By Meher Preetam Korukonda, Mohammad Shahidehpour, and Le Xie

Cyber Resilience in Virtual Power Plants

A multiscale multilayer approach toward secure energy management.

A CYBER-RESILIENT CYBERPHYSICAL SYSTEM (CRCPS) is designed to continue operating safely and reliably even in the presence of cyberattacks, failures, or unexpected disruptions. Unlike a regular CPS that focuses purely on preventing and mitigating cyberattacks, a CRCPS emphasizes the system's ability to withstand, recover from, and adapt to adverse conditions while maintaining its critical functionalities. This approach addresses not only how well the system can defend against cyberthreats but also how quickly it can recover and resume normal operations after an attack or failure.

A CRCPS is designed to resist cyberattacks by incorporating strong cybersecurity measures. However, resistance alone is not enough; the system must also anticipate potential failures and prepare to withstand them. Furthermore, in the event of an attack or failure, it should degrade its services or functionalities in a controlled

Digital Object Identifier 10.1109/MELE.2025.3535023
Date of current version: 28 February 2025

©SHUTTERSTOCK.COM/METAMORWORKS

manner rather than fail completely. This ensures that critical operations continue at a reduced capacity without causing catastrophic failures. For example, in a smart grid, if part of the network is compromised, the system reroutes power to ensure that essential services, such as hospitals or emergency responders, continue to receive electricity.

CRCPSs are equipped with advanced monitoring and detection tools that enable them to quickly identify anomalies, cyberthreats, or system malfunctions. These systems automatically respond to contain threats and prevent further damage. After a cyberattack or failure, a resilient CPS must have the mechanisms in place to quickly restore operations to normal. This includes backup systems, redundancy, and recovery protocols that ensure minimal downtime and data loss.

Cyber resilience is often achieved through proper design of redundancy in both the cyber and physical components. Multiple pathways or backup systems are available so that if one part of the system fails, others can take over without disruption. A cyber-resilient CPS incorporates security across all layers of the system, from sensors and actuators in the physical layer to software, data, and communication networks in the cyber layer. Every component of the system must contribute to overall resilience. For instance, in a smart factory, every piece of equipment, from the control systems to the machines themselves, has security measures in place, such as encrypted communications, secure access control, and tamper detection.

Instead of relying on a single defense mechanism, a CRCPS uses a layered security approach that combines multiple techniques such as encryption, authentication, intrusion detection, access control, and cyberphysical detection and mitigation of such attacks. This increases the system's ability to resist and recover from attacks. Even when under attack, a resilient CPS prioritizes maintaining operational continuity, particularly for critical processes. The system should ensure that any disruptions are minimized and that high-priority tasks continue to function without interruption.

A CRCPS includes comprehensive incident response plans that detail how to respond to various cyberattacks or system failures. These plans ensure quick decision making and resource allocation to mitigate the impact and recover from an incident. A smart energy grid has predefined protocols for responding to cyberattacks that target control systems. These protocols include automatic isolation of the affected section, rerouting power, and activating backup generation systems.

In summary, a CRCPS must have the following characteristics, as depicted in Figure 1:

- It must anticipate potential risks, both known and unknown, and be prepared to address them before they can cause damage. This involves proactive threat intelligence, risk assessments, and scenario planning.
- Even under severe cyberattacks or technical failures, a CPS must ensure that its most critical functions survive and remain operational. This might include shutting down nonessential functions to preserve vital operations.
- It must be agile enough to adapt to changing circumstances in real time. This includes reconfiguring components, rerouting data flows, or adjusting operational parameters based on the current threat landscape or operational conditions.
- The system must be built to withstand both random failures and targeted attacks without losing functionality. This can be achieved through redundant hardware, multiple communication pathways, and diverse defense mechanisms.
- Cyber resilience involves rapid recovery after a cyber incident, including restoring normal system functions, repairing damage, and learning from the incident to improve future resilience.

Boosting Cyber Resilience in Virtual Power Plants

In the pursuit of building secure and resilient virtual power plants (VPPs), little is known about the security risks in this new form of CPS. An effective solution requires joint efforts from both engineering and human science aspects. On one hand, we need an efficient engineering approach to secure the cyber resources (e.g., communication networks), which are often transparent to most grid applications by design. The solutions must address the unique challenges of VPPs (e.g., no interference to real-time operations) and also leverage their opportunities (e.g., a relatively fixed infrastructure). On the other hand, individuals working within a cyber system

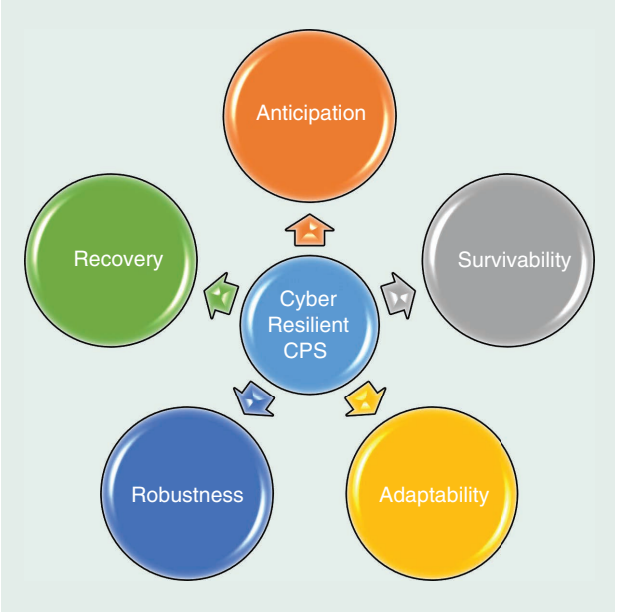

Figure 1. The key features of a CRCPS.

(human in the loop) play important roles beyond autonomous operations and remain a critical, and possibly the weakest, link in securing the cyberspace. We must also study the impact of humans (i.e., operators, in this case) on system security.

The focus of this article is on the cyber resilience of VPPs. The various control objectives in a VPP are enforced through a systematic hierarchical control structure, which is structured across several layers to effectively coordinate and optimize distributed energy resources (DERs), including renewables, batteries, and demand response assets. In a VPP, a sensor layer that supplies data to various control layers is present. In addition, a VPP also encompasses fast primary controllers for individual DERs that operate in the microsecond timescale, followed by an upstream aggregate controller that ensures coordinated control of the multi-DER in the millisecond timescale. The aggregator gateway, operating in the millisecond-to-second timescale, serves as a utility-control-center client and as a server to VPP facilities. The exchange of data between different layers of operation in the VPP is facilitated by an underlying communication network (see Figure 2), which needs to be shielded against any type of cyber intrusion.

