a reported 77 cybersecurity-related incidents were uncovered in critical infrastructure including the process industry with a vast majority of attacks on energy and oil production industries (Iaiani et al., 2021). The lack of adequate prevention of cyber-attacks endangers the balance of the economy, environment, and society. For instance, in 2021, the oil pipeline system in the United States, Colonial Pipeline, endured a cyber-attack, which stalled the transportation of oil to much of the eastern United States, causing...
Skyrocketing gas prices (Tsvetanov and Slaria, 2021) and volatile supplies of fuel. In 2015, Ukraine encountered the BlackEnergy malware attack that forced over 200,000 people without power and electricity (Böröcz et al., 2021). The aforementioned examples of cyber-attacks are stark reminders of the repercussions of cyber-attacks and their impact on societal welfare, which are reasons for a greater need for well-established cybersecurity systems. Therefore, the design and implementation of cyber-defense in OT domains that involves industrial control, operation, and supply chain management systems remain an ongoing systems and control engineering research issue of great practical importance.

Chemical and manufacturing industries have adopted firewall isolation, multi-factor authentication, and developed cyber protection protocols over the past decade to improve cybersecurity, particularly in the context of IT tasks. However, with the integration of IT and OT in the framework of Industry 4.0 and the development of intelligent, targeted cyber-attacks that have access to the technical details of the control system and production processes in the plant that aim to modify the operator and control system actions applied to a chemical process, the need for OT task cybersecurity has grown significantly. Earlier efforts to enhance the cybersecurity of the OT space started around 2010 but gained momentum around 2017 by taking advantage of industrial process operation and automation groups. Today, OT cybersecurity is viewed as a key concern across the entire chemical sector and aims to establish cybersecurity standards and raise the level of protection across chemical plants. In particular, to enhance cybersecurity and physical security of process operations, the fundamental cybersecurity research roadmap (a framework, whose key components are summarized in Fig. 1) proposed originally by the National Institute of Standards and Technology (NIST) (2018) that has influenced the efforts of many companies including Dow, has proposed a five-step plan to detect and mitigate the impact of cyber-attacks with recovery plans: identify, detect, protect, respond, and recover. However, within this five-step framework, there are many key research questions that need to be considered. Specifically, despite a series of recent efforts over the past five years, designing efficient detection methods and suitably optimal, yet secure, operation control and supply chain strategies for chemical processes in the presence of intelligent cyber-attacks remains an important, fundamental research issue. Furthermore, while the development of most of the existing cyber-attack detection methods still depends partly on human analysis, the increased use of data and the design of stealthy cyber-attacks pose challenges to the development of timely detection methods with high detection accuracy. In the following paragraph, we provide an overview of results on the development of machine learning-based cyber-attack detection schemes as this is a topic central to cybersecurity approaches in the OT space, and it is covered in greater detail later on in the manuscript (please see “Machine Learning-Based Cyber-Attack Detection” section).

Machine learning, a method of data analysis that can help engineers learn from data, identify patterns, and make decisions with minimal human intervention, has attracted increasing attention and has shown promising potential for use in the detection of cyber-attacks. While the use of machine learning methods in solving classification, regression, and clustering problems has a long history (please see Schmidhuber (2015), Kramer (2016) for a broad overview), over the last decade, we have witnessed many efforts to address these problems more efficiently with machine learning algorithms by taking advantage of more advanced and powerful computing resources/platforms, and many free and open-source software libraries. To detect cyber-attacks, machine learning methods can be utilized to solve classification problems to determine the existence of cyber-attacks in the chemical plant and its control systems using an abundance of industrial process data that is generated by machines and devices under normal operations and under cyber-attacks. Machine-learning methods deployed for cyber-attack detection were presented in a number of works (Tsai et al., 2009; Buczak and Guven, 2015; Olay et al., 2015). Using various machine-learning classification methods, cyber-attacks on power systems were distinguished from process disturbances in Hink et al. (2014), and a behavior-based intrusion detection algorithm was developed to identify the type of attack (Junejo and Goh, 2016). Similarly, the detection of cyber-attacks in a chemical process was realized via the development of feedforward artificial neural networks in Wu et al. (2018a), where compromised signals were rerouted to a secure sensor upon detection. In Shon and Moon (2007), a hybrid approach using support vector machines and genetic algorithms was implemented and compared to existing network intrusion detection systems in industry. An overview of recent research directions for applying supervised and unsupervised machine learning techniques to address the problem of anomaly detection was presented in Omar et al. (2013).

Among many machine learning methods, neural network and many of its variances have demonstrated remarkable performance. For instance, a Long Short-Term Memory (LSTM) recurrent neural network (RNN) was used to build a classifier model for the intrusion detection system in Kim et al. (2016). The anomaly detection algorithm outlined in Goh et al. (2017) also used a LSTM network as a predictor to model normal behavior of a water treatment testbed and used the Cumulative Sum (CUSUM) method to identify anomalies. A multi-layer data-driven cyber-attack detection system was proposed in Zhang et al. (2019a) where four classification methods including k-nearest-neighbor, decision tree, bootstrap aggregating, and random forest, were used to detect cyber-attacks including man-in-the-middle, denial-of-service, data exfiltration, data tampering, and false data injection attacks based on network and host system data. Many variants of convolutional neural networks with different topologies, parameters, and structures were analyzed for the task of intrusion detection in cybersecurity of network traffic in Vinayakumar et al. (2017), which have shown significant improvement over conventional classifiers. These recent literature contributions have demonstrated the feasibility of machine-learning algorithms in anomaly detection including anomalies caused by cyber-attacks. At any large-scale chemical production plant, a tremendous amount of data is being collected and archived daily in the historian. Using neural-network algorithms, the data can be utilized to train effective detection devices for monitoring and guarding the plant against malicious cyber-attacks.

Besides the detection of cyber-attacks, efforts are made to improve cyber and physical security through a variety of fundamental operation and control methods that address the following aspects: security by design, advanced recovery, advanced threat detection, secure remote access, and combined safety (Fig. 1). This work will discuss recent works within the elements of Fig. 1 in context of cybersecurity of process control and operation systems and supply chains. Specifically, to guarantee the process performance and to mitigate the impact of cyber-attacks, process control systems, e.g., model predictive control (MPC) and economic MPC (EMPC), utilizing encrypted signals may be employed to operate the process with secure remote access in the presence of cyber-attacks. With regard to security by design and advanced
recovery, a cyber-secure two-tier control architecture can be developed and integrated with ML-based detectors to enhance process cybersecurity by reconfiguring the control system to stabilize the process at the original steady state upon the detection of a cyber-attack. Additionally, to account for the interactions among control, cybersecurity, and safety systems, the integration of attack detection and control policies as well as combined control and safety systems have been pursued and will be discussed. Safety Technology (ST) will also be an important component in the IT-OT framework of cybersecurity to integrate safety with cybersecurity. Finally, directions for future research in the context of cybersecurity of process control, process operation systems, and supply chains will be discussed.

2. Class of nonlinear process systems

To describe mathematically the various types of cyber-attacks as well as detection and mitigation control methods, we need to introduce a suitable notation, a class of process systems, and specific stabilizability assumptions. Specifically, we consider the class of continuous-time nonlinear systems represented by the following state-space model:

\[ \dot{x} = f(x, u, w) : = f(x) + g(x)u + h(x)w, \quad x(0) = x_0 \]

where the \( n \)-dimensional state vector is denoted by \( x \in \mathbb{R}^n \) and \( u \in \mathbb{R}^k \) denotes the \( k \)-dimensional manipulated input vector bounded by \( u \in U \). The set \( U \) defines the maximum value \( u_{\text{max}} \) and the minimum value \( u_{\text{min}} \) for input vectors, i.e., \( U := [u_{\text{min}} \leq u \leq u_{\text{max}}] \subset \mathbb{R}^k \). \( w \) is the disturbance vector, where \( W := \{ w \in \mathbb{R}^l \mid |w| \leq \theta, \theta \geq 0 \} \). \( f(\cdot), g(\cdot), \text{ and } h(\cdot) \) are sufficiently smooth vector and matrix functions of dimensions \( n \times 1, n \times k, \text{ and } n \times q \), respectively. We assume that \( f(0) = 0 \) without loss of generality, and therefore, the origin is a steady-state of Eq. (1). Additionally, we assume there exists a feedback controller that can stabilize the system at the origin. Specifically, we assume there exists a continuously differentiable Lyapunov function \( V(x) \) and a Lyapunov-based controller \( u = \Phi(x) \in U \) such that the origin of the nominal system \( (x(t) \equiv 0) \) is rendered asymptotically stable for the states in an open neighborhood \( D \) around the origin. The stability region \( \Omega_p := \{ x \in D \mid V(x) \leq \rho \}, \rho > 0 \) is characterized as a level set of \( V \) within \( D \). Throughout this manuscript, \( |\cdot| \) is used to denote the Euclidean norm of a vector. Set subtraction is denoted by “\( \setminus \)”, i.e., \( A \setminus B := \{ x \in \mathbb{R}^n \mid x \in A, x \notin B \} \).

