# Deep Reinforcement Learning for Cybersecurity Assessment of Wind Integrated Power Systems

**XIAORUI LIU**, (Student Member, IEEE), **JUAN OSPINA**, (Member, IEEE),
**AND CHARALAMBOS KONSTANTINOU**, (Senior Member, IEEE)

Center for Advanced Power Systems, FAMU-FSU College of Engineering, Florida State University, Tallahassee, FL 32310, USA

Corresponding author: Charalambos Konstantinou (ckonstantinou@ieee.org)

**ABSTRACT** The integration of renewable energy sources (RES) is rapidly increasing in electric power systems (EPS). While the inclusion of intermittent RES coupled with the wide-scale deployment of communication and sensing devices is important towards a fully smart grid, it has also expanded the cyber-threat landscape, effectively making power systems vulnerable to cyberattacks. This article proposes a cybersecurity assessment approach designed to assess the cyberphysical security of EPS. The work takes into consideration the intermittent generation of RES, vulnerabilities introduced by microprocessor-based electronic information and operational technology (IT/OT) devices, and contingency analysis results. The proposed approach utilizes deep reinforcement learning (DRL) and an adapted Common Vulnerability Scoring System (CVSS) score tailored to assess vulnerabilities in EPS in order to identify the optimal attack transition policy based on $N - 2$ contingency results, i.e., the simultaneous failure of two system elements. The effectiveness of the work is validated via numerical and real-time simulation experiments performed on literature-based power grid test cases. The results demonstrate how the proposed method based on deep $Q$-network (D$Q$N) performs closely to a graph-search approach in terms of the number of transitions needed to find the optimal attack policy, without the need for full observation of the system. In addition, the experiments present the method's scalability by showcasing the number of transitions needed to find the optimal attack transition policy in a large system such as the Polish 2383 bus test system. The results exhibit how the proposed approach requires one order of magnitude fewer transitions when compared to a *random* transition policy.

**INDEX TERMS** Cybersecurity assessment, contingency analysis, cyberattacks, deep reinforcement learning.

## I. INTRODUCTION

The power grid is the cornerstone of all critical infrastructures. The safe and secure functionality of electric power systems (EPS) is directly related to every aspect of the economy and society. In the last decades, worldwide energy demand has significantly increased and is estimated to continue to do so by nearly 50% by 2050 [1]. Due to the increasing energy demand as well as the need to enhance system efficiency and asset reliability, the technological modernization of the power grid infrastructure has become an immediate priority for governments and energy stakeholders around the world [2]. This modernization, alongside environmental concerns, are driving factors for the integration of renewable energy sources (RES) to the power grid. For example,

The associate editor coordinating the review of this manuscript and approving it for publication was Guangya Yang.

the U.S. Energy Information Administration (EIA) indicates that, in 2019, the wind was responsible for generating approximately 42% of RES generated power at utility-scale facilities in the U.S., and 7.3% of the total U.S. electricity generation, making it the most popular RES [3]. Even though wind integration aids in accommodating the increasing power demand, its intermittent nature introduces challenges related to the mismatch between supply and demand. For instance, short-term wind power fluctuations occur on a second or sub-second timescale during which load balancing methods do not yet operate. Thus, to ensure system stability, critical aspects such as optimal location, power flow, and generation variance must be taken into consideration when interconnecting wind energy systems.

Traditionally, contingency analysis has been used to assess physical power system security in EPS [4]. This is achieved by calculating the power flow of all the system elements in the

**TABLE 1.** NERC TPL-001-4 contingency categories.

| Category | Contingency Case |
|----------|------------------|
| $A$ | No contingency |
| $B$ | $N-1$ |
| $C$ | $N-1-1$ or $N-2$ |
| $D$ | $N-k, k > 2$ (cascading) |

event of single or multiple failures. In essence, a contingency is the failure or loss of any element such as a circuit breaker, generator, or transmission line. Contingencies can be planned or unplanned. Planned contingencies include events resulting from scheduled maintenance and proactive emergency preparedness, while unplanned events include fluctuating wind injections, cyberattacks, human errors, etc. The North American Electric Reliability Corporation (NERC) requires system operators to meet the $N-1$ security constraint and classifies systems into four main categories [5]. These categories are shown in Table 1.

The intermittent nature of wind power generates challenges when performing security studies of power systems based on contingency analysis. Their intermittency can rapidly change the most critical contingencies of the system or create a number of contingencies ($\lambda$) that exceeds the maximum number of contingencies that the system can handle ($k$); thus leading to cascading scenarios. A prime example of insufficient security margins is the widespread power outage across the U.K. in 2019 [6]. The near-simultaneous loss of two-generation sites, one being an offshore wind farm and the other one a gas-fired power station, resulted in a massive under-frequency event. Load shedding mechanisms responded immediately causing a major disturbance that affected nearly one million people.

Additionally, the power grid is experiencing a rapid move towards a more interconnected system. Operational technology (OT) electronic devices are deployed and operated at all scales of the power system and are often being designed and retrofitted with information technology (IT) devices to support communication processes and protocols that enhance the controllability and observability of the system. The use of such digital electronic devices with software applications [7], modules, drivers, commercial-off-the-shelf (COTS) hardware [8], and network resources is a double-edged sword [9]. On one hand, it assists in the development of the future modern and advanced grid in terms of optimizing asset utilization, addressing disturbances, providing better power quality, and accommodating all storage and generation options with grid-support functions. On the other hand, the coupling between such cyber-electronic devices and physical components in power systems has altered the threat model [10], [11].