Even in the event of a cyberattack, a fail-safe control action must be initiated after timely detection to provide uninterrupted power, at least to critical infrastructures. On the VPP system level, emerging technologies like software-defined networking (SDN) and blockchain enhance the security of tertiary controls through system correctness verification, security policies enforcement, vulnerability detection, and threat mitigation.

Even when under attack, a resilient CPS prioritizes maintaining operational continuity, particularly for critical processes.

On the individual level, investigating the relationship among individual differences (personality traits, interests, and attitudes) between different system operators and cyberattack detection and response is an interesting direction that should be analyzed. These studies can lead to personnel selection and training plans that promote cyber agility and accuracy of the human-in-the-loop component of VPPs.

Thus, the fundamental innovation that is needed for cyber secure and resilient VPPs against a baseline traditional power grid is the development of a multiscale multilayer anomaly detection (AD) and mitigation framework (see Figure 2), which is discussed in greater detail in the next sections.

Safeguarding VPP Sensor Data

In VPPs, the sensing layer is intricately linked to various control layers that are responsible for regulating the output voltages and currents of DERs. This sensor layer typically exhibits inadequate cybersecurity capabilities due to a mixture of legacy systems and/or resource-constrained computational assets that may make them vulnerable to physical and remote attacks. Physical manipulation necessitates that an attacker possesses physical access to the targeted inverter or converter. The attacker can then inflict direct harm on controlling sensors, such as smart meters, or introduce data that disrupt sensor readings, including electromagnetic

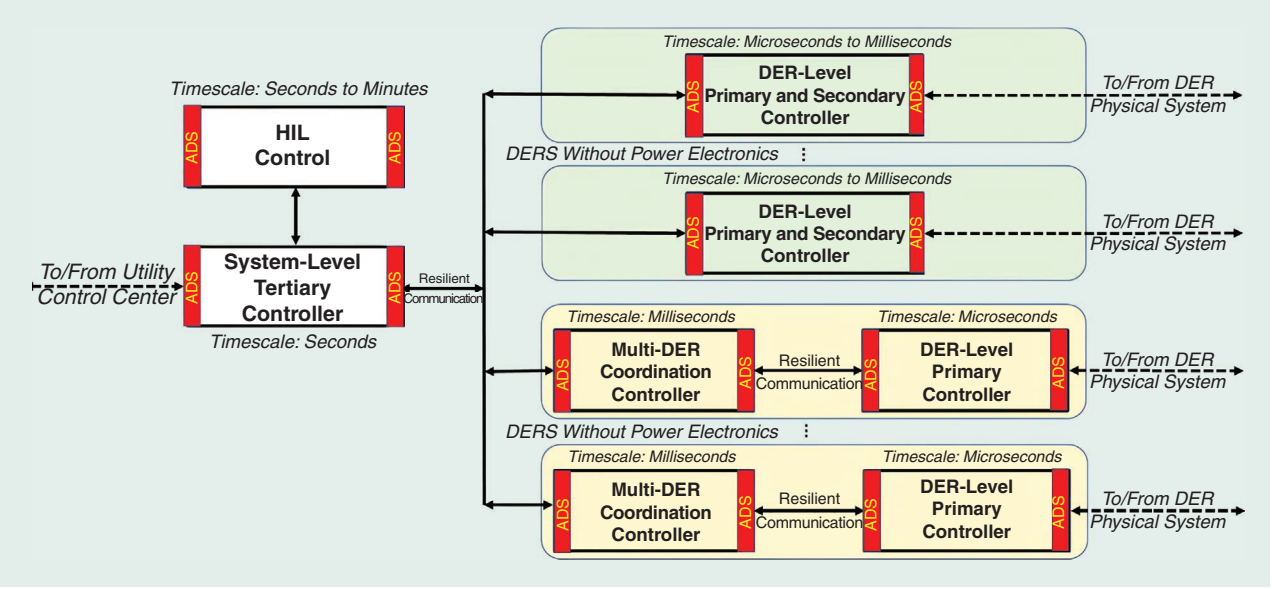

Figure 2. A multiscale multilayer resilient control scheme with an ADS for VPPs.

actuators. Remote manipulation may be achieved via a man-in-the-middle attack, which entails fabrication of network nodal identities to intercept, alter, insert, or discard data within the communication channel. If the firmware update data are compromised by a malicious actor, a man-in-the-middle attack could potentially alter sensor data, resulting in system malfunction. Various types of man-in-the-middle attacks exist, including both replay and data injection. Replay attacks frequently entail adversaries duplicating intercepted legitimate data transmissions, presenting a substantial risk to information security. Data injection attacks involve the introduction of noise into sensor data through lower-order harmonics. An unobservable fake data injection constitutes a sophisticated cyberattack in which an attacker aims to circumvent detection by introducing measurement data that closely resemble authentic data, hence jeopardizing the stable operation of the system.

The cyber shield attack detector uses the difference between the expected and actual measurements with a dynamic watermark to determine the presence of a cyberattack.

Broadly speaking, measurement-based AD mechanisms can be widely classified into the following categories: detection using relevance in spatial and temporal data, securing critical sensors, and using watermarking mechanisms. In temporal detection, the current state of the system is estimated by using historical measurements. An example of doing so is using a Kalman filter. In spatially relevant detection, the information from redundant sensors is utilized for cross-checking the validity of the measurement data. Here, the estimation is carried out only using real-time measurements. Another method for preventing cyberattacks is to provide additional computational resources to critical sensors and protect them. The aforementioned methods may be facilitated by increasing the computational resources available at the edge. Such resources may enable the deployment of advanced attack detection algorithms and/or cryptographic tools that can help increase sensor-level trust. High-trust sensor streams can then be used to accurately reconstruct the system's state. Watermarking is another technique that is suitable for deployment, especially at the sensor level, due to its low real-time computational requirement. The concept is that by injecting a known noise as a probe input of the system, an expected effect of such input should be found in the true measurement output due to the system's dynamics. Figure 3 shows how a dynamic watermark is carefully added to the modulation index of a solar inverter. The cyber shield attack detector uses the difference between the expected and actual measurements with a dynamic watermark to determine the presence of a cyberattack. A combination of these techniques can be explored to provide effective detection against sophisticated cyberattacks in the sensing layer.