3. Background and description of cyber-attacks

From the perspective of process control systems as well as process operation and supply chains, cyber-attacks are malicious signals that can compromise actuators, sensors, communication channels between devices, and the operation and control system algorithms. With respect to control system cybersecurity, cyber-attacks modify the control implementation using process and control system information in an attempt to disrupt closed-loop performances. A comprehensive review in Ashibani and Mahmoud (2017) includes an analysis on security issues, requirements, and possible solutions at various layers of the OT architecture. A review of possible weaknesses in corporate networks and in production environments is presented in Ashgar et al. (2019). In Amin et al. (2012), a hierarchical attack on automated canal systems was described with various deception attacks in different cyber layers and a field-operational test attack was reported on the Gignac canal system located in Southern France.

Sensor attacks strategically modify the feedback measurements of the attacked states, from which the controller receives and subsequently computes a control action that is different or contrary to its actual optimal value based on the true plant state. Attacker attacks also have access to the plant model and controller design details, which aim to diverge the system away from its ideal operating point. However, instead of altering the sensor measurements, actuator attacks modify the direction and magnitude of the control actions without being detected by sensor monitoring tools. Common detection strategies include designing excitation signals that are superimposed on the control commands to increase the detectability of the attack and developing an input observer to detect attacks as well as estimate the magnitude of the attack (Muniraj and Farhood, 2019). In addition to the detection of actuator attacks, an isolator was developed to identify the affected actuator(s) in the network. As intelligent cyber-attacks are adaptive to the process and control system behavior, we may assume that they are as powerful as having access to the measurement feedback signals (sensor attack), control command signals (actuator attack), or auxiliary information such as the threshold and bias parameters in detection methods such as cumulative sum (CUSUM) (Mohanty et al., 2007; Cárdenas et al., 2011). Being aware of the process and controller behavior, the attacks will therefore have information on the stability region of the process, as well as the existing alarm triggers imposed on the input and output variables. Among sensor cyber-attacks, some common attack types are denial-of-service attacks, replay attacks, and deception attacks — such as min–max, geometric, and surge attacks (Cárdenas et al., 2011). The formulations of the aforementioned three deception and replay attacks are presented below.

3.1. Min–max cyber-attack

Min–max attacks are designed to induce the maximum destabilizing impact within the shortest time without being detected. In order to stay undetectable by classical detection methods such as CUSUM, which detects cyber-attacks by calculating the cumulative sum of the deviation between the expected and measured states based on the process model of Eq. (1), min–max attacks are introduced using the falsified state values furthest from the equilibrium point (minimum or maximum) such that the system does not exit the closed-loop stability region \( \Omega_p \). In this way, the min–max attacks ensure that the attacked state measurements fed to the control system do not exit the stability region and do not trigger any conventional detection alarms. The min–max attack can be formulated as follows:

\[
\dot{x}(t_i) = \min_{\epsilon \in \mathbb{R}^n} \max_{\epsilon \in \mathbb{R}^n} \{ x \mid V(x(t_i)) = \rho \}, \quad \forall \ i \in [t_0, t_0 + L_a]
\]

where \( \rho \) defines the level set of the Lyapunov function \( V(x) \) that characterizes the stability region \( \Omega_p \) for the system of Eq. (1). \( \dot{x} \) is the compromised sensor measurement at each sampling step, \( t_i \) marks the time instant that attack is added, and \( L_a \) denotes the time duration of the attack in terms of sampling periods.

3.2. Replay cyber-attack

In a replay attack, the attacker first records segments of the system output corresponding to a nominal operating condition where large oscillations occur. The attacker then intercepts and resets the current process state measurements to these pre-recorded values. Replay attacks can be represented by the following equations:

\[
\dot{x}(t_i) = x(t_i), \quad \forall \ k \in [k_0, k_0 + L_s], \quad \forall \ i \in [t_0, t_0 + L_s]
\]

where \( x(t_i) \) is the true plant measurement, \( L_s \) represents the length of the attack in terms of sampling periods, and \( \dot{x} \) is the series of replay attacks introduced at time \( t_i \) duplicating previous plant measurements that are recorded starting from time \( t_{k_0} \). As previous plant outputs are obtained from legitimate closed-loop measurements and given by secure sensors, these state values are supposedly inside the stability region and the operating envelope. Therefore, by replicating these values and feeding them back to the controller, classical detectors will not be able to recognize the abnormality caused by replay cyber-attacks.

3.3. Geometric cyber-attack

Geometric cyber-attacks aim to deteriorate the closed-loop system stability slowly at the beginning, then geometrically increase their
impact as time progresses, with their maximum damage achieved at the end of the attack duration. Initially, the attacker adds a small constant $\beta$ to the true measured output where $\beta$ is well below the maximum allowable value as defined in a min–max attack. At each subsequent time step, this offset is multiplied by $(1 + a)$, where $a \in (0, 1)$, until it reaches the maximum allowable attack value.

The formulation of a surge attack can be written in the following form:

$$\hat{x}(t_i) = x(t_i) + \beta \times (1 + a)^{i-L_i}, \quad \forall \ i \in [i_0, i_0 + L_a]$$

where $\hat{x}$ is the compromised sensor measurement, $\beta$ and $a$ are parameters that define the magnitude and speed of the geometric attack.

### 3.4. Surge cyber-attack

Surge attacks act similarly as min–max attacks initially to maximize the disruptive impact for a short period of time; then they are reduced to a lower value by introducing a bounded noise $\eta_i \leq \eta(t_i) \leq \eta_u$ ($\eta_u$ and $\eta_l$ are the upper and lower bounds of the noise, respectively) such that the cumulative error between state measurements and their steady-state values will not exceed the threshold defined by some statistic-based detection methods such as CUSUM. The formulation of a surge attack is presented below:

$$\hat{x}(t_i) = \min_{x \in \mathbb{R}^s} \max_{x \in \mathbb{R}^s} \{x | V(x(t_i)) = \rho\}, \quad \text{if} \quad i_0 \leq i \leq i_0 + L_s$$

$$\hat{x}(t_i) = x(t_i) + \eta(t_i), \quad \text{if} \quad i_0 + L_s < i \leq i_0 + L_a$$

where $i_0$ is the start time of the attack, $L_s$ is the duration of the initial surge, and $L_a$ is the total duration of the attack in terms of sampling periods. To illustrate the pattern and effect of the four cyber-attack types discussed above, Fig. 2 shows the true concentration values when min–max, replay, geometric, and surge cyber-attacks are introduced.

### 4. Machine learning-based cyber-attack detection

The first step in the cybersecurity roadmap is to detect and identify cyber-attacks by developing advanced threat detection and protection methods. Cyber-attack detection carried out using data-based approaches, and more specifically, machine-learning methods, have been studied (Huang et al., 2007; Omar et al., 2013; Agrawal and Agrawal, 2015). Machine learning can be utilized to develop detection algorithms based on the time-series data from the dynamic operation of the system of Eq. (1) (Wu et al., 2018a). Depending on the training data, the neural networks can be used to distinguish between “attack” and “no attack” (two classes), or to identify the type of attack (multiple classes). While under attack, data collected from individual sensors can also be used to locate the corruption where the neural network model distinguishes between multiple classes with each class representing one problematic sensor. In our study, a feedforward artificial neural network is used for supervised classification. Through a series of nonlinear transformations, each layer in the neural network consists of a series of nonlinear functions of the weighted sum of inputs or neurons (i.e., activation functions), yielding values for the neurons in the subsequent layer from the previous layer.