In the past, the threat model has been solely focused on physical threats. However, due to the integration of such network-controlled components, the security challenges need to consider both the cyber and physical nature of the grid, addressing the growing number of emerging threats [12].

Some examples of these potential threats are presented in [13]–[15], where it has been demonstrated that attackers can leverage publicly available sources by using open-source intelligence (OSINT) techniques combined with open-source exploitation methods in order to spoof GPS signals of phasor measurement units (PMUs). Another example is presented in [16], where a real-world attack within the Ukrainian power system is accomplished by injecting malicious firmware in serial-to-Ethernet gateways at targeted substations. Attackers were able to trip circuit breakers and cause a blackout that affected approximately 225,000 customers.

In this work, we provide an effective way for system operators, at both the local and international level, to assess the cyberphysical security of EPS, which takes into consideration wind uncertainty together with cyber-based aspects such as quantitative scoring systems of vulnerabilities identified in IT/OT devices supporting the grid infrastructure. The assessment follows a step-by-step process, from an attacker's point of view, designed to identify the most critical system points an adversary can leverage to compromise the targeted EPS. Our contributions are summarized as follows:

(1) We propose a cybersecurity assessment approach that considers adversaries that make use of OSINT modeling techniques to construct power system models. Such models are then used in tandem with contingency analysis that takes into account wind intermittent generation to identify the critical cyber and physical vulnerabilities of the EPS. The assessment process is performed without the need for full observability of the system since it models the state of the power system as a partially observable Markov decision process (POMDP) that is solved using deep $Q$-networks (DQNs). The solution given by the proposed DQN reveals the optimal attack transition policy an adversary would follow to potentially induce cascading failures in the assessed cyberphysical EPS.

(2) We propose an adapted version of the Common Vulnerability Scoring System (CVSS) based on contingency analysis results and information from the power and communication networks that reveal cyber and physical vulnerabilities within system nodes. The adapted CVSS is used to generate a transition graph designed to assess the complexity of each possible attack path based on various adversarial strategies.

(3) We evaluate the performance of the proposed methodology using real-time simulations on test power systems highlighting the method's scalability. Our results showcase that a fewer number of transitions is needed compared to a *random* transition policy in order to find the optimal attack transition.

The rest of the paper is organized as follows. Section II presents related work. In Section III, we introduce the methodology of the proposed cybersecurity assessment approach. Section IV presents contingency analysis studies for various power system test cases using real-time simulation. In Section V, we demonstrate the effectiveness of the proposed cybersecurity assessment approach and compare it to different transition techniques. Finally, Section VI concludes the paper and provides directions for future work.

## II. RELATED WORK

In this part, we explore some of the state-of-the-art approaches being proposed by researchers that aim to address issues related to (1) $N - k$ contingencies simulations considering intermittent RES generation, (2) assessing the severity of electronic devices security vulnerabilities, and (3) vulnerability and risk assessments methods for cyberphysical EPS.

### 1) METHODS FOR $N - K$ CONTINGENCY ANALYSIS

Towards reducing the occurrence of cascading failures, existing research efforts have focused on proposing efficient methods that can perform studies based on $N - k$ contingency scenarios. Due to the size and complexity of power systems, these ''what-if'' contingency scenarios are based on computationally expensive optimal power flow processes. Research in this area aims to address the computational overhead of $N!/[k!(N - k)!]$ simulations for $N - k$ contingencies. For example, the work in [17] describes a fast-bounding case which requires a small online memory model. Other efforts compute the active power flow change at lines and the voltage change at buses to evaluate the severity of $N - 1$ and $N - 2$ contingencies [18]. In [19], a graph-based power model analysis is presented for contingency ranking. In [20], a heuristics pruning approach for identifying $N - k$ contingencies is discussed while a topology-based algorithm that considers whether the generator or line is in densely populated areas is presented in [21]. The authors use the concepts of closeness and betweenness centrality to determine the component's importance for a $N - k$ criterion.

One of the main challenges of performing contingency studies in power systems with high penetration of wind is that the uncertain nature of wind causes high variability when identifying the most critical $N - k$ contingencies of the system. Existing studies do not often take into account this variability [22]–[24]. The authors in [23] and [25] demonstrate some of the effects that intermittent power generation has in critical contingency identification. They present probabilistic power flow studies that show how the variable nature of power flow, due to wind fluctuations and uncertainties, can alter the number and location of the most critical contingencies recognized by system operators. The correct identification of these critical contingencies is of paramount importance as they can be potentially leveraged by adversaries in order to cause major disruptions in EPS [26], [27].

### 2) SEVERITY ASSESSMENT OF ELECTRONIC DEVICES SECURITY VULNERABILITIES

The wide-scale integration of information and communication technologies in the form of digital electronic devices into the electrical grid expands the list of possible attack vectors that adversaries could exploit to cause major disruptive events. Hence, in order to ensure the secure operation of the entire system, it is essential to consider the inherent vulnerabilities introduced by the grid-supporting IT/OT infrastructure. One scoring system that is widely used for device-level vulnerability assessments in the IT industry is CVSS [28]. The CVSS can assess the severity of software, hardware, and firmware vulnerabilities by using numerical scores. One example of its use can be found in [29]. Here, the authors utilize CVSS to estimate the probability of successfully exploiting identified independent vulnerabilities, including zero-days, existing in components connected to the LAN of a supervisory control and data acquisition (SCADA) system. Another example of CVSS use can be found in [30], where a CVSS-based cyber asset impact score is presented providing a real-time cyber impact severity score that can be used as a basis for processes such as vulnerability management, isolation of cyber assets, and system reconfiguration.