Securing Hierarchical Control Operation of VPPs

The operation of DERs and loads in a VPP requires that control schemes be able to appropriately address the different dynamic requirements and objectives in the grid. This hierarchical control architecture of a VPP is used for the incorporation of other particular economic, technical, or environmental objectives, which set the participation of a VPP. One of the reasons for adopting a hierarchical control structure is the hierarchical nature of the operational objectives of VPPs. The economic objectives might be achieved via real-time pricing, critical-peak pricing,

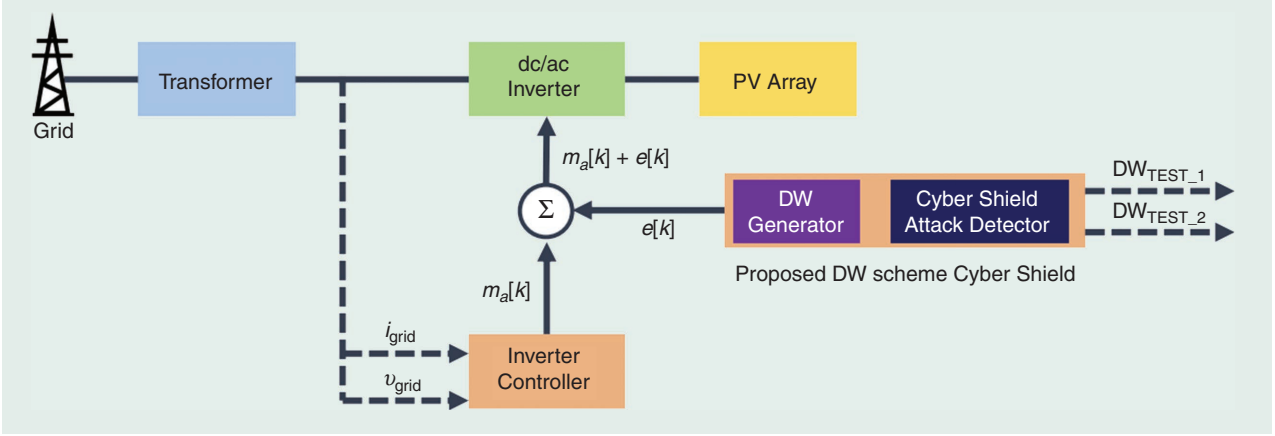

Figure 3. A watermarking implementation framework in a solar farm for detecting cyberattacks. PV: photovoltaic; DW: dynamic watermark.

time-of-use rates, or incentive-based demand. The environmental objectives might include an operation of the grid with lower emissions to meet the environmental targets set by regulatory schemes. Similarly, the technical objectives might include optimizing network operation by minimizing power losses, regulating voltage and frequency variations, and improving power quality. All three different control layers—primary, secondary, and tertiary—are designed for specific objectives based on the timescale at which they need to be fulfilled. The interlinking data transfer among these layers of control provides ample opportunity for attackers to destabilize the entire system. In this section, we discuss the security risks and possible solutions that are specific to different levels of control in VPPs.

The secondary control layer groups multiple DERs into localized clusters to maintain balanced power quality and coordinate energy flows within a specific area.

Primary Layer

At the primary control layer, each individual DER operates independently to maintain stability by managing local parameters like voltage and frequency. This local control is fast and essential for responding to sudden fluctuations in generation or demand at the device level. It rapidly generates the desired control sequences for the DER inverters guided by local measurements and global consideration, including the coordinated operation of the respective VPP and the dispatch consideration of the aggregate multi-DER control layer. Clearly, if the coordination is affected due to a cybersecurity breach, it will have a direct impact on the local DER control, where the manifestation may be small but rapid in a shorter time frame but aggregates quickly over time given that the inverters operate at a fast scale. On the other hand, this layer is also susceptible to attacks on the sensor level. At the aggregate multi-DER control layer, the manifestation of this cybersecurity, depending on the number of DERs affected, could be appreciable on a spatial scale (due to aggregation) but on a slower temporal scale.

At the primary VPP control layer (i.e., the DER level), the core vulnerabilities of DERs to cyber breach are potentially multifold. The inverter's resilient and stable operation is vulnerable to information manipulations and interruptions between the aggregate control and primary control layers. Control schemes are typically formulated in the information-centric domain; however, handshaking of this information among control nodes in a distributed control scenario needs to follow communication protocol(s) that are conventionally data oriented and that focus on structured packetization of the data for node-to-node or node-to-multinode transfer. When such a packet is corrupted during transmission, it introduces delay because retransmission is required and/or error-correction coding needs added computation time to receive the data at the receiving node if possible. So, a key challenge in this layer of VPP control is the impact of loss of data, or the loss of information in the data packet and how the primary control mechanisms that yield resiliency and stability to such loss in VPPs can be devised.

As such, any control resiliency needs to recognize the manifestation of security breach on the spatiotemporal scales. There have been several works that focus on coordinated control of power electronics for partitioning as well as mitigation of the variable latencies associated with communication delays. One of the interesting observations critical to cybersecurity at this layer is the fundamental difference between data and information transfer. What it essentially means is that a disruption in information flow, and not necessarily data flow, may have a dominant influence on the system's control performance. In other words, if the coordination control of a VPP is information based, then, one may be able to reduce data-exchange needs and instead communicate only when tangible control information changes occur, thereby reducing the susceptibility of the communication network to security breach.

Considering that a typical DER inverter in a VPP operates with a switching period of roughly 50 μs, information manipulation, such as the delay margin for the primary controller, is limited. In such cases, an event-triggered, resilient self-learning control with minimum communication becomes useful for AD and mitigation. Such a control strategy will ensure reduced dependency and susceptibility to communication failure during a security breach and further leverages spatiotemporal multiscale AD to adjust the DER control modes and set points to ensure fault tolerance and/or self-healing with minimal communication. Although the AD system (ADS) detects the abnormality of the DER controller, it is important to note that such an information manipulation and interruption may happen due to either a compromised aggregator controller and/or a compromised communication channel between the DER's primary-layer control and the aggregate control layer that aims for coordination among multiple DERs.

Secondary Layer

The secondary control layer groups multiple DERs into localized clusters to maintain balanced power quality and coordinate energy flows within a specific area. This layer provides load balancing and synchronizes the output from

various DERs, enhancing reliability and reducing the impact of fluctuations on the broader grid. By managing clusters of DERs together, secondary control ensures consistent, high-quality power output, adjusting for local demand changes and enhancing the resilience of each cluster. It also reduces the voltage and frequency deviations that occur at the primary layers due to the droop-based power control reference's reference signals, which are generated at the primary layer.

The control references in the secondary layer are generated by using multiagent techniques that rely heavily on a distributed communication structure among various DERs to provide a coordinated control solution. A cyberattack may be introduced to the VPP infrastructure through the communication medium, aggregators, and coordination control layer to interrupt its harmonious operation. A severe cyberattack typically spreads throughout the grid gradually (i.e., stealthy-attack scenarios) to make the detection of such an attack extremely difficult at early stages by using conventional protection schemes and single-layer AD mechanisms.