The structure of a neural network model with two hidden layers is shown in Fig. 3, with each input unit representing a nonlinear function $p(\cdot)$ of the full state measurements at each sampling time and an output vector representing the probability of each class label. The two-hidden-layer feedforward neural network is mathematically formulated as follows:

$$\theta_j^{(1)} = g_1\left(\sum_{i=1}^{N_p} w_{ij}^{(1)} p(\hat{x}(t_i)) + b_j^{(1)}\right)$$

$$\theta_j^{(2)} = g_2\left(\sum_{i=1}^{N_p} w_{ij}^{(2)} \theta_j^{(1)} + b_j^{(2)}\right)$$

where $g_1$ and $g_2$ are activation functions.
the testing dataset. Additionally, to reduce false alarm rates, a sliding NN model is calculated by the ratio of the number of data samples model. The classification accuracy (i.e., the test accuracy) of the trained strate the performance of the neural network since the test dataset is that will depend on the requirement of the machine-learning detector.

\[ y = g \left( \sum_{i=1}^{h_1} a^{(1)}_{ij} x_i + b^{(1)}_j \right) = \left[ a^{(2)}_1, a^{(2)}_2, \ldots, a^{(2)}_H \right]^T \]

(6c)

where \( \theta^{(3)}_j, j = 1, \ldots, h_2, l = 1, 2 \) are the neurons in the first (\( l = 1 \)) and second (\( l = 2 \)) hidden layers, respectively. The output node is represented by \( \theta^{(3)}_j, j = 1, \ldots, H \), where \( H \) is the number of class labels. In general, the number of layers is determined through trial-and-error to achieve the best classification accuracy and computational efficiency. The input node \( p(x(t)) \) receives the state measurement at time \( t_i \), where \( i = 1, \ldots, N_T \) is the length of the time-varying trajectory. \( a^{(l)}_i \) and \( b^{(l)}_i \) represent the weights connecting neurons \( i \) and \( j \) in consecutive layers (from \( l = 1 \) to \( l \)), and the bias term on the \( j \)th neuron in the \( l \)th layer, respectively. Based on the information received from the previous layer as well as the optimized biases, weights, and the nonlinear activation function \( g(z) \), each layer calculates an output and sends it to the next layer. Examples of the activation functions include the softmax function \( g(z) = \frac{e^z}{\sum e^z} \), the hyperbolic tangent sigmoid transfer function \( g(z) = \frac{2}{1+e^{-2z}} - 1 \), and some other common functions such as the sigmoid, radial basis functions, and Rectified Linear Unit (ReLU).

The classification accuracy of the test dataset is utilized to demonstrate the performance of the neural network since the test dataset is independent of the training dataset and is not used in training the NN model. The classification accuracy (i.e., the test accuracy) of the trained NN model is calculated by the ratio of the number of data samples with correct predicted classes to the total number of data samples in the testing dataset. Additionally, to reduce false alarm rates, a sliding alarm verification window in Fig. 4 is implemented, where the number of positive attack detections \( D_i = 1 \) within this window needs to surpass a threshold before a cyber-attack alarm is confirmed. The size of this verification window and the threshold value are determined based on the closed-loop evolution of the process as these two parameters have a direct impact on the detection time and alarm rate.

5. Attack-resilient MPC approaches exploiting sensor redundancy

5.1. Tracking MPC

Upon the detection of an attack on the sensors providing real-time state measurements to the control system, advanced recovery strategies have been developed to mitigate the impact of attacks. Specifically, one of the most common approaches adopted in industry is to switch to an accurate measurement from redundant, secure sensors. We present the resilient control strategies in the framework of Lyapunov-based MPC that can be represented by the following optimization problem:

\[
\mathcal{J} = \min_{u(t)} \int_{t_k}^{t_{k+N}} L_i(\hat{x}(t), u(t))dt
\]

(7a)

s.t. \( \dot{x}(t) = F(x(t), u(t), 0) \)

(7b)

\( x(t_k) = x(t_k) \)

(7c)

\( u(t) \in U, \forall t \in [t_k, t_{k+N}) \)

(7d)

\( V(x(t_k), u(t_k)) \leq V(x(t_k), \Phi(x(t_k))) \), if \( V(x(t_k)) > \rho_{\text{min}} \)

(7e)

\( V(\hat{x}(t)) \leq \rho_{\text{min}}, \forall t \in [t_k, t_{k+N}) \), if \( V(x(t_k)) \leq \rho_{\text{min}} \)

(7f)

where \( x(t) \) is the predicted state trajectory, \( S(\Delta) \) is the set of piecewise constant functions with period \( \Delta \), and \( N \) is the number of sampling periods in the prediction horizon. \( V(x(t_k), u(t_k)) \) represents the time derivative of \( V(x) \), i.e., \( \frac{d}{dt} V(x(t_k), u(t_k), 0) \), \( \Phi(x) \) is the stabilizing control law assumed for the nonlinear system of Eq. (1). The cost function \( L_i(\hat{x}(t), u(t)) \) satisfies \( L_i(0, 0) = 0 \) and \( L_i(\hat{x}(t), u(t)) > 0 \), \( \forall(\hat{x}(t), u(t)) \neq (0, 0) \) such that the minimum value of the cost function will be attained at the equilibrium of the system of Eq. (1). We assume that the states of the closed-loop system are measured at each sampling time instance and will be used as the initial condition in the MPC optimization problem of Eq. (7) in the next sampling time. Specifically, based on the measured state \( x(t_k) \) at \( t = t_k \), the above optimization problem is solved to obtain the optimal solution \( u^*(t) \) over the prediction horizon \( t \in [t_k, t_{k+N}) \). The first control action of \( u^*(t) \), i.e., \( u^*(t_k) \), is sent to the control actuator to be applied over the next sampling period. Then, at the next sampling time \( t_{k+1} := t_k + \Delta \), the optimization problem is solved again, and the horizon will be rolled one sampling time. Specifically, the MPC optimization problem minimizes the objective function of Eq. (7a) over
5.2. Economic MPC

In addition to utilizing sensor redundancy in the context of tracking MPC, one can develop a similar approach for Economic MPC (EMPC), which is another form of MPC that directly integrates process economic considerations with process control to dynamically optimize process economics through time-varying operation. A number of past works have been developed to address stability, safety, and computational efficiency issues in EMPC (Heidarinejad et al., 2012; Angel et al., 2011; Müller et al., 2013; Ellis et al., 2014; Wu et al., 2018b). To handle the cyber-attacks that compromise both closed-loop stability and process economic benefits under EMPC, the attack-resistant Lyapunov-based EMPC design, which combines open-loop and closed-loop control, is developed and represented by the following optimization problem:

\[
\mathcal{J} = \max_{u \in \mathcal{S}_U} \int_{0}^{T_{N_p} + t_F} l_x(\bar{x}(t), u(t)) dt \\
\text{s.t.} \quad \dot{\bar{x}}(t) = F(\bar{x}(t), u(t)), \\
\quad u(t) \in U, \forall t \in [T_{N_p}, T_{N_p} + t_F] \\
\quad \bar{x}(T_{N_p}) = \bar{x}(t_F) \\
\quad V(\bar{x}(t)) \leq \rho_{\text{secure}}, \forall t \in [T_{N_p}, T_{N_p} + t_F], \text{ if } \bar{x}(T_{N_p}) \in \Omega_{\text{secure}} \\
\quad V(\bar{x}(T_{N_p}), u) \leq V(\bar{x}(T_{N_p}), \Phi(\bar{x}(T_{N_p}))), \text{ if } \bar{x}(T_{N_p}) \in \Omega_{\text{secure}} \setminus \Omega_{\text{secure}}
\]

where \( \Omega_{\text{secure}} \) is the set that the process will be operated within such that the system will not immediately lose stability when under malicious cyber-attacks. \( \rho_{\text{secure}} \) is the number of sampling periods in one material constraint period, which is the prediction horizon for open-loop control. Since it is common that chemical processes are subject to periodic feed stock constraints, which are specified as part of the input constraint set \( U \), we also require, for example, the quantity of feed materials to be limited within a fixed period of time \( t_{N_p} \). During this period of time (termed material constraint period), the total feed material is constrained to a constant value \( C \), i.e., \( \int_{0}^{T_{N_p}} \bar{u}(t) dt = C \), where \( \bar{u}(t) \) represents feed material used at every sampling period. Therefore, the material consumption constraint renews every \( t_{N_p} \). If the total operation time is longer than one material constraint period, this material consumption constraint results in cyclic operation of the plant, and consequently, the cyclic behavior of the state-space trajectory. At the start of a new material constraint period, the total consumption limit is renewed, as new feed materials become available to be used again for the next constraint period. In the presence of cyber-attacks, the attack-resilient EMPC is implemented as follows. At time \( t_k \), the EMPC in the open-loop control mode receives the state measurement \( x(t_k) \) and computes the optimal trajectory of \( N_p \) control action that will be applied in a sample-and-hold manner until the end of this material constraint period. In the case that there are no cyber-attacks or process disturbances, this optimal trajectory of control actions would yield maximum economic benefits while meeting all input and state constraints. While at the closed-loop operation, if the feedback measurement is no longer reliable and cannot be used for closed-loop control, the open-loop control actions that were calculated at the beginning of the material constraint period will be used as a substitute until the end of the material constraint period.