### 3) VULNERABILITY AND RISK ASSESSMENT METHODS FOR CYBERPHYSICAL EPS

Several researchers have focused on developing system-wide security assessment tools aimed to identify possible vulnerabilities and attack vectors which can be subsequently used to produce optimal control policies designed to guide secure operations of cyberphysical EPS. The work presented in [31], proposes power system emergency control mechanisms based on D$Q$Ns to maintain the reliable operation of the system by performing dynamic breaking of generation and under-voltage load shedding. Other researchers have also made use of D$Q$Ns to perform cybersecurity analysis studies in EPS. One example is presented in [32], where the authors propose a D$Q$N-based cyberphysical topology attack designed to trip critically-targeted transmission lines with the objective of exceeding power flow line constraints in the system. This research demonstrates how the disconnection of essential transmission lines could cause system collapse, and how attackers can find out what type of topology attack needs to be performed in order to cause cascading failures. A similar approach is taken by researchers in [33], where a robust D$Q$N-based contingency management mechanism is proposed to provide remedial action when contingencies exist in the system.

Other works have focused on more traditional ranking mechanisms to improve EPS cybersecurity. For example, the research presented in [21] assesses system vulnerability from the cyberphysical security perspective using contingency ranking methods and a cyber-intrusion ranking methodology. Similarly, in [34], an operational reliability impact assessment framework has been developed. In this study, the authors incorporate cyberphysical threats in the assessment of the EPS operation. Another approach is presented in [35], describing an overload risk assessment method based on $N - 1$ contingency analysis and wind penetration.

Compared to the existing work, our proposed cybersecurity assessment approach assesses the cyberphysical security of EPS leveraging the use of deep reinforcement learning (DRL) paradigms while considering the intrinsic interactions between the system's network and physical components. Our work considers physical-based aspects, such as contingency analysis and wind uncertainty, together
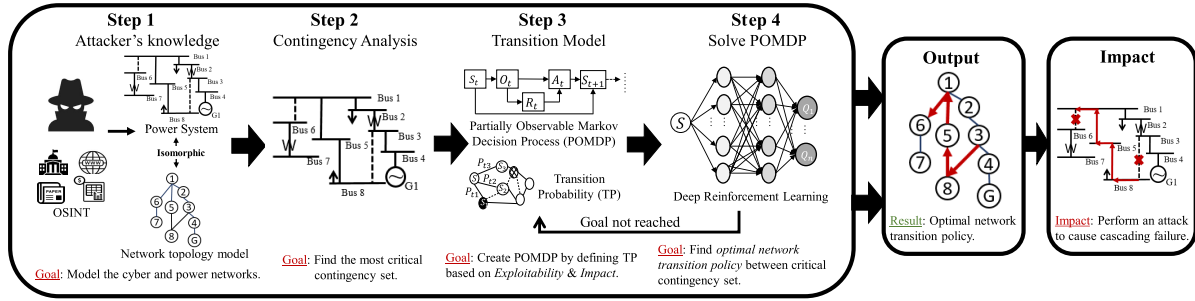
**FIGURE 1.** Graphical depiction of the major steps of the proposed cybersecurity assessment process and the optimal attack transition policy given as output.

with network-based aspects, such as vulnerabilities stemming from OSINT methods. The proposed assessment framework reveals potential threats that can be utilized by attackers to cause serious disruptions in EPS, and at the same time, assist system operators to prioritize the deployment of cybersecurity mechanisms, and thus, contribute towards reducing the risk of cyberattacks.

## III. CYBERSECURITY ASSESSMENT METHODOLOGY

In this section, we provide the methodology of the proposed approach. Fig. 1 shows the step-by-step process that our cybersecurity assessment approach follows. In step 1, the assessment process determines the threat model based on the adversary objectives and capabilities. Specifically, our work considers an attacker that leverages OSINT techniques. The OSINT methods are utilized in step 2 in order to run contingency analysis with the objective of identifying the set of $k$ critical contingencies of the system. Our results focus on two contingencies that assess the power system condition when two components are lost, i.e., $k = 2$. However, the proposed approach can be extended to consider a higher number of contingencies. To proceed with the assessment process without the need for full system observability [36], in step 3, the proposed approach creates a POMDP by defining a transition probability (*TP*) based on the proposed adapted version of the CVSS score metric. The score evaluates the difficulty of each network transition in the generated system graph. Then, in step 4, the POMDP is solved using a DRL model designed to find the optimal attack policy between the previously identified contingencies. Finally, the output of the cybersecurity assessment process evaluates the potential threat by revealing the optimal attack transition policy between the identified contingency pair which could cause cascading failures in the physical system. The details of each step are presented in the following subsections.