It is important to note that the low-inertia characteristic of VPPs makes them all the more vulnerable to anomalies, which could result in frequency and voltage instability if malicious activities are not detected in a timely manner for protection and the healing process. In this context, stealthy attacks in which the effect is not immediately apparent to the controller can cause catastrophic events. Stealthy attacks target the control signal of a system, and the effect is not immediately apparent to the controller and the ADS, while causing some of the system states to diverge, which results in instability of a VPP. This may result in a situation where the system's divergence is kept hidden until a time that the system is operating at a tipping point when individual DER controllers start to fail, causing a chain-reaction-effect attack; thus, a conventional intrusion detection mechanism for measurements is theoretically rendered ineffective. Thus, there is a need for a mechanism that prevents malicious operation active and reactive power (PQ) set points for DERs that are induced by a stealthy intruder who breaches the aggregate multi-DER coordinated control layer of a VPP.

The goal at this layer is to reduce the probability of success of stealthy-attack scenarios and mitigate their potential harmful impacts at the multi-DER coordination control layer of a VPP. This is achieved by introducing a guard control unit (see Figure 2) with decision-making capability in the primary DER coordinated controller, which ensures resiliency during an event at an early stage and consequently minimizes the attack surface with minimal communication requirements. The guard control unit includes a real-time normal operation region identification and trajectory operation estimation for a cluster of DERs in a VPP with the ability to pinpoint the misbehaving subset of DERs in timely manner. Optimal coordinated corrective actions for autonomous voltage and frequency restoration may be developed by reconfiguration in the operation mode and set point of the healthy DERs. There is a scope for developing autonomous frameworks for adjusting system parameters that introduce imperfect knowledge of the system model for the intruder who minimizes the stealthy-attack surface, thereby minimizing the probability of success of the stealthy-attack scenarios.

Most of the ADS solutions available at this level are focused on three distinct directions. One is the identification of already-available attack signatures, but it suffers from a lack of agility in the context of continuously adapting attackers. The second is represented by neural networks and usually requires large sets of labeled training data that cover all the possible fundamental attack strategies. The third is focused on time-based stochastic models, like Markov ones, and makes implicit (and often unrealistic) assumptions regarding the stationarity of the modeled system. The ADS's challenge at the multi-DER coordination control layer of a VPP is due to DER dynamic variations, which are influenced by the multitude of heterogenous devices, which in turn are influenced by external factors and the severely underdetermined and unobservable character of the system. It is important to design a real-time safe-operation region identification framework for networks of DERs in a VPP domain supported by machine learning approaches that can predict the behavior and trajectory operation of a subset of DERs for detecting malicious activity at aggregate control layers and consequently provide corrective actions for healing the DERs' coordination. The proposed guard control unit and decision-making module, with its trajectory operation prediction and normal operation region assessment, provides another layer of resilience that is supported by multilayer AD.

Tertiary Layer

Cyberattacks at the tertiary control layer can have serious implications for a VPP's ability to participate in system operations and electricity markets, impacting both their operational effectiveness and financial performance. The tertiary layer, responsible for overall systemwide optimization, forecasting, and market participation, is particularly vulnerable to cyberthreats because of its central role in economic dispatch, scheduling, and demand forecasting. Successful cyberattacks on this layer disrupt these core functions, leading to misinformed decision making, loss of efficiency, and compromised reliability in both market and system operations.

For instance, an attack that corrupts demand forecasts or manipulates economic dispatch decisions can cause a VPP to misallocate resources, resulting in inefficiencies and imbalances that can destabilize the grid. If a VPP's dispatch schedules are altered or delayed, it may fail to meet real-time demand, leading to financial

penalties, increased operational costs, and reputational damage. In competitive electricity markets, accurate forecasting is critical for positioning VPPs for profitable participation. A cyberattack that injects false data or delays communications could cause a VPP to submit erroneous bids or fail to respond to price signals, impacting its profitability and credibility as a reliable market participant.

Furthermore, attacks at the tertiary level can undermine trust in a VPP within the energy market and among stakeholders. Regulatory bodies, market operators, and other market participants may view a compromised VPP as a risk, potentially leading to increased scrutiny or limitations on its market access. This layer is responsible for aggregating and securely dispatching large amounts of DER power to the grid, and a compromised VPP can become a point of vulnerability in the energy system, with repercussions that impact market stability and grid reliability on a larger scale.

In practice, current solutions to security issues at this level are dominated by applying commercial off-the-shelf Internet security techniques, such as firewalls and antivirus or antispyware software, to secure power grid control systems. However, those security solutions can only provide fine-grained protection for single devices. Various gaps exist, including systemwide visualization, real-time monitoring capability, strictly defined communication paths, and a deny-by-default security model. These gaps can be reduced by applying an SDN architecture in the VPP communication network. Investigating SDN technologies in the context of power grids is a very new but promising research topic. Recent works include applications in substation automation, reliability evaluation, quality-of-service optimization, and fast failover mechanism. However, those works are still in the very early stage and do not particularly focus on security. The unique features offered by SDN, such as global visibility and direct programmability of network control, help with developing effective and efficient means to detect and mitigate cyberthreats. Specific research tasks in this context include building 1) novel intrusion detection systems and 2) innovative cross-layer verification frameworks. Figure 4(a) summarizes many known cyberattacks and their implications in the context of a VPP, while Figure 4(b) lists the components and benefits of a VPP transformed through application of SDN technology.

Apart from SDN, deploying blockchain technology at the tertiary layer has great potential for enhancing cyber resilience by providing a secure, decentralized, and transparent framework that safeguards data integrity, authenticates transactions, and enables automated operations without a central point of failure. In this layer, blockchain can record and manage economic dispatch, demand response, and energy storage operations transparently. It can also enable secure and decentralized bidding in energy markets, allowing a VPP to trade energy with the main grid while ensuring that transactions are verified and auditable. Smart contracts on the blockchain can automate trading and dispatch decisions based on predefined conditions like market prices or energy availability, making operations more efficient and resilient against external tampering. Additionally, blockchain can facilitate secure and compliant data exchanges of a VPP with external entities. It provides a transparent ledger of all grid interactions, including compliance reporting, the ancillary services provided, and inter-VPP collaborations. Blockchain's immutable record enhances regulatory trust and enables a VPP to offer grid services such as frequency regulation or emergency support, with assurances of data integrity and transaction security.