At the end of the material constraint period, a cyber-attack detector is activated to determine any occurrence of an attack and the reliability of the control system is reassessed. The detector will provide information on the security status of the feedback measurements over the latest material constraint period. Upon mitigating the impact of a confirmed attack and/or confirming the security of the control system, closed-loop control with secure feedback measurement can be reactivated as a new material constraint when the period starts. The operation of EMPC around a secure operating region is illustrated in Fig. 6 and the attack-resistant strategy of switching from closed-loop to open-loop control is illustrated in Fig. 7.

6. Integrated attack detection and control policies: Additional recent results

A key issue for cyber-attacks in chemical process industries is that they can impact process safety by directly adjusting process states or potentially the equipment condition (Niem et al., 2020). This
motivates a fundamental understanding of the nature of the interactions between the control design and cyber-attacks on various components of the control loop. For example, Durand (2018) elucidated a challenge with attempting to thwart the falsification of sensor measurements using a randomized control law selection. This challenge prevents attacks from causing an issue that may require certain control laws to be used in different regions of state-space, which an attacker can use to provide attacks that are destabilizing. Randomness was further explored in the context of taking advantage of noise in quantum computation for adding randomness to control action selection in Rangan et al. (2022a) and for similar reasons, was not able to thwart cyber-attacks on the sensor measurements. However, designing cyberattack detection policies in tandem with control laws can aid in forcing attacks to reveal themselves by setting expectations for what a non-attacked process state trajectory should appear as using the control theory and then by detecting whether the control-theoretic requirements are achieved using the detection policy (in the spirit of other active attack detection policies such as dynamic watermarking (Satchidanandan and Kumar, 2016)).

Integrated attack detection and control policies have been explored in the context of Lyapunov-based economic model predictive control (EMPC) (Heidarnejad et al., 2012) of nonlinear systems when sensors (Durand and Wegener, 2020; Oyama and Durand, 2020), actuators (Rangan et al., 2022b), or sensors and actuators at the same time (Oyama et al., 2022b) can be attacked. These policies have considered indicators of attacks such as whether the Lyapunov function is decreasing along the state measurement trajectory, is comparing state predictions with state measurements, or is adding redundancy in state estimates that can be used to provide cross-checking of whether state measurements are correct. These provide different safety guarantees (in the sense that the closed-loop state is maintained within an expected region of state-space at least for some time after an attack) when the sensors are attacked, when the actuators are attacked, or when different detection policies are combined and both sensors and actuators are attacked. The case that sensor measurements are attacked has also been considered for the case that the process dynamics could change at the same time (Oyama et al., 2021; Rangan et al., 2021). Through an extension of the LEMPC-based integrated detection and control policies to this case through a two-tier attack detection policy, safety for at least some time period after an attack on the sensors, process dynamics change, or both can be guaranteed under sufficient conditions.

Fig. 7. Demonstration of attack-resilient EMPC control strategy by switching from closed-loop control actions to pre-calculated open-loop control actions when the state measurements reach the boundary of the secure operating region.
detect a range of cyber-attacks. Since switching may excite the process dynamics resulting from the potential of false alarms, a switching condition was presented to minimize false alarms (Narasimhan et al., 2022b).

7. Encrypted control

In addition to detection and recovery, another way to enhance the cybersecurity of control systems is to establish secure remote access. Encryption-based control using the encryption of the communication signals (e.g., semi-homomorphic encryption methods) can be developed to ensure secure communication in the sensor–controller and controller–actuator links in the presence of cyber-attacks. Homomorphic Encryption (HE) allows the performing of arithmetic operations such as addition and multiplication in the ciphertext (encrypted message) space such that no decryption on messages is needed in order to perform these operations. Unlike conventional control schemes, encrypted control systems compute encrypted inputs based on encrypted states and encrypted controller parameters without intermediate de-

It is demonstrated in Eq. (9) that the addition operation can be carried out with the encrypted numbers directly, and therefore, no decryption is needed at this stage. Following the additive homomorphism property, a semi-encrypted product can be computed as follows. Given \(m_1, m_2 \in Z_M\) such that \(m_1, m_2 \in Z_M\), the multiplication of \(m_1\) and \(m_2\) can be written as addition of \(m_1\) with itself for \(m_2\) times. Therefore, using the additive homomorphism property of Eq. (9), the following equation is obtained for all \(r \in Z_M\).

\[
E_M(m_1, m_2, r_1, r_2) = E_M(m_1, r_1)E_M(m_2, r_2) \mod M^2 = c_1c_2 \mod M^2
\]  

(9)

Note that the product calculated in Eq. (10) is semi-encrypted since only \(c_1\) is encrypted and \(m_2\) is not. This also explains why Paillier cryptosystem is not a fully homomorphic scheme.

Since the messages/numbers to be encrypted in Paillier cryptosystem are required to be a set of integers, quantization of the signals is needed to map real numbers to integers to encrypt-decrypt the communication signals in the closed-loop system of Eq. (1) (Darup et al., 2017). Specifically, we first map the set of real numbers to the set \(Q_{1,\delta}\) as follows.

\[
g_{1,\delta} : \mathbb{R} \rightarrow Q_{1,\delta}
\]

\[
g_{1,\delta}(a) = \arg \min_{b \in Q_{1,\delta}} |a - b|
\]

(11)

where \(Q_{1,\delta}\) is a set of rational numbers between \(-2^{\delta-1}d^{-1}\) and \(2^{\delta-1}d^{-1} - 2^{-\delta}\) separated from each other with a resolution of \(2^{-\delta}\), i.e., \(q \in Q, 3\delta \notin [0, 1]^{1}\), such that \(q = -2^{\delta-1}d^{-1}\beta + \sum_{i=1}^{\delta-1} 2^{-i-1}\beta_i\). Subsequently, we map the set of rational numbers to the set of integers \(Z_{\delta,2}\) as follows:

\[
f_{1,\delta} : Q_{1,\delta} \rightarrow Z_{\delta,2}
\]

\[
f_{1,\delta}(q) := 2^{\delta}q \mod 2^{\delta}
\]

(12)

While the sensor data is encrypted and utilized by the controller to compute control actions without decryption, the control actions need to be decrypted before sending to the actuator to be applied to the system of Eq. (1). Therefore, the inverse operation \(f_{1,\delta}^{-1} : Z_{\delta,2} \rightarrow Q_{1,\delta}\) is defined as follows:

\[
f_{1,\delta}^{-1}(m) := \begin{cases} \frac{1}{2^{\delta}} \left( t - 2^{\delta} \right) & \text{if } t \geq 2^{\delta-1} \\ t & \text{otherwise} \end{cases}
\]

(13)

While Eq. (13) maps the decrypted inputs back to the rational number space, it can be observed that there will be some error in the input due to the difference between the actual control input and the one mapped to its closest rational number within the set \(Q_{1,\delta}\). Therefore, to address this issue, the control system should be designed to ensure a certain degree of robustness with respect to potential encryption process errors. For example, the quantizations of the state measurements and controller matrices can be modeled as artificial disturbances to the system of Eq. (1) (i.e., \(h(x)\)) and accounted for in the design of a robust control scheme.

Darup et al. (2017) describes the encrypted control law evaluation only for a linear system having the control law of the form \(u = kx + b\). This limitation is imposed by the nature of the Partially Homomorphic Cryptosystems which only allow the addition and multiplication operations in the encrypted message space. Thus, in the case of nonlinear systems or nonlinear control laws, it is important to modify our approach. A nonlinear control law can be defined as \(u = \Phi(x)\). The sensor measures the states, encrypts them using the Public Key and sends them to the controller establishing a secure communication of the signals. The controller then decrypts the states and performs the nonlinear control law calculations. Once the control action has been computed,
the controller then encrypts it (using the Public Key) and sends it to the actuator. At the actuator, the control input is decrypted and the control action is applied to the nonlinear system. We use quantization functions to convert the states and controller parameters to the integer message space in order to prepare them for encryption. This quantization of real numbers induces some loss of data or quantization errors. Thus, in the case of nonlinear systems, it is important to model these quantization errors as disturbances to the system and the controller should be able to handle these disturbances. A schematic of this encrypted control scheme for nonlinear systems is shown in Fig. 8. More work needs to be done in this direction to address the stability, robustness, and performance issues for explicit nonlinear control as well as model predictive control.