### A. STEP 1: THREAT MODEL

OSINT refers to a collection of techniques and methodologies that can be used to gather, analyze, and exploit publicly available information (e.g., public government data, commercial data, social media, etc.) to characterize aspects of a particular target system. Existing work has demonstrated

that the U.S. power grid infrastructure could be effectively profiled using OSINT techniques [37], [38]. Following a similar approach, we consider a threat model where an attacker can leverage publicly available information using OSINT methods to collect sufficient EPS data (e.g., line parameters, the status and location of circuit breakers, system topology, generation sites and capacity, etc.). Also, the attacker is able to acquire data to calculate power flow and therefore run contingency analysis [14]. Depending on the degree of system contingencies (e.g., $N - 1$ secure system), the adversary can leverage the ranking results of contingency algorithms to identify which system elements if "removed" can lead to an insecure power system state. Although a plethora of public power system information is available, it is unlikely that the attacker will ever have full knowledge and real-time observability of the system [39]. In our approach, it is assumed that the attacker, in spite of having the necessary information to perform contingency analysis via OSINT techniques, he/she does not have the full state information of the system. Specifically, while the adversarial agent is transitioning through the cyber system network to exploit vulnerabilities in the identified double contingency nodes, he/she is unaware of his/her position relative to the contingencies and the cyber network transition complexities (based on the adapted CVSS) of the different attack paths.

In addition, we assume that *the cyber system network graph is isomorphic with the physical system graph*, indicating that the topology of the communication network is mapped with the topology of the physical system. Therefore, we model the environment as a POMDP in which the agent may only access the current state and make an observation for obtaining possible actions in each state (Step 3). Based on the observation results for each state-action combination, the network *TP* is calculated. This probability reveals the transition complexity between different states. By leveraging this methodology, a D$Q$N-based algorithm is then utilized to identify the optimal attack transition policy between the critical contingency elements (Step 4).

### B. STEP 2: CONTINGENCY ANALYSIS

In order to find the attacker's optimal attack transition policy, we first need to identify the set of critical double

contingencies of the physical power system (e.g., simultaneous $N-2$ or consecutively $N-1-1$). We utilize a fast pruning $N-2$ algorithm to find all the thermal constraint violations via linear power flow approximation [40]. The algorithm is initiated based on the set of all $N-2$ pairs. The contingency candidate list is pruned using line outage distribution factors (LODFs). LODFs describe the power flow impact on other lines when a line outage occurs. The pruning approach is based on the thermal constraints of lines, running until the number of contingencies included in the set does not change. If the LODF exceeds its thermal constraint, it is added to the contingency candidate set. The line overload condition can be written as $A_{xy} \cdot B_{xc} + A_{yx} \cdot B_{yc} > 1$, where $x$ and $y$ are lines experiencing outages, $z$ is an arbitrary line experiencing power flow changes, and $c$ is a possible constraint. Matrix $A_{xy}$ can be calculated by $A_{xy} = (1 + L_{xy} \cdot f_y/f_x)/(1 - L_{yx} \cdot L_{xy})$, where $L$ is the LODF shown in Eq. (1). This equation describes the change in the flow through line $x$, where $f_x$ is the original flow, and $f_x'$ is the flow after the outage. Correspondingly, $f_y$ represents the flow through line $y$ before the line is tripped. Matrix $B_{xc}$ is calculated by $B_{xc} = f_x \cdot L_{zx}/(f_z^{critical} \pm f_z)$, where $f_z^{critical}$ is the bound value, and the $\pm$ sign corresponds to the conditions $f_z' < -f_z^{critical}$ and $f_z' > f_z^{critical}$, respectively. Eq. (2) shows the power flow variance experienced by line $z$ when lines $x$ and $y$ are experiencing outages.

$$L_{xy} = \frac{f_x' - f_x}{f_y} \tag{1}$$

$$f_z' - f_z = \frac{L_{zx} \cdot (f_x + L_{xy} \cdot f_y)}{1 - L_{yx} \cdot L_{xy}} + \frac{L_{zy} \cdot (f_y + L_{yx} \cdot f_x)}{1 - L_{yx} \cdot L_{xy}} \tag{2}$$

## C. STEP 3: POMDP TRANSITION MODEL BASED ON ADAPTED CVSS

After finding the most critical contingency set, the process advances to create the corresponding POMDP of the cyberphysical-graph environment by calculating the corresponding $TP$ between the different nodes of the system. Generally, POMDPs are used to model the response and outcomes of systems when different actions are performed at specific states. In our environment, observations made by the attacker do not provide full state information, i.e., the agent does not know apriori how many nodes the system has nor their respective states, and he/she needs to observe the environment to determine potential actions, hence the selection of POMDP system modeling. POMDPs can be mathematically modeled as a 6-tuple $(\mathcal{S}, \mathcal{A}, \Omega, P, R, \mathcal{O})$, where $\mathcal{S}$ is the set of all possible states in a given environment, $\mathcal{A}$ contains all the agent's potential actions, $\Omega$ is a set which includes all possible observations, $P$ is the $TP$ for each state, $R$ is the reward function for performing different actions, and $\mathcal{O}$ represents conditional observation probabilities. The notation of POMDP tuples is summarized in Table. 2. At the current state $s$, given the $TP$ and observation $o$, the agent takes action $a$ to move to the next state $s'$. As a result of this state-action pair, the agent receives reward $R$. This process repeats until the terminal state is reached. In this POMDP formulation,

**TABLE 2.** Notation of POMDP tuples.

| POMDP tuples | |
|---|---|
| $\mathcal{S}$ | State set |
| $\mathcal{A}$ | Action set |
| $\Omega$ | Observation set |
| $P\ (TP)$ | Transition probability |
| $R$ | Reward function |
| $O$ | Observation probabilities set |

the $TP$ for each state is an essential factor that must be determined adequately according to the process being modeled. In our case, the $TP$ relies on the cyber system vulnerabilities, i.e., vulnerabilities that exist in electronic devices, and their potential impacts related to the physical system, i.e., the identified power system contingencies.