Human-in-the-Loop Approach to Cyber Resilience

Human behaviors within a cyber system and their impact on system security are increasingly gaining traction in cyberphysical energy systems. An emerging stream of

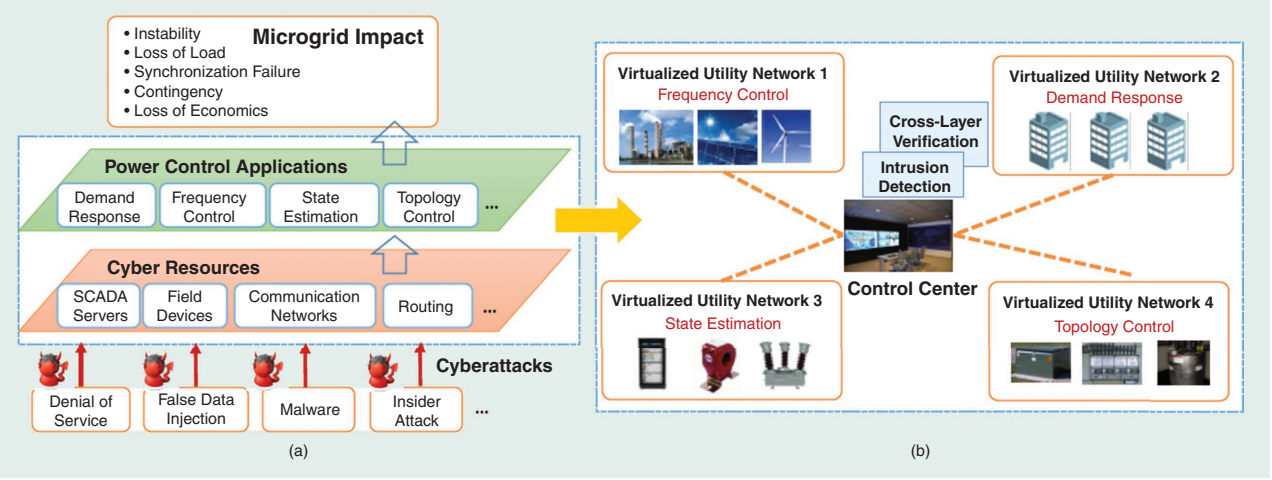

Figure 4. The transformation to a cyberattack-resilient VPP with SDN Technology. SCADA: supervisory control and data acquisition.

research investigated how individual characteristics of end users and inside employees, particularly their personality traits and attitudes (technology acceptance), influence security-related behaviors. A recent survey of network operators strikingly shows that 89% of operators are never sure that their configuration changes are bug-free, and 82% are concerned that changes would cause problems with existing functionality. Therefore, in addition to engineering solutions to boost VPP infrastructure protection, there is a critical need to examine how a human in the loop impacts VPP security and how to enhance this link.

This emerging line of research can be furthered in three important ways. First, although existing solutions focus on traditional forms of cyberspace (e.g., the Internet), this can be extended to critical infrastructures, in particular, VPPs, by identifying potential security risks and the role of operators for system security. Second, past research on the effects of individual differences of human in the loop focused exclusively on personality traits and largely overlooked other individual factors that are potentially predictive.

Personal interests were recently shown in the organizational psychology literature to be powerful predictors of job performance with effects over and above those of cognitive ability and personality traits. Cognitive and emotional states (e.g., stress and cognitive load) were shown in the human factors literature to heavily influence safety behaviors. We can apply the latest advances in organizational psychology, personality psychology, and human factors into cybersecurity research and examine the influence of individual interests, personality traits, and attitudes as well as cognitive and emotional states on operators' competency in reacting to potential cyberattacks. Third, although existing studies focus primarily on the impact of individual differences on end-user security behaviors (e.g., change the default password) and inside employee compliance or violation behaviors, future research can focus on examining the behaviors of VPP operators, and specifically, their competency in detecting and counteracting potential cyberattacks to the system.

By using historical data and experience, it is possible to develop a series of computerized scenarios to simulate real situations (both normal operations and malicious cyberattacks) that system operators will potentially encounter. Specific behavioral guidelines for system operators can be developed for all the scenarios, incorporating both the cybersecurity and organizational science perspectives. Developed computerized scenarios and behavioral guidelines can be presented to a group of system operators in industries like ComEd. By using a combination of survey and controlled experimental studies, it is possible to examine the effects of 1) personality traits (using Big Five personality factors and narrow facets within the Big Five), 2) vocational interests, 3) attitudes (technology acceptance), and 4) cognitive and emotional states on system operators' competencies to 1) accurately detect malicious cyberattacks, 2) react immediately to cybersecurity threats, and 3) comply with security behavioral guidelines during the process of handling cyberattacks.

An interesting direction is testing and seeing how relatively stable individual difference variables effect influence operators' competencies in securing a CPS, and how attitudes and cognitive and emotional states moderate the effects of individual differences on system operators' cybersecurity competencies. Findings on desired individual differences attributes can be used to develop selection and staffing plans to ensure that system operator candidates have the strongest cybersecurity competencies. Findings on the effects of attitudes can further be used to develop training plans that improve system operators' cybersecurity competencies. Findings on the effects of cognitive and emotional states can then be used to develop job designs to optimize system operators' functioning in a CPS.

Integrated Cyberphysical Analysis and Testing

A CPS combines physical processes with computational and communication systems, requiring specialized testing methodologies and platforms to ensure their security, reliability, and performance. The testing methodologies for a CPS need to account for the interaction between both cyber (software and networks) and physical (hardware, sensors, and actuators) components, and their interdependence. This section deals with the cyberphysical testing methodologies and platforms that are used to evaluate such systems.

Augmented Risk Assessment

In current practice, a cybersecurity risk assessment and a contingency analysis for VPPs are considered two separate tasks and are normally conducted by two different groups of experts. A traditional VPP cybersecurity risk assessment, which is performed by cybersecurity engineers, focuses only on IT security configurations such as firewalls, antivirus software, operation system and application patches, user access, and so on. On the other hand, a VPP contingency analysis i.e., operational tasks (OT) tasks, are performed by system operators by simulating scenarios that include physical contingencies due to physical faults and operational errors. The convergence of IT and OT, however, demands an integrated risk assessment due to the increasing security interdependencies. Malfunctioning DER firmware, for example, may be considered a low-risk threat in traditional risk assessment but could cause a cascading failure between IT and OT systems due to the loss of generation–demand balance and system stability. In this sense, the cybersecurity risk assessment must be augmented to incorporate a physical contingency analysis to accurately quantify the risk of cyberthreats. It is important to develop a proper risk assessment framework that can evaluate the risk of cyber contingencies as well as a

combination of cyber and physical contingencies. It will greatly improve the understanding of a VPP security posture under a complicated IT/OT environment and provide valuable insights for cyberattack detection and mitigation.

A cybersecurity risk assessment framework should be developed that leverages the physical principles derived from the physical models as well as the statistical patterns derived from the data stream to achieve a holistic risk assessment in the era of IT/OT convergence. As shown in Figure 5, such a framework will identify the cyber, operational, and physical and synthetic vulnerabilities that involve both IT and OT. It will perform a contingency analysis by using a simulation-based approach to evaluate the negative impact of identified vulnerabilities and identify the assets and systems that are critical for the overall security posture. The temporal change of parameters, such as operation conditions, control mode switching, and renewables forecasts, should be considered in the dynamic risk assessment framework. An augmented attack tree model can be developed to incorporate the output of both vulnerability and contingency analyses, and then provide the necessary information for risk assessment.