8. Control architecture design for handling cyber-attacks: Decoupling stability and performance objectives

To enhance the robustness of MPC to cyber-attacks, a two-tier control architecture was designed by Chen et al. (2020b) to allow convenient reconfiguration of the control system to stabilize the process to its operating steady state upon successful detection of cyber-attacks. Specifically, we consider the following class of continuous-time nonlinear systems:

\[
\begin{align*}
\dot{x}(t) &= f(x(t), u_c(t), u_a(t)) \\
y_c(t) &= h_c(x(t)), \quad y_a(t) = h_a(x(t))
\end{align*}
\]

where \( x \in \mathbb{R}^{n_x} \) is the state vector, \( y_c(t) \in \mathbb{R}^{n_y} \) represents the vector of state measurements that are sampled continuously (e.g., reactor temperature), and \( y_a(t) \in \mathbb{R}^{n_y} \) represents the vector of networked state measurements that may be sampled asynchronously at \( t = t_k \) (e.g., reactor product concentration); \( u_c \) and \( u_a \) are the manipulated input vectors, which are constrained by \( u_c \in \mathbb{R}^{n_u}, u_a \in \mathbb{R}^{n_u} \) \( \in \mathbb{R} \).

Through \( y_c \) and \( y_a \), we assume measurement of the full state vector \( x \) can be obtained at \( t_k \). The cyber-secure control architecture integrates a lower-tier control system that uses the dedicated sensor measurements, \( y_c(t) \), to ensure stability of the steady-state of the closed-loop system and an upper-tier, high-performance control system (e.g., MPC) that uses both dedicated \( y_c(t) \) and networked \( y_a(t) \) sensor measurements to improve closed-loop performance significantly above what could be achieved with the lower-tier control system.

Specifically, we assume that for the lower-tier controller, there exists an explicit feedback controller \( u_c(t) = \phi_c(y_c(x)) \in U \) that can stabilize the closed-loop system of Eq. (14) using only the continuous measurements \( y_c(t) \). The Lyapunov-based MPC (LMPC) of Eq. (7) can be used as the upper-tier controller to fully utilize the networked (potentially asynchronous) state measurements \( y_a(t) \) and to compute \( u_a(t) \) that improves the overall closed-loop performance over what can be achieved with \( \phi_c(x) \) while not jeopardizing the stability properties achieved by \( u_c(t) \). Upon detection of an attack on the sensors providing networked asynchronous state measurements to the two-tier control system, the control system reconfiguration logic allows for two mitigation plans. First, the control system can deactivate the upper-tier controller completely and operate the system under the stabilizing lower-tier control system only, which uses cyber-secure, dedicated sensor measurements. Since the lower-tier controllers are capable of driving the process to its operating steady state with secure continuous measurements, the effect of the cyber-attacks is fully eliminated in the closed-loop system in this case and the process is stabilized to the operating steady-state. Second, if a sensor isolation detector is also implemented, it will be activated once a sensor attack is verified. Subsequently, the upper-tier controller can choose to switch the compromised sensor to its redundant back-up sensor with secure readings. By abandoning the corrupted sensor and using its back-up sensor using a secure sensor-controller communication, the upper-tier controller remains functional and is able to drive the process to its steady state with better closed-loop performance. In the extreme case that both continuous and asynchronous sensor measurements are attacked, the upper-tier controller will be shut off and the lower-tier controllers will reroute their continuous measurement signals from the corrupted sensors to their respective secure back-up sensors. The two-tier control design, where the networked sensor measurements, \( y_a(t) \), used only by the upper-tier controller may be under potential cyber-attack, is illustrated in Fig. 9. In addition to shutting off the upper-tier control system, the use of encryption of the signals of the upper-tier control system may be employed at the expense of reduced closed-loop performance in order to improve its robustness to signal quantization errors.

9. Application to a chemical process example

We use a chemical process example as a benchmark to demonstrate the application of integrated data-based attack detectors and cyber-secure MPC schemes that minimize the impact of cyber-attacks on process operation. Specifically, machine learning detectors via feedforward neural network are developed using sensor measurements under nominal and noisy operating conditions in Chen et al. (2020b), and applied online to a simulated reactor-reactor-separator process. Two reactions take place in series (A → B → C) in both CSTRs and the overhead vapor from the flash tank is recycled to the first CSTR. The performance-improvement LMPC receives asynchronous measurements on the mass fractions of A and B in each of the three vessels (\( x_{A1}, x_{B1}, x_{A2}, x_{B2}, x_{A3}, x_{B3} \)), all of which can be subject to cyber-attacks), and manipulate the fresh feed flowrate into the second CSTR, \( F_{20} \). Three
Fig. 9. Two-tier control-detector architecture showing lower-tier controllers using continuous secure sensor measurements and an upper-tier MPC using both continuous (secure) and networked (vulnerable to cyber-attacks) sensor measurements, where secure back-up sensors, if available, can be used to replace the compromised networked sensor for upper-tier MPC (Chen et al., 2020b).

Fig. 10. State-space plot showing the evolution of true process states (blue trajectories) and attacked state measurements (red trajectories) over two material constraint periods under the resilient LEMPC when (a) min–max, (b) geometric, and (c) surge attacks, targeting the temperature sensor are successfully detected by a NN detector at the end of the first material constraint period, \( t = 0.06 \text{ h} \), where the dash-dotted ellipse is the stability region \( \Omega \rho \) and the dashed ellipse is \( \Omega \rho \text{sec} \) (Chen et al., 2020a).

Table 1

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Training (%)</th>
<th>Testing (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal, One Attack</td>
<td>99.6</td>
<td>92.2</td>
</tr>
<tr>
<td>With Noise, One Attack</td>
<td>99.9</td>
<td>100</td>
</tr>
<tr>
<td>With Noise, Two Attacks</td>
<td>98.2</td>
<td>91.4</td>
</tr>
<tr>
<td>With Noise, Sensor Isolator</td>
<td>99.6</td>
<td>99.0</td>
</tr>
</tbody>
</table>

The process description and parameter values are given in Chen et al. (2020b), and are omitted here. An upper-tier Lyapunov-based MPC, which uses networked sensor measurements to improve closed-loop performance, is coupled with lower-tier cyber-secure explicit feedback controllers to drive a nonlinear multivariable process to its steady state. Although the networked sensor measurements may be vulnerable to cyber-attacks, the two-tier control architecture ensures that the process will stay immune to destabilizing malicious cyber-attacks. Simulation results demonstrate the effectiveness of these detection algorithms in detecting and distinguishing between multiple classes of intelligent cyber-attacks that may occur at different locations of the sensor network. Upon the detection of cyber-attacks, the two-tier control architecture allows convenient reconfiguration of the control system to stabilize the process to its operating steady state. The training and testing accuracy for detecting the presence of an attack, the attack type, or the location of the attack are given in Table 1. Furthermore, a modified Lyapunov-based EMPC using combined closed-loop and open-loop control action implementation schemes was proposed in Chen et al. (2020a) to optimize economic benefits in a time-varying manner while maintaining closed-loop process stability and resiliency against various types of cyber-attacks. Data-based cyber-attack detectors are developed using sensor data via machine-learning methods and these detectors are periodically activated and applied online in the context of process operation. With FNN detectors trained and applied online, the closed-loop state evolution under the resilient EMPC is shown in Fig. 10 where the process is exposed to three types of sensor cyber-attacks (Chen et al., 2020a).
10. Control architecture design for handling cyber-attacks: Decentralized and distributed control

In addition to constructing control architectures where control systems are structured according to closed-loop stability and performance objectives, decentralized and distributed control systems provide an efficient solution to many challenges of controlling large-scale industrial processes (Christofides et al., 2013) and may provide certain advantages with respect to robustness to cyber-attacks in comparison with centralized control systems. We will discuss decentralized and distributed control systems in the context of model predictive control.

In a decentralized MPC system, no communication is established between the different local controllers, therefore each controller does not have any knowledge on the control actions calculated by the other controllers. While this may lead to reduced closed-loop performance, it may be beneficial in the context of cyber-attacks as the control systems can operate independently. Specifically, for each subsystem, a separate MPC is designed to regulate the states $x_j$ of the subsystem $j = 1, \ldots, N_{sys}$, and optimize the respective control actions. Each decentralized MPC $j = 1, \ldots, N_{sys}$, the optimization problem of MPC $j$ is as follows:

$$ J_j = \min_{u_j} \int_{t_0}^{t_f \in U_j} L(\hat{x}(t), u_j(t)) dt $$  

(15a)

subject to

$$ \dot{\hat{x}}(t) = F(\hat{x}(t), u_j(t), 0) $$  

(15b)

$$ \hat{x}(t_0) = \hat{x}(t_0) $$  

(15c)

$$ u_j(t) \in U_j, \forall t \in [t_k, t_{k+1}) $$  

(15d)

$$ V(\hat{x}(t)) \leq \rho_j, \forall t \in [t_k, t_{k+1}) $$ if $\hat{x}(t_k) \in \Omega_{p_j}$  

(15g)

The control actions optimized by MPC $j$, denoted by $u_{j \ast}$, will be applied to the corresponding control actuators in subsystem $j$. Note that while full-state feedback measurements could be available to all MPCs, each MPC in the decentralized MPC only has the information of the process dynamics of its respective subsystem.