Considering the cyber network system vulnerabilities as well as the optimal attack transition policy between the identified contingencies (physical vulnerabilities), a $TP$ for each transition step (between cyberphysical system nodes) can be determined. These probabilities aid in the traversal agent's decision making since the values reveal the difference in complexity and difficulty for each transition, i.e., how vulnerable is the cyberphysical system at each node, i.e., bus, from the point of view of the attacker transition policy. In each step, the node's identified cyber and physical characteristics including the electronic device vulnerabilities, thermal limits of lines, and power generation are considered. A graphical illustration of this procedure is shown in Fig. 2. In this work, we compute the $TP$ using an adapted version of CVSS v3.1.
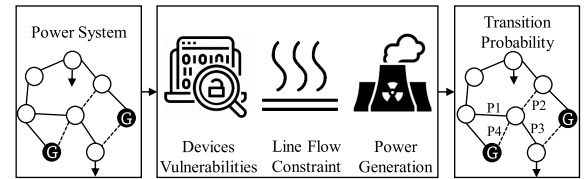


**FIGURE 2.** Overview of the transition probability (*TP*) assessment.

CVSS is a vulnerability scoring system generally used in the IT industry to assess the severity of the identified computer system's vulnerabilities. Although there exist temporal and environmental metrics in CVSS, their main aim is to reflect how vulnerabilities change over time or demonstrate uniqueness to a particular user's environment [28]. For our application, base metrics portray a better picture regarding how the cyber and physical vulnerabilities at each power system node affect the transition difficulty of the threat. More specifically, the base score provides a comprehensive assessment of the intrinsic characteristics of identified vulnerabilities using quantitative *Exploitability* and *Impact* metrics as shown in Fig. 3. The range of scores goes from 0 to 10, with 10 being the most severe – maximum value.
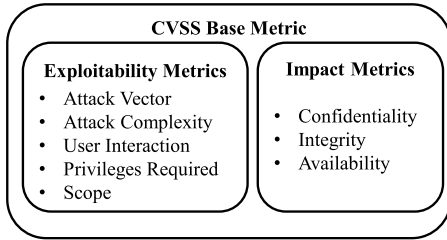
**FIGURE 3.** Outline of the Common Vulnerability Scoring System (CVSS) base metric.

### 1) EXPLOITABILITY METRIC

This metric describes the difficulty and technical means by which software, hardware, or firmware vulnerability can be exploited. In our case, the exploitability represents the difficulty of vulnerability exploitation for each electronic device that exists in a particular node of the cyber-layer of the power system. In other words, it represents the complexity of the transition based on the type of node (i.e., $PQ$ or $PV$ power system bus) to which the agent is transitioning to. The overall score of this metric is determined by five sub-metrics, described below.

*a) Attack Vector (AV)* – This metric is defined as one of the following categories: *network*, *adjacent network*, *local network*, or *physical*. In a network attack, an adversary exploits a vulnerable device bound to the network stack. This type of attack is conducted through the Open Systems Interconnection (OSI) layer 3. In an EPS, an attacker may conduct a network attack by manipulating TCP-level packets flowing across a substation network. In an adjacent attack, the adversary also exploits vulnerable devices bound to the network stack. However, the attack cannot be performed across the boundary of OSI layer 3. In essence, the attack is limited to the same shared physical or logical network. An example of this type of attack is an Address Resolution Protocol (ARP) flooding attack that leads to a denial-of-service targeted at the control and monitoring devices connected to a LAN segment of a microgrid [41]. In a local attack, a direct path to the vulnerable element is required (e.g., local terminal, remote terminal, or deceive legitimate users into executing malicious instructions). In an EPS, this type of attack could be performed by executing malicious code in local control or monitoring electronic devices accessed via a local or remote terminal. Finally, for a physical attack, actual physical interaction between the attacker and the target is necessary. In an EPS, this means that the attacker must compromise the targeted electronic devices through physical means (e.g., causing physical damage to the devices).

*b) Attack Complexity (AC)* – This metric represents the amount of effort an attack on the vulnerable electronic device would require. The value of this metric, *high* or *low*, depends on the security level of the electronic devices as well as the adversary's capabilities and skills. In EPS, generation buses can be considered of more significance than load buses in regards to power grid operation and, consequently,

possible threats. Typically, additional security mechanisms are in place to protect bulk generation infrastructure [42]. This is accomplished by using electronic security devices, physical barriers, or security monitoring equipment. Hence, as part of our CVSS vulnerability scoring, $PV$ and $PQ$ buses are considered of *high* and *low* attack complexity, respectively.

*c) User Interaction (UI)* – This metric reveals whether user interaction is required to exploit a certain electronic device. It quantifies the amount of participation required from a human user, different from the attacker, to successfully compromise the targeted device. For example, attackers could attempt to deceive the system operator to give them access to the control room via phishing or malware attacks. Due to the importance of $PV$ buses, we assume that the attacker will require UI to manipulate a $PV$ bus. On the contrary, it is assumed that attackers would not need to obtain special permission from another human user to access $PQ$ buses. The values for this metric are: *required* for $PV$ buses, and *none* for $PQ$ buses.

*d) Privileges Required (PR)* – This metric determines the level of privileges needed to carry out an attack, i.e., it evaluates the level of privileges that are required by the attacker before successfully compromising the vulnerable electronic device. Similarly to the previous metric, we designate its values according to the type of power system bus being evaluated: *high* for $PV$ buses, and *low* for $PQ$ buses.