Cosimulation-Based Hardware-in-the-Loop Testbed Design

In any CPS, the power network and its components, along with information and communications technology infrastructure, constitute two elements of a broader, heterogeneous system. Communication networks, like all digital systems, are represented as a series of discrete events (e.g., transmitting and receiving packets, packet buffer overflows, and so on), whereas power systems are generally modeled as continuous-time functions that utilize differential-algebraic equations despite the occurrence of discrete events in power systems when the statuses of breakers, switches, and relays change. Cyberphysical testing methodologies and platforms are essential to ensuring the security, reliability, and performance of such complex, interconnected systems. These testing approaches, ranging from hardware in the loop (HIL) to cosimulation, provide comprehensive ways to evaluate both the cyber and physical aspects of systems like smart grids.

Historically, interdomain interactions have been investigated by meticulously replicating the subsystem of interest for a certain domain while simplifying the other components. Such treatment, however, may result in issues as the absence of depth in the system's features neglects fundamental interconnections. Coupled simulations, also called *cosimulations*, meet functional needs by modeling multidomain systems through the use of different simulation tools that work together as a single simulation platform for analysis. A cosimulation consists of a collection of interconnected simulators that collaborate with one another. Each simulator possesses a distinct solution and operates concurrently and autonomously on its respective model. By dynamically linking the input and output variables of the simulators, it is possible for one simulator to use its output as an input for another and vice versa. Cosimulation technologies with real-time emulation capabilities deliver a high degree of realism and facilitate HIL simulations.

HIL simulations integrate the actual dynamics of physical power systems with the control and adaptability of computer-based models. A primary advantage of HIL tests is their capacity to assess the performance of novel technologies in authentic operational settings. HIL testing provides a more precise assessment of novel technologies through the integration of physical apparatuses. Researchers and industry professionals widely use platforms such as RTDS, OPAL-RT, GridLAB-D, and MATLAB/Simulink to conduct these tests in real-time, simulated, or hybrid environments.

The majority of cyberphysical testbeds that have been built to date possess either cosimulation or HIL capabilities. Some studies have considered both issues. Nevertheless, they are mostly in the transmission system domain. There is a need to construct a comprehensive HIL cosimulation testbed to evaluate the effects of cyberphysical contingencies in a networked multi-VPP scenario as well as the efficacy of various detection and mitigation strategies that are designed to improve cyberphysical resilience.

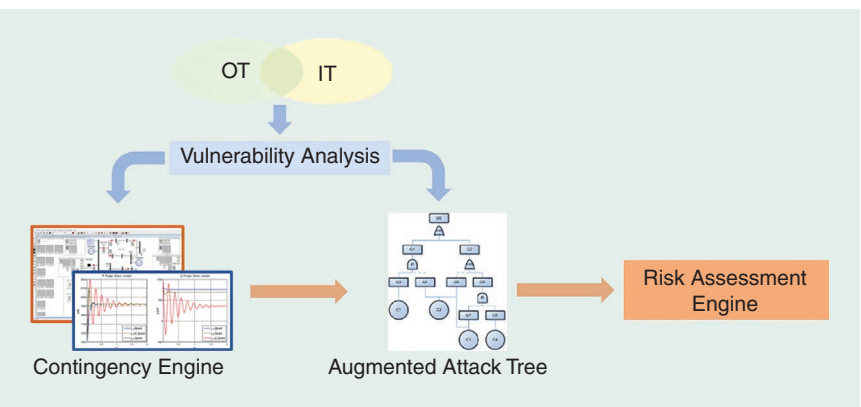

Figure 5. An augmented cybersecurity risk assessment considering IT and OT.

Proof of Concept: Data Replay Attack on a Solar Farm

In this section, we describe our experience with the detection of a data replay attack on an HIL model of a solar farm connected to a distribution grid. The attack is carried out on the sensor data of the solar inverters to destabilize functionality of the solar farm. Timely detection of this attack by using the proposed dynamic watermarking technique will fortify the sensor layer and ensure cyber resiliency of

the distribution grid. Our system is a 2-MW solar field connected in parallel with distribution loads to the substation, as shown in the schematic in Figure 6(a). The substation is connected to the solar farm and the loads via a 12/16/20-MVA 69 × 12.47-kV transformer. The dc power generated by the solar farm is converted into ac power with the help of a solar inverter. This facility provides enough power for approximately 600 customers annually. The solar field has 7,784 solar panels, each with the capacity of generating 350 W at 38.9 V.

Using our facilities at Illinois Institute of Technology (IIT), we designed an HIL testbed that replicates the exact operation of the solar farm. Figure 6(b). shows the schematic of the testbed that was established in the Galvin Center for Electricity Innovation at IIT. The solar farm is emulated in real time by using a SynDEM research kit, while the distribution system with the substation and loads are modeled inside RTDS. The SynDEM smart grid research kit is a multifunctional power electronic converter. The kit can be reconfigured to obtain different power electronic configurations that are versatile for different applications, and uses automatic code generation tools to code in different control algorithms directly through MATLAB/Simulink.

On the hardware end, there are two SynDEM research kits: one configured as a grid simulator and the other as the solar farm. At the point where the solar inverter is connected to the distribution system, the three phase