To achieve better closed-loop control performance compared to decentralized MPC, distributed MPC systems may be developed to take advantage of some level of communication that may be established between the different controllers. Specifically, iterative distributed MPC systems (one of several DMPC architectures discussed in Christofides et al. (2013)) allow signal exchanges between all controllers, thereby allowing each controller to have full knowledge of the predicted state evolution along the prediction horizon and yielding better closed-loop performance via multiple iterations at the cost of more computational time. For example, both controllers communicate with each other in a two-MPC system to cooperatively optimize the control actions. The two controllers solve their respective optimization problems independently in a parallel structure and at the end of each iteration they will exchange solutions with each other. The optimization problem of MPC 1 in an iterative distributed LMPC at iteration $c = 1$ is presented as follows:

$$ J_1 = \min_{u_{1 \ast} \in \Xi_1} \int_{t_0}^{t_f} L(\hat{x}(t), u_{1 \ast}(t), \Phi_1(\hat{x}(t))) dt $$  

(16a)

subject to

$$ \dot{\hat{x}}(t) = F(\hat{x}(t), u_{1 \ast}(t), 0) $$  

(16b)

$$ u_{1 \ast}(t) \in U_1, \forall t \in [t_k, t_{k+1}) $$  

(16c)

$$ \hat{x}(t_0) = \hat{x}(t_0) $$  

(16d)

$$ V(\hat{x}(t)) \leq \rho_1, \forall t \in [t_k, t_{k+1}) $$ if $\hat{x}(t_k) \in \Omega_{p_1}$  

(16f)

where the variables and constraints are defined following those in the decentralized MPC design. For each control action $j$ corresponding to subsystem $j = 1, \ldots, N_{sys}$, the following takes place:

$$ \dot{\hat{x}}(t) = F(\hat{x}(t), u_j(t), 0) $$  

(15b)

$$ u_j(t) \in U_j, \forall t \in [t_k, t_{k+1}) $$  

(15d)

$$ V(\hat{x}(t)) \leq \rho_j, \forall t \in [t_k, t_{k+1}) $$ if $\hat{x}(t_k) \in \Omega_{p_j}$  

(15g)

While both distributed and decentralized MPC systems are designed to alleviate the computational complexity for solving large-scale optimization problems for multiple subsystems as opposed to centralized MPC, the vulnerability to cyber intrusions also increases with the expansion of communication networks. The work in Chen et al. (2021) investigates the effect of different types of standard cyber-attacks on the operation of nonlinear processes under centralized, decentralized, and distributed model predictive control systems. The robustness of the decentralized control architecture over distributed and centralized control architectures was analyzed. Considering the inherent structure and operating requirement of both systems, the decentralized control system was found to exhibit greater robustness against potential cyber-attacks at the expense of a small performance loss versus centralized and distributed MPC.

While isolation and handling of actuator faults in nonlinear processes under continuous, synchronous measurements have been studied in Gani et al. (2007), Mhaskar et al. (2008), Ohran et al. (2008a), McFall et al. (2008), Ohran et al. (2008b), detection and handling of cyber-attacks in cooperative, distributed control architectures for nonlinear processes is a challenging task due to cyber-attack intelligence. Additionally, it cannot be addressed with the aforementioned process monitoring and control methods dealing with the centralized control systems because cyber-attacks may not only affect the sensor measurements going to the controllers but also the inter-controller communication. Therefore, as shown in Fig. 11 (Chen et al., 2021), a machine-learning-based detector can be developed to detect and isolate cyber-attacks in the context of sequential DMPC. Subsequently, a resilient control strategy can be employed that orchestrates the reconfiguration of the control system. This strategy determines if the MPC algorithms should be reconfigured or new backup control loops (e.g., switching from distributed MPC to decentralized MPC where there is no communication between the controllers) should be activated in the presence of cyber-attacks in order to preserve closed-loop system stability.

A chemical process example of two CSTRs in series with the reaction $A \rightarrow B$ taking place in both reactors is simulated in Chen et al. (2021) to demonstrate the robustness of decentralized control architectures and the effectiveness of the neural-network detection scheme in maintaining the closed-loop stability of the system. The process description and parameters can be found in Chen et al. (2021) and are omitted here. The following Figs. 12–13 from Chen et al. (2021) show the true closed-loop state trajectories under the decentralized control-detector system. The proposed control-detector architecture and detection methodology can be extended to other applications of model predictive control or other methods of advanced control systems, in general.
programming, and linear/nonlinear programming have been widely used to improve cyberinfrastructure security in many ways involving prevention/protection, detection, mitigation, response and recovery from a cyber-threat. In Sun et al. (2018), recent research results on the cybersecurity of a smart grid were discussed and a cyber-power system test-bed was used to demonstrate the impact of attacks and the effectiveness of cybersecurity solutions. More recently, Smetana et al. (2021) introduced the integration of food system technologies with cyber–physical system technologies and pointed out the need for the development of efficient defense mechanisms to address potential cyber-food safety risks and hazards.

In another recent work, Cheung et al. (2021) provided an overview of research works on cybersecurity in supply chain management. It was pointed out that the measures for enhancing cybersecurity can be classified into three broad categories: precautionary measures, real-time recovery measures, and aftermath measures, which are very similar to the practices introduced for process control systems (i.e., advanced threat detection, security by design, secure remote access, and advanced recovery) and reviewed in the earlier part of this manuscript. Specifically, some common precautionary measures for supply chains are the identification of vulnerabilities in cyberspace, secure access, authentication, data protection, firewall, and gateway development. Machine learning, game theory, Bayesian analysis, and attack path generation and analysis methods have been applied to identify and locate vulnerabilities while blockchain technology has been widely used for data protection and authentication (Koheri, 2017; Taylor et al., 2020). For example, game theory has been widely used in the supply chain to optimize a defender’s strategy by modeling each player’s (i.e., attack and defender) behavior and strategies and to capture the interaction between two opposing players. To provide a high-level description of this approach, consider an attacker with a set of $N$ potential attacking strategies $\{s_a\} \in S_a$ and a defender with a set of $N_d$ potential defense strategies $\{s_d\} \in S_d$, where $S_a, S_d$ are the space of all possible strategies for attacker and defender, respectively, and $i$ represents the strategy index (Colbert et al., 2020). Given an attack strategy $a_i$ and a defense strategy $d_j$, the attacker suffers a cost $C_{a(i)}$ to penetrate the security layer and accomplish its goal and the defender spends a cost $C_{d(j)}$ to apply its strategy. Additionally, given a strategy tuple $(i, j)$, it is assumed that the attacker succeeds in attacking the $i$th system with probability $p_i(s_a(i), s_d(j))$, $i = 1, \ldots, N$. Assuming the attacker gains a benefit $b$ by successfully attacking a network of $N$ subsystems and both the attacker and defender have complete knowledge of the system, the utility $u_a$ for the attacker can be calculated as follows:

$$u_a(s_a(i), s_d(j)) = b \cdot P_i(s_a(i), s_d(j)) - C_{a(i)}$$  \hspace{1cm} (18)

Similarly, the utility $u_d$ for the defender is as follows:

$$u_d(s_a(i), s_d(j)) = b \cdot [1 - P_i(s_a(i), s_d(j))] - C_{d(j)}$$  \hspace{1cm} (19)

Therefore, for both the attacker and defender, the objective is to select the optimal strategy that maximizes their utilities, for which a number of strategy selection methods have been developed in literature. Interested readers may refer to Zhu et al. (2010), Do et al. (2017), Attiah et al. (2018), Cheung and Bell (2021) for the applications of game theory in cybersecurity. In addition to the optimization-based approaches discussed above, laws, policies, regulations, and standards (e.g., National Institute for Standards and Technology (NIST)) are another important precautionary measure to provide guidelines for companies. With regards to real-time recovery, component isolation and recovery, real-time monitoring as well as communication and interaction between supply chain partners are some common measures to mitigate the impact of cyber-attacks on the supply chain networks. Finally, aftermath measures such as data backup, resilient infrastructure design, and system restoration are needed to ensure full recovery of the network and to refine the precautionary and real-time recovery plans.
12. Industrial cybersecurity with IT and OT integration