*e) Scope (S)* – This metric demonstrates whether or not compromising a particular electronic device will cause implications beyond its security scope. If the scope metric is defined as *changed*, attacking the corresponding electronic device will result in a detrimental implication beyond its security scope, i.e., will affect the other elements in the system. If the scope is defined as *unchanged*, it will only cause implications to elements under the same security scope. In our context, when a $PQ$ bus is attacked, no major disturbances are observed in other system's elements since generation is not directly affected, thus its scope can be defined as *unchanged*. However, if a $PV$ bus is compromised, more severe effects on surrounding nodes of the physical EPS network, caused by power stability issues, are observed. In this case, the scenario needs to be characterized by *changed* scope.

Following the description of the exploitability metrics, Table 3 shows a detailed comparison between the metrics values found in different available scoring systems. These scoring systems are the CVSS v3.1 [28], CVSS v2.0 [43], and the Industrial Vulnerability Scoring System (IVSS) [44]. CVSS v3.1 is the most up to date scoring system which provides the most accurate way of capturing the main characteristics of vulnerability via numerical scores. IVSS is an outdated scoring system and not widely used and supported by the community. Other quantitative risk assessment scoring systems, such as CCSS [45] and CMSS [46], were also considered when selecting the appropriate scoring system. However, all of these scoring systems are based on the previous version of CVSS, i.e., CVSS v2.0.

**TABLE 3.** Exploitability submetrics comparison of different score systems: CVSS v3.1, CVSS v2.0, and IVSS.

| Score System | Metric | Abb. | Metric Value | Num. Value |
|---|---|---|---|---|
| **CVSS v3.1** | Attack Vector | AV | Network | 0.85 |
| | | | Adjacent network | 0.62 |
| | | | Local network | 0.55 |
| | | | Physical | 0.2 |
| | Attack Complexity | AC | Low | 0.77 |
| | | | High | 0.44 |
| | User Interaction | UI | None | 0.85 |
| | | | Required | 0.62 |
| | Privileges Required | PR | None | 0.85 |
| | | | Low | 0.62 if S = Unchanged 0.68 if S = Changed |
| | | | High | 0.27 if S = Unchanged 0.50 if S = Changed |
| **CVSS v2.0** | Access (Attack) Vector | AV | Network | 1.0 |
| | | | Adjacent network | 0.646 |
| | | | Local network | 0.395 |
| | Access (Attack) Complexity | AC | Low | 0.71 |
| | | | Medium | 0.61 |
| | | | High | 0.35 |
| | Authentication (v3.1 - Privileges Required) | AU (PR) | None | 0.704 |
| | | | Single | 0.56 |
| | | | Multiple | 0.45 |
| **IVSS** | Access (Attack) Vector | AV | Remote (v3.1 Network) | 1.0 |
| | | | Local network (v3.1 Adjacent) | 0.7 |
| | | | Local host (v3.1 Local) | 0.4 |
| | | | Physical | 0.2 |
| | Exploit (v3.1 Attack) Complexity | EC (AC) | Low | 1.0 |
| | | | Moderate | 0.5 |
| | | | High | 0.2 |
| | User Interaction | UI | No | 1.0 |
| | | | Yes | 0.3 |
| | Authentication (v3.1 - Privileges Required) | AU (PR) | None | 1.0 |
| | | | User (v3.1 - Low) | 0.6 |
| | | | Admin (v3.1 - High) | 0.2 |

## 2) IMPACT METRIC

In CVSS, the impact metric is used to evaluate different exploitation methods and capture the effects of successfully exploited vulnerabilities. This metric is determined using three factors: *confidentiality* ($C$), i.e., the effect on system information disclosure, *integrity* ($In$), i.e., how detrimental the modification of system data would be, and *availability* ($A$), i.e., the system accessibility after an adverse effect has occurred. During an attack, an adversary can cause *high*, *low*, or *no impact* in each specified factor. For our study, the impact metric is designed to capture the effect of different exploited vulnerabilities in the EPS. During an attack on a *PV* or a *PQ* bus, the system may experience varying degrees of impacts related to *total loss*, *some loss*, or *no loss* of the confidentiality, integrity, and availability of certain grid-supporting devices. More specifically, if the attacker is able to attack a *PV* bus, we assume a worst-case scenario since

the attacker demonstrated to have enough information and skills to attack a highly secure system and possibly have the means to exploit additional vulnerabilities. This, in turn, may result in a total loss of integrity, confidentiality, and availability. Using this assumption, the impact of compromising a *PV* will cause *high* impact on confidentiality, integrity, and availability. On the other hand, despite existing research demonstrating the importance of load altering attacks on power system stability [47], manipulation of *PQ* buses and load change attacks will likely not result in interruption of the operation of the generator, load, or transmission line in the system due to frequency load shedding protections [48]. Under these circumstances, the impact of a compromised *PQ* bus will not be significant enough when compared to the impact of a compromised *PV* bus [49]. Thus, we assume that compromising *PQ* buses will have *low* impact in all three categories. Finally, the *no impact* value is used when an attacker compromises an electronic device that is not connected to any *PV* or *PQ* bus.