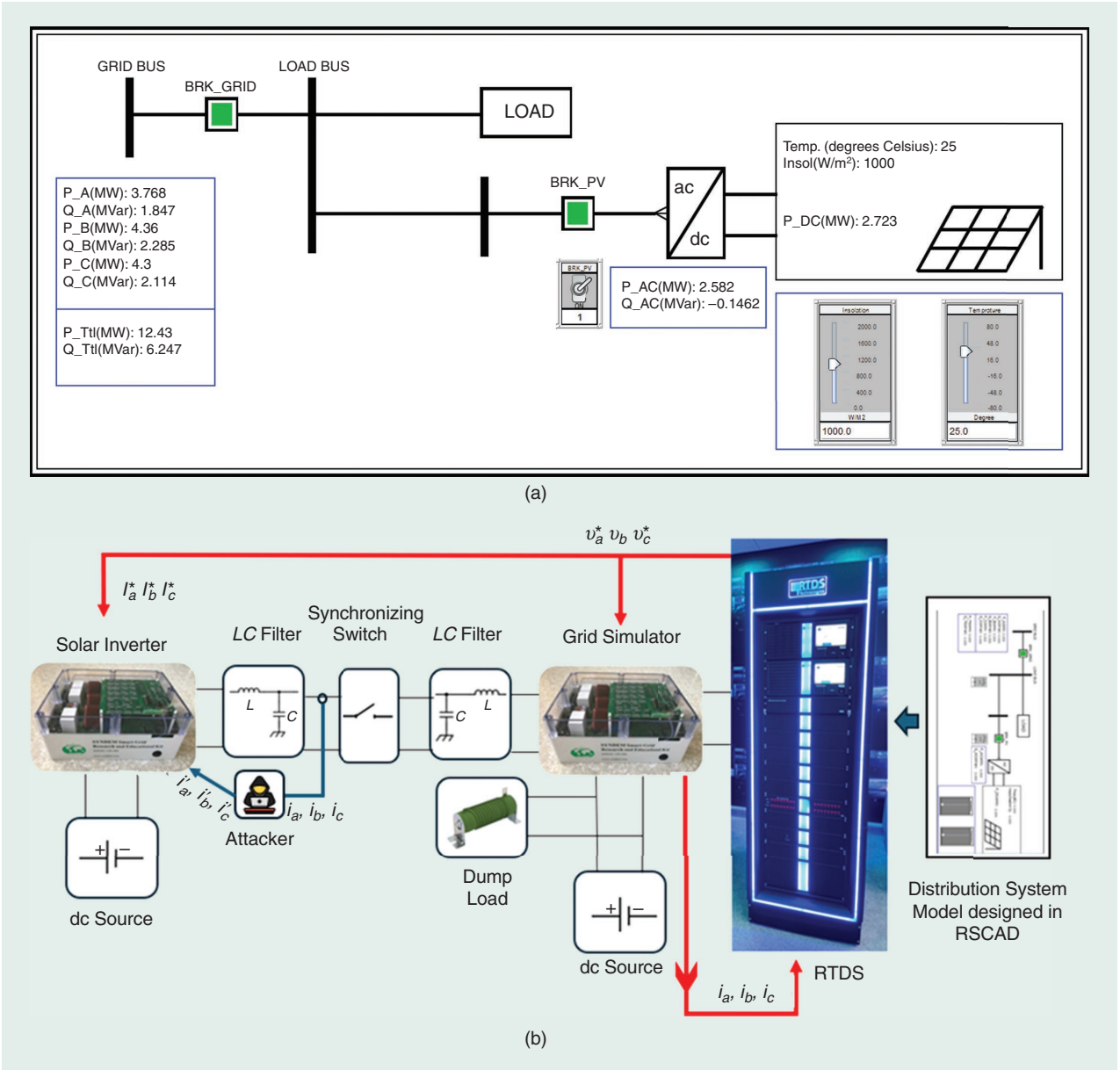

Figure 6. (a) A schematic of the HIL testbed of the grid-connected solar farm. (b) A schematic of the solar farm connected to the distribution system. temp: temperature.

voltages, v_a^*, v_b^*, and v_c^*, are recorded from the RTDS simulation and sent as reference voltages the grid simulator by using the analog output port (GTAO). The grid simulator uses a voltage controller to mimic the reference signals in hardware. Another SynDEM research kit is used to function as the solar farm, on which a current controller is designed to generate the necessary power output from the solar panels. The SynDEM kit does not have the inherent capability to function as a solar inverter by itself. We achieved the characteristics of the solar inverter by communicating real-time current reference signals I_a^*, I_b^*, I_c^* from RTDS to the solar inverter based on the real-world data collected on the field. The current-controlled solar inverter mimics the current references in real time, thereby replicating the operation of the solar farm. Further, the exchange current values i_a, i_b, i_c between the grid simulator and the solar inverter are communicated by the grid simulator back to RTDS through the analog input port (GTAI). These values are plugged into the RSCAD simulation in the form of current sources to propagate the effect of the external solar inverter in the distribution system.

For emulating the data replay attack on our solar farm-based power hardware-in-loop (PHIL) testbed, we intercepted and manipulated the instantaneous current values being fed to solar inverter's controller. We recorded the current waveforms of the healthy inverter i_a, i_b, i_c for some time and replayed the manipulated feedback signals i_a', i_b', i_c' to the controller of the solar inverter to destabilize it. We implemented this with the help of a delay logic, which is equivalent to replaying the current waveforms after a given time period. Figure 7 shows the B-phase output current of the solar inverter that was recorded before and after the data replay attack. In the initial 5 s, there is no cyberattack, and the output current is very uniform. However, at t=5 s, a data replay attack is initiated on all phases of current measurement. After the attack begins, the output current becomes unstable and starts oscillating from a very high to a very low value. At t=32 s, the cyberattack is switched off and the system comes back to

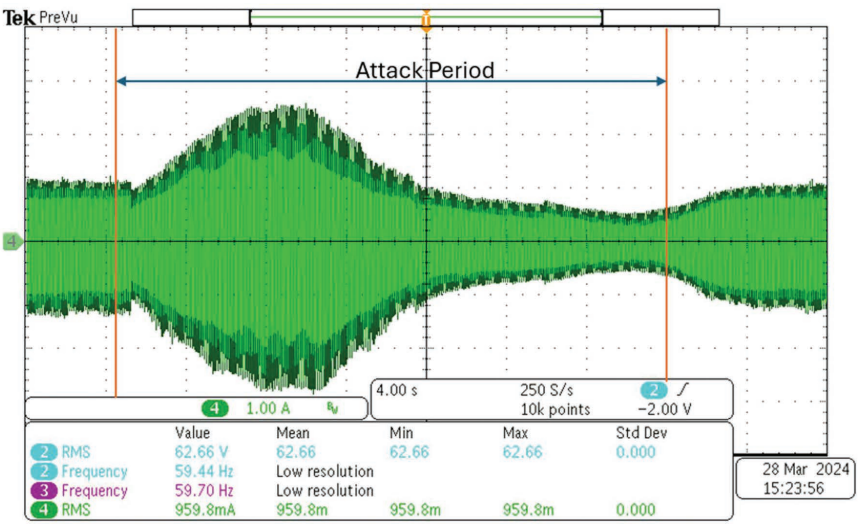

Figure 7. The output current waveform before and after the cyberattack. RMS: root mean square; min: minimum; max: maximum; std dev: standard deviation.