Operational technology (OT) cybersecurity has gained increasing attention since 2010. To handle recent cyber events that have driven the need for more regulations and measures to combat cyber-threats in chemical industries, major chemical companies, such as Dow, have developed very significant cybersecurity programs. For example, Dow established its first generation cybersecurity program in 2017 and has greatly improved the program in the following years to keep pace with the evolving threats. While this article has mainly addressed cybersecurity concerns in OT space, it is noted that both IT and OT are utilized in industry to develop cybersecurity solutions to protect software, hardware, infrastructure, people, and data. It is important for process engineers to understand both IT and OT cybersecurity landscapes to be able to develop frameworks for detection and control/learning system design that integrate the best policies from both domains to create workable solutions. On the one hand, the connection to IT network enables constant monitoring of the performance and condition of equipment and systems, and allows the industrial systems to obtain a more detailed view of individual equipment and conduct a more comprehensive analysis of the entire plant through big data. On the other hand, traditional OT systems do not have cybersecurity features such as encryption and authentication systems for secure data access and the equipment with long life cycles in OT systems cannot be regularly updated with patch systems due to stability concerns. Therefore, to allow for digital modernization of chemical industries,

Fig. 12. Closed-loop trajectories of true states. The two-CSTR process is operated under the decentralized MPC system when surge attacks are added to the temperature sensor $T_1$ of the first CSTR at $t = 0.30$ h and detected by the 2-class FNN detector at $t = 0.32$ h, after which all sensors are switched to their secured back-up sensors and the true process states are driven back to the ultimate bounded region $\Omega_s$ around the operating steady state (Chen et al., 2021).

Fig. 13. Closed-loop trajectories of true states. The two-CSTR process is operated under the decentralized MPC system when geometric attacks are added to the temperature sensor $T_1$ of the first CSTR at $t = 0.30$ h and detected by the 2-class FNN detector at $t = 0.35$ h, after which all sensors are switched to their secured back-up sensors and the true process states are maintained within the ultimate bounded region $\Omega_s$ around the operating steady-state (Chen et al., 2021).
advanced cybersecurity solutions with IT and OT integration need to be developed and broadly implemented.

Compared to traditional IT cybersecurity, OT solutions are unique purpose-built technologies and protocols for systems that have been operated much longer than IT systems. Since upgrades or changes in the OT space generally require plant shutdowns which are not easily done, cyber-assessment and cyber-protection packages with minimum disruption to operations should be developed and deployed at high priority plants. Additionally, the International Society of Automation (ISA) has provided a guidance (i.e., International Electrotechnical Commission (IEC) 62443) for companies to evaluate the cost of any potential attacks from four aspects: consequences, threats, recovery, and investment in the development of OT cybersecurity solutions.

13. Cybersecurity and safety

Since a primary objective of cybersecurity in OT space is to ensure the safe operation of physical assets at all times, there is a clear need to combine safety and security concerns with control systems to handle cyber-attacks on safety-critical systems that have the potential to cause real harm in the physical world. Novel safety and security methods have been developed in important recent works (e.g., (Hashimoto et al., 2013; Ahooyi et al., 2016; Albalawi et al., 2018; Zhou et al., 2019)) to proactively detect attacks and operation hazards as well as to improve safety assurance against cyber-attacks. Furthermore, in Wen et al. (2022b), Amin et al. (2022), Amin and Khan (2022), the authors have developed data-driven methods for process safety analysis and proposed a holistic framework for process safety and security analysis.

Khan et al. (2015) reviews the evolution of the methods and models for process safety and risk management in the last few decades and discusses the current trend of process safety and risk related developments. In El-Kady et al. (2022), the authors provide a review of physical and cyber threats and their defense/protection measures in digitalized process systems. Additionally, Arunthavanathan et al. (2021) provides a review of methods in process safety, where the integration of dynamic fault detection and diagnosis with risk assessment tools has been demonstrated to significantly improve safety in process facilities. Following this direction, many recent efforts have been made in this area to improve process safety. For example, in Kopbayev et al. (2022), machine learning methods have been utilized to detect the fault and take early corrective actions in order to improve process safety. In Cheded and Doraisswami (2021), a novel integrated framework that integrates model-free and model-based methods was developed for fault detection and isolation to tackle the issue of process safety. While both the ordinary process faults due to process upsets or forced-induced faults (i.e., cyber-attack) can lead to safety and security threat in process operations, the main difference between cyber-attack and process fault

is that the intelligent cyber-attacks with knowledge of the plant layout and of the control structure are programmed to disrupt plant operation, which is fundamentally different from ordinary sensor and actuator faults. In addition to fault detection research, the conflicts in human-automated systems due to contradictory observations and actions under cyber-attacks is another deeper and more implicit phenomenon that may bring risks to process safety. To address this issue, human-centered design is important to reduce the occurrences of human-automated system conflict in automation and digitalization of process operations, as pointed out in Wen et al. (2022a).

In the intersection of cyber-security with safety systems and automatic feedback control systems, process safety systems such as alarms systems, emergency shutdown systems, and safety relief devices can provide the last line of defense in the event of an abnormal situation due to cyber-attacks. To prevent the system states from leaving their safety limits prior to the successful detection of cyber-attacks, safety systems can be integrated with control systems to reduce the physical risks of cyber-attacks ranging from simple unplanned downtime in operations to a plant explosion or release of hazardous materials (Wu and Christofides, 2021). To integrate safety systems with control systems, Safeness Index functions $S(x)$, a function of the (closed-loop) process states that characterizes the “safeness” of a process operation, are adopted as a safety metric for the activation/deactivation of safety systems (Albalawi et al., 2017; Wu et al., 2018c; Zhang et al., 2019b). Safe and unsafe operations can then be evaluated by comparing the value of $S(x)$ with the threshold value that is pre-determined using process first-principles knowledge or past plant data. Additionally, because the Safeness Index function can provide information on both measured and estimated states, its use in the alarm system can help manage the trade-off between measuring fewer states (which may lead to missed alarms) and more states (which leads to instrumentation expenses and possibly more occurrences of alarm overloading).

In addition to integrating process safety metrics into the decision making models of safety systems, integrating the actions of safety systems and control systems may be beneficial as well, as pointed out in Wu and Christofides (2021). Specifically, in the traditional process safety paradigm, process variables are stabilized at their set-points by basic process control systems under normal operation; when the control system fails to operate the process in a safe operating region in the presence of disturbances or cyber-attacks, the safety systems (e.g., alarm systems, emergency shutdown systems (ESS), and safety relief devices) are activated to prevent further unsafe operation. However, since the process dynamics is changed after the activation of safety systems (e.g., the opening of a pressure relief valve to prevent high pressure in a chemical reactor), the actions taken by the safety systems should be taken into account in the reconfiguration of control systems. For example, the cyber-secure control system proposed in
the previous section can be integrated with safety systems that take actions based on whether \( S(x) \) crosses the threshold. We assume secure, redundant sensors or reliable state estimations are available to the control, alarm, emergency shutdown, and relief systems with standard industrial practice. Additionally, the actions taken by the alarm, ESS, and relief systems are assumed to be on-off type actions to simplify the discussion. In the case that safety systems are triggered due to cyber-attacks, the safety-based (lower-tier) control system continues to regulate the process state, while the upper-tier MPC needs to switch to secure, redundant sensors or encrypted secure channels to obtain the true state, and update the prediction model to account for the change in system dynamics. The safety system will be taken off-line after process states enter the safe operating region, and subsequently, the two-tier control system switches to the initial process model.

14. Future research directions

14.1. Actuator cyber-attack detection and handling

Similar to sensor cyber-attacks, actuator cyber-attacks also have access to the plant model and controller design details, aiming to diverge the system away from its ideal operating point. However, instead of altering the sensor measurements, actuator attacks modify the direction and magnitude of the control actions without being detected by sensor monitoring tools. Common detection strategies include active detection methods that design excitation signals to be superimposed on the control commands to increase the detectability of the attack and developing an input observer to detect attacks as well as estimating the magnitude of the attack (Muniraj and Farhood, 2019). Unlike the passive detection methods that use regular operation data to determine if the operation is being affected by a cyberattack, active detection methods that apply some perturbation to the closed-loop process system through the control system can actively probe systems for cyberattacks. Conceivably, active detection methods may ensure that a process is free of a wider range of possible cyberattacks than passive detection methods. Future work developing novel active detection methods and, potentially, extending these methods to aid identification and mitigation may prove fruitful. Furthermore, it is noted that certain actuator attacks are undetectable by an observer-based controller (Ayas and Djouradi, 2016); thus, a machine-learning-based detection method may provide new insights. In addition to the detection of actuator attacks, an isolator may need to be developed to identify the affected actuator(s) in the network. Subsequently, to mitigate the effect of actuator attacks, machine learning methods may be utilized to identify cyber-attack patterns and predict future attack actions. Based on that, a resilient control system may be developed to compensate the effect of attacks without having to shut down the entire plant. Additionally, in the case that a safety-critical actuator is under attack, a controller that can operate the system in the presence of actuator attacks needs to be developed to account for the unavailability of the affected actuators due to a physical intervention of maintenance personnel.