Based on the exploitability and impact metrics, CVSS can be calculated as shown in Eq. (3) [28]:

$$CVSS = \begin{cases} \min\{E + I, 10\}, & \text{if } S \text{ unchanged} \\ \min\{1.08 \cdot (E + I), 10\}, & \text{if } S \text{ changed} \end{cases} \quad (3)$$

$$E = (AV \cdot AC \cdot UI \cdot PR) \cdot 8.22 \quad (4)$$

$$I = \begin{cases} 6.42 \cdot I_b, & \text{if } S \text{ unchanged} \\ 7.52 \cdot (I_b - 0.029) - \\ \quad 3.25 \cdot (I_b - 0.02)^{15}, & \text{if } S \text{ changed} \end{cases} \quad (5)$$

where $I_b = 1 - [(1 - C) \cdot (1 - In) \cdot (1 - A)]$. $E, AV, AC, UI, PR, S$, and $I$ represent the exploitability metric, attack vector, attack complexity, user interaction, privileges required, scope, and impact metrics, respectively. The calculated CVSS value is used as a major factor in the computation of the *TP* within our transition model.

The traditional CVSS scoring method provides a detailed calculation process that assesses the impact of exploiting a vulnerability with different attack vectors. However, it cannot be used directly for our application since it fails to consider important factors when used to evaluate complex cyberphysical systems. In particular for power systems, it does not take into account features such as system topology, power generation, and line constraints. Since we assume the adversarial agent does not have full topological information, we include power generation and line constraint calculations in our proposed *TP* calculation. Since generators provide varying amounts of power to a system depending on the current state of the grid, the relative importance of a generator (and hence its attack impact) is determined by its power output. In addition to considering the difficulty of transitioning to certain system nodes, we also examine the overload percentage of the transmission lines. If the power flow across that line is near its flow constraint, the line could be more easily affected by changes in the surrounding system. Taking each of the aforementioned aspects into consideration, we define the *TP*

as follows in Eq. (6):

$$TP = \frac{CVSS}{10} * \frac{G}{\sum_{k=1}^{n} G_k} * \frac{Pf}{\lambda^{critical}} \qquad (6)$$

where $G$ is the power generation of a connected generator, $n$ is the total number of generation units in the system, $Pf$ represents the power flow through transmission lines, and $\lambda^{critical}$ is the line power flow constraint for the connected transmission line. For a power system operating under normal conditions, the range of $G/\{\sum_{k=1}^{n} G_k\} \in [0, 1]$ and $\lambda^{critical} \in [0, 1]$. Since the *CVSS* score $\in [0, 10]$, we scale it by dividing by 10. A smaller *TP* value represents a cyberphysical node vulnerability of low severity, i.e., the node has lower possibilities to be exploited by attackers since it has a lower CVSS score, and it is less important in terms of overload percentage, generation amount, and thermal limits. On the contrary, a *TP* represents a cyberphysical node vulnerability of high severity.

### D. STEP 4: SOLUTION OF ADVERSARIAL MODEL

After formulating and defining the corresponding POMDP, in this step, we develop an algorithm to solve the model and yield the optimal transition policy for the considered threat. Due to the complexity of EPS, it is important to have a mechanism to solve sequential decision-making problems efficiently. In our studies, we develop a DQN-based DRL algorithm.

#### 1) Q-LEARNING

*Q*-learning is an off-policy RL algorithm designed to find the optimal action an agent needs to take at the current state. All the actions that the RL agent can take are evaluated using a *Q*-value which determines how good is a particular action in the current state. Eq. (7) shows the 'update' rule used to calculate new *Q*-values at each state-action pair, where $s$ is current state, $a$ is the action at $s$, $s'$ is the new state, and $a'$ includes all the future potential actions. Using the learning rate $\alpha \in [0, 1]$, the new *Q*-value ($Q_{new}(s, a)$) is updated using the current *Q*-value ($Q_{old}(s, a)$), the estimated optimal future value ($\max_a Q(s', a')$), and the immediate reward ($R$). $\gamma \in [0, 1]$ represents a discount factor that determines the importance of immediate rewards compared with potential long-term rewards. A higher *Q*-value demonstrates that a series of actions will produce a higher total accumulated reward. These actions are referred to as the *optimal policy*.

$$\begin{aligned} Q_{new}(s, a) &= Q_{old}(s, a) \\ &+ \alpha \cdot (R + \gamma \cdot \max_a Q(s', a') - Q_{old}(s, a)) \end{aligned} \qquad (7)$$

Traditionally, *Q*-learning is implemented using *Q*-tables. However, this approach is not practical nor scalable for solving large state-action environments. To solve this issue, researchers in [50] proposed the replacement of *Q*-tables with deep neural networks, also known as DQNs.

#### 2) DEEP Q-NETWORK (DQN)

In order to address the computational overhead of *Q*-learning when dealing with large, uncertain, and dynamic
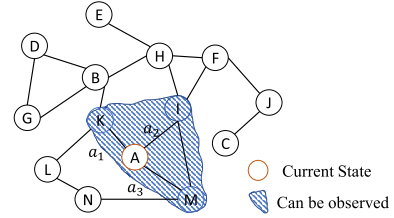