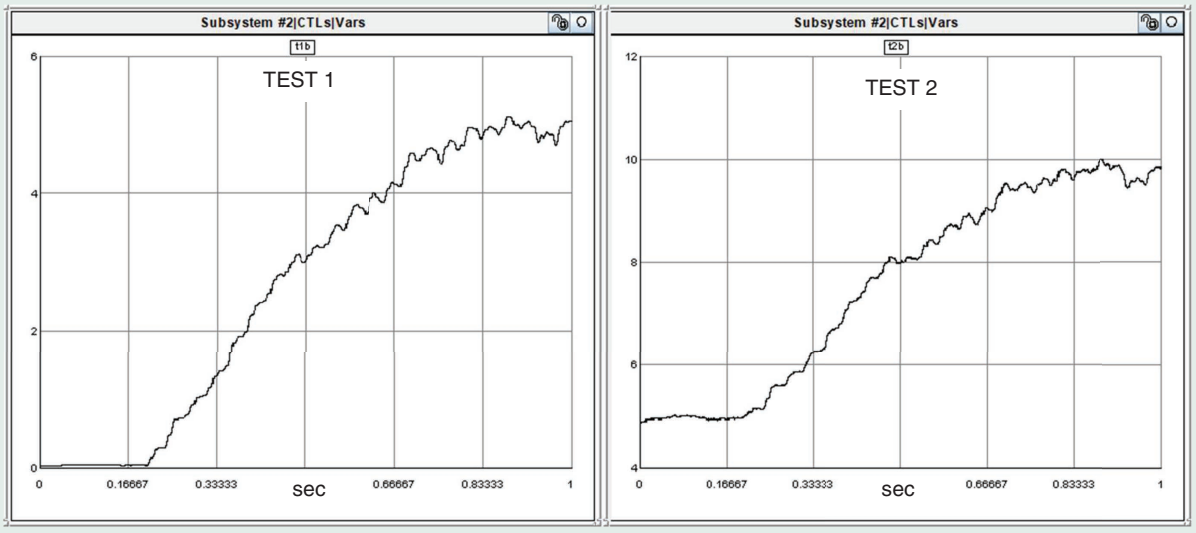

Figure 8. Dynamic watermarking indicators during a cyberattack for (a) test 1 and (b) test 2 results.

stable operation. The destabilizing effect of the data replay attack is quite evident from these results.

To detect the cyberattack, we designed two dynamic watermarking tests (tests 1 and 2), one in the presence and the other in the absence of the watermark and implemented them by using a moving average filter. We injected a dynamic watermark (ω_m) with a variance of 0.005 into the modulation index output of the controller. The inverter current is compared with the expected value generated from a system identification model, and the error between the expected and actual current values is subjected to the dynamic watermarking tests, the outputs of which indicate the presence of a cyberattack when they cross a threshold value.

Figure 8(a) and (b) show the outputs of tests 1 and 2 on the B phase of the solar inverter current before and after the cyberattack, respectively. Initially, the solar inverter kicks off with its normal operation, and the output of tests 1 and 2 is low. However, after the cyberattack is initiated, both tests record a very high value compared to that of the values before the cyberattack. In both cases, we can see that the test 2 results exceed those of test 1 that too exactly by the variance of the watermark signal. This is because test 2 contains watermark information along with the measurement noise as opposed to test 1, which contains only the measurement noise. The results reinforce the utility of dynamic watermarking for detecting cyberattacks on distribution grids. This technique can further be applied on the sensor layer of VPPs to make them cyber resilient.

Summary and Conclusions

VPPs are seen as a powerful platform for building distributed electric power systems that are efficient, secure, sustainable, reliable, and resilient. VPPs can also play a role in mitigating the impact of electric power system disturbances under unexpected but catastrophic events like natural disasters and cyberattacks. Although VPPs pose significant potentials in sustaining local power services under emergency operating conditions, they are also subjected to cyber incidents. Just like many other modern critical infrastructures, VPPs increasingly adopt Internet technology to boost control efficiency, which unfortunately opens up a new front to a potential "cyber Pearl Harbor." It is thus of significance to identify cyber vulnerabilities in VPP operations and deploy effective measures to address cyberthreats. Also, it is critical to pay further attention to strengthening the capabilities of VPPs for sustaining power services in the event of cyber incidents and protect local customers when cyber-induced power outages occur. In this context, cyber-resilient VPP operations will be particularly eminent to maintain highly reliable and resilient power supplies.

The hierarchical management and control structure of a VPP necessitates extensive use of communication. To ensure safe and secure operation, CPSs like VPPs must ensure the confidentiality, integrity, and availability of their information while making sure that cyber contingencies do not ensue a major power blackout. The vulnerabilities of such a CPS need to be studied by using integrated methodologies and tested on platforms that encode the heterogeneity of both the cyber and physical components. The control systems that are deployed in such a CPS must not only regulate the necessary quantities in a VPP but also ensure that they quickly detect and respond to cyberattacks so as not to drive the system into instability.

All the more, it is important to note that apart from finding only technical solutions, it is also necessary to consider human characteristics, behaviors, and choices in building solutions for such a complex cyberphysical applications like a VPP. Although a lot of applications of cyber resilience are available at the transmission level, much work needs to be carried out at the distribution level, especially within the VPP framework. Hence, in this article, all the relevant future research directions have been indicated, following which, VPPs can truly become cyber resilient.

Acknowledgment

This work was supported in part by U.S. Department of Energy Grant DE-CR0000042 (Midwest Center for Microgrid Cybersecurity).

For Further Reading

T. M. Aljohani, "Cyberattacks on energy infrastructures as modern war weapons—Part I: Analysis and motives," *IEEE Technol. Soc. Mag.*, vol. 43, no. 2, pp. 59–69, Jun. 2024, doi: 10.1109/MTS.2024.3395688.

J. Ye et al., "A review of cyber–Physical security for photovoltaic systems," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 10, no. 4, pp. 4879–4901, Aug. 2022, doi: 10.1109/JESTPE.2021.3111728.

W.-H. Ko et al., "Robust dynamic watermarking for cyber-physical security of inverter-based resources in power distribution systems," *IEEE Trans. Ind. Electron.*, vol. 71, no. 7, pp. 7106–7116, Jul. 2024, doi: 10.1109/TIE.2023.3303614.

A. Gholami, F. Aminifar, and M. Shahidehpour, "Front lines against the darkness: Enhancing the resilience of the electricity grid through microgrid facilities," *IEEE Electrific. Mag.*, vol. 4, no. 1, pp. 18–24, Mar. 2016, doi: 10.1109/MELE.2015.2509879.

M. P. Korukonda, R. Prakash, S. Samanta, and L. Behera, "Model free adaptive neural controller for standalone photovoltaic distributed generation systems with disturbances," *IEEE Trans. Sustain. Energy*, vol. 13, no. 2, pp. 653–667, Apr. 2022, doi: 10.1109/TSTE.2021.3123184.

Biographies

Meher Preetam Korukonda (mkorukonda@iit.edu) is with Illinois Institute of Technology, Chicago, IL 60616 USA.

Mohammad Shahidehpour (ms@iit.edu) is with Illinois Institute of Technology, Chicago, IL 60616 USA.

Le Xie (xie@seas.harvard.edu) is with Harvard John A. Paulson School of Engineering and Applied Sciences, Boston, MA 02134 USA.