14.2. Encrypted control

Since implementing encryption to encrypt-decrypt the communication signals involves the quantization of the signals and calculations using large integers, which may result in significant delays in order to ensure error-free signal encryption-decryption, the encryption-decryption scheme should be tuned to ensure that the calculations can be done with the available computational resources for a specific operating region in the state-space. Once the region of operation size is increased, the computational burden of the encryption-decryption scheme increases as larger deviations from the steady state correspond to larger numbers that need to encrypt using fixed-point operations that are more computationally expensive. This trade-off needs to be carefully studied, and quantitative computational formulas need to be developed to determine how the size of the allowable operating region should be influenced by the presence of potential cyber-attacks such that encryption can be used with allowable computational resources, need to be developed.

14.3. Incorporation of domain knowledge in the design of machine-learning-based cyber-attack detectors

Process dynamics and control strategies can be used to determine the most cost-effective and flexible frameworks for providing security to process networks and computing devices. This is important because an overly conservative cybersecurity policy can impede progress toward an efficient next-generation manufacturing framework; better understanding how the physics of the process help to dictate what types of security measures are required is important for preventing the negative impacts of attacks without getting in the way of process adaptability. For example, the machine-learning detector presented above is built using all the input variables available with an attempt to capture all possible relationships between inputs and outputs. However, in the case of large-scale chemical process networks, several issues may arise if taking all inputs into the training of the detector, especially when the outputs are not sensitive (or are fully decoupled) to some of the inputs or process states. In the example of two CSTRs in series, the states of the first CSTR influence the states (and thus, the dynamic behavior) of the second CSTR, but the states of the second CSTR do not influence the states of the first CSTR. This is important information that can be used as specific constraints on the structure of the machine learning detector for the entire two-CSTR system to improve its sensitivity to noise. Second, the detector structure may become complicated in terms of more layers and neurons in order to find a good approximation between all inputs and outputs, which increases the computational burden required for training the detector both off-line and on-line. Motivated by the above, one method for optimizing the detector structure is to perform an input selection (also termed as feature selection in machine learning) to select a subset of relevant features for use in detector construction using direct information of process structural relationships from process-directed graphs. By carrying out an input selection, the detector structure is simplified, which reduces the training times and avoids the burden of dimensionality. Additionally, another approach to improve the performance of the detector in terms of better prediction accuracy and less computation time is to incorporate chemical process structural knowledge in constructing it. Specifically, constraints will be imposed using process-directed graph information on some of the weight parameters in the detector such that the connected inputs, which have no impact on the output variables, exhibit no correlation to the outputs in the training process of the detector.

14.4. Decentralized learning for data security and privacy

Improving process data security is another important direction, particularly when this data is being operated upon using control laws or machine learning algorithms, to provide flexibility in manufacturing without concerns for data privacy. Therefore, further advances in techniques and frameworks for promoting privacy are needed to provide
tractable solutions for industry. For example, developing a machine-learning-based detector for large-scale distributed systems requires a tremendous volume of data to be collected from all subsystems through various mediums of communication such as Internet and wireless networks, and then processed in a central server or cloud for training. However, as the communication mediums are vulnerable to attackers, the machine-learning-based detector developed in a local server or cloud could be misguided and unable to detect the target cyber-attacks in the presence of data tampering or data manipulation. In addition, as machine learning approaches have been widely used to develop data-driven models for chemical processes that can be incorporated in advanced process control schemes (e.g., MPC), data security and privacy is also of great importance and is gaining increasing attention. While centralized learning can process data and develop machine learning models in a centralized manner for large-scale distributed systems by taking advantage of a high performance computing cluster/cloud, data security and privacy becomes a big issue due to insecure communication links. To alleviate the security concerns, decentralized learning and federated learning methods that distribute a pretrained model to all subsystems, and allow each subsystem to develop and update its own model locally without sharing the raw data with the central server/cloud (AbdurRahman et al., 2020; Li et al., 2020; Ghimire and Rawat, 2022; Khan et al., 2021b). The updated model parameters will be sent to the server for model aggregation, and finally, the updated model will be distributed to all subsystems. The idea of decentralized learning has shown its great potential in developing privacy-aware machine learning models, and needs to be further explored in the development of machine-learning-based detectors.

14.5. Cybersecurity, safety, operation and control: Engaging vendors

The interface of cybersecurity, safety, operation and control will certainly be explored further in the years to come. Despite the recent efforts to detect and mitigate the impact of cyber-attacks on process control systems, the impact of cyber-attacks on process safety has received very limited attention. How plant operators and control systems should work together to safely handle a cyber-attack with minimal performance loss and without costly plant shut-downs is an important question that needs to be studied. Engaging vendors that design, build, and implement safety and control systems to account directly in their architecture and implementation for cybersecurity concerns as well as monitor and analyze evolving cybersecurity threats should be an important consideration and a potential avenue to bring academic advances on the industrial floor. In this context, it is important that the cybersecurity solutions that are implemented in the OT space can work and cooperate effectively with multiple control system platforms developed by different vendors.

14.6. Industrial cases studies

In addition to the simulation studies of chemical reactors discussed in this manuscript, it is also important to implement the operation of the machine-learning-based detector and MPC to handle potential cyber-attacks in a variety of chemical process networks and energy systems (for example, gas pipeline networks). Novel detector-controller architectures need to be developed to improve the robustness of the entire pipeline network to cyber-attacks which is a critical need for the existing US pipeline networks. It is particularly important to build case studies using large-scale process simulators and incorporate as many as possible practical concerns based on direct industrial feedback to test the effectiveness and applicability of the methods developed by academics.

14.7. Cybersecurity awareness education and training

Cybersecurity concerns and cybersecurity mitigation methods are absent from today’s chemical engineering curriculum at both undergraduate and graduate levels. With the digitalization in process operations, the scope of safety concerns has broadened due to the failure of process control and software systems. Therefore, process safety and cybersecurity concerns should be taught in the classroom to educate and train the next generation of chemical engineers in response to the process industry’s emphasis on digital solutions in process operations (Khan et al., 2021a). Process control and process design courses as well as chemical engineering labs could be good starting points to introduce cybersecurity issues to raise awareness of cybersecurity concerns among our students who, in their vast majority, go to work in industry. In addition, the organization of short courses and workshops to communicate recent academic advances of cybersecurity approaches to engineers in industry and inform academics of industrial cybersecurity issues should be pursued. It is important to point out that while the present manuscript addresses OT cybersecurity concerns within a chemical process context, cybersecurity issues are present in all industries employing chemical engineers from chemical to pharmaceutical to food and materials industries.

15. Conclusion

This work presents an overview of recent research results on cybersecurity in process control, process operations, and supply chains. The design and implementation of cyber-defense OT methods including machine-learning-based cyber-attack detection, resilient control strategies, and their integration with MPC, encryption–decryption algorithms, and cyber-secure control architectures were discussed. Chemical process examples were used to demonstrate the efficiency and effectiveness of machine-learning-based detection schemes, and the robustness of attack-resilient MPCs and decentralized MPCs against several most common intelligent cyber-attacks discussed in the open literature. Additionally, an overview of cybersecurity issues in process operations and supply chains was presented, followed by the integration of IT and OT into industrial practices, as well as integrated safety and cybersecurity solutions for safety-critical systems. The paper concluded with a discussion of future directions for academic research, vendor engagement, academia-industry dialogue, and educational needs.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We would like to thank Professor Helen Durand (Wayne State University), Professors Matthew Ellis and Nael El-Farra (University of California, Davis), Professor Victor Zavala (University of Wisconsin, Madison), Dr. Scarlett Chen (Google), Mrs. Connie McAda (Dow), and Mr. Atharva Suryavanshi, Mr. Matthew Tom and Mr. Fahim Abdullah (University of California, Los Angeles) for their research results and contributions to this manuscript. We would also like to gratefully acknowledge the financial support from the US National Science Foundation and the US Department of Energy.