**FIGURE 4.** Example of the observation process for state A.

environments, DQNs generalize the approximation of the *Q*-value function using artificial neural networks rather than storing every solution in a table. For our environment modeled as a POMDP, we assume that the DQN agent starts in a random initial state $s$ (a node in the cyberphysical network) and transitions to the next state $s'$ occur by taking actions (i.e., moving through nodes/buses in the system) based on observations until it reaches both nodes of the contingency pair, regardless of transition order. As shown in Eq. (8), an observation function $O$ generates the observations for each potential action $a'$ given state $s$. The state $s$ refers to the bus where the agent is currently located during the solution process including all bus-related corresponding information and measurements (i.e., circuit breaker status, power generation, power consumption, voltages, etc.). At every step, the agent makes an observation $o$ which helps to indicate which possible movements (or actions) the agent can take according to the observed available neighboring buses and branches connected to the current bus. Then, based on the observation, an action $a$ (movement to another bus) is performed in order to transition to the next state (i.e., the next bus). Fig. 4 presents an overview of this process. It can be seen that if the current state of the agent is bus A, then the attacker (agent) can make an observation $o$ to obtain the next available bus to move into, which can be either K, I, or M, by taking $a_1$, $a_2$, or $a_3$, respectively. Using the observation, the agent is also capable of revealing whether or not a contingency is present in the current bus where it moved into, since the contingency pairs are known to the attacker from the analysis in step 2 (Section III-B).

$$O(o|s, a) = O(o|s', a') \qquad (8)$$

Once each potential transition is determined, the *TP* for each transition needs to be computed (as defined in Eq. (6)). These calculated results will be utilized to determine the security index, $SI'_i$, when making a transition from $s$ to $s'$ as shown in Eq. (9), where $\gamma$ is a discount factor, and $\Delta C_p$ is the line flow difference between the current state and each potential transition state. Finally, the maximum value of the security index which represents the bus with the highest vulnerabilities' score and overload value, will be used to compute the corresponding reward function. As shown in Eq. (10), the reward function considers the overall benefit of different transitions as it takes into account the security index

of each potential state, $SI_i'$.

$$SI_i' = \max_{a \in \mathcal{A}} \gamma \cdot \sum_{s' \in \mathcal{S}} TP\left(s'|s, a\right) \cdot \Delta C_p \tag{9}$$

$$R = \sum_{s' \in \mathcal{S}} TP\left(s'|s, a\right) \cdot [\Delta C_p + SI_i']) \tag{10}$$

State-action-rewards tuples are stored in the replay memory set $M$ for recording agent's experiences. This memory set assists in independently training the neural network. All environmental information of the current state (weights, biases) is stored in the action-value parameter $\theta$. In each step, the DQN combines multi-layered neural networks with existing $Q$-learning algorithms to approximate $Q(s, a; \theta)$. $\theta^-$ will change as the result of changing $\theta$. Eq. (11) demonstrates the updated target value given by the current state and action, where the target action-value parameter $\theta^-$ is equal to $\theta$ at the beginning of the iterations. When this number of iterations is reached during training, $\theta^-$ is updated to prevent an obstructed learning process [51]. Using the parameters described, the loss-function value can be calculated as shown in Eq. (12) for each state-action pair. It represents the error between the predicted $Q$-value and the target $Q$-value. The goal is to determine an *optimal policy* that minimizes the error and ensures that the training result will be as close as possible to the target value, where the target value is the estimated expected return of the actions taken by the DQN.

$$y_i = R + \gamma \max_{a'} Q\left(s', a'; \theta_i^-\right) \tag{11}$$

$$\mathcal{L}_i\left(\theta_i\right) = \mathbb{E}_{(s,a,R,s') \sim \mathcal{M}} \left[\left(y_i - Q\left(s, a; \theta_i\right)\right)^2\right] \tag{12}$$

The agent performs an action that is selected according to the designated exploration-exploitation ($\epsilon$-greedy) strategy of Eq. (13). Such a strategy controls the degree of exploitation over exploration. At each step, if exploration is being performed with probability $\epsilon$, the algorithm selects a random action $a_t$ from the action set. During exploitation with probability $1 - \epsilon$, the action with the maximum $Q$-value is taken. The target values $\theta^-$ will only be updated once the desired number of iterations has been reached [51]. The overall learning process is presented in Fig. 5 and the notation of the DQN parameters is summarized in Table 4. This learning process is repeated until a terminal state is reached, i.e., both contingency pairs have been finally "visited" by the agent.

$$a_t = \begin{cases} random\ a', & \epsilon \\ \arg max_a Q(s', a'; \theta), & 1 - \epsilon \end{cases} \tag{13}$$

### E. STEP 5: OUTPUT OF THE ASSESSMENT PROCESS

An attacker with sufficient OSINT can aggregate enough power system information (e.g., power generation, capacity, load consumption, topological data, etc.) to perform contingency analysis and identify critical system elements. These identified critical contingency elements can be leveraged to generate cyberattack transition policies following the process
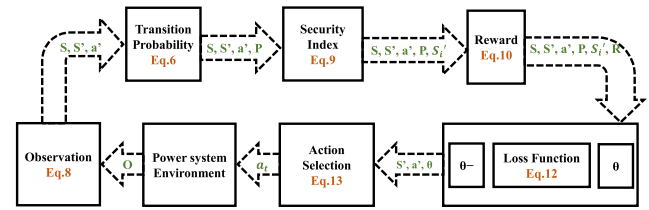


**FIGURE 5.** Process of how the DQN model is utilized as part of the proposed cybersecurity assessment.

**TABLE 4.** Notation of DQN parameters.

| DQN parameters | |
|---|---|
| $\alpha$ | Learning rate |
| $\gamma$ | Discount factor |
| $s$ | Current node |
| $s'$ | Next node transitioned from $s$ |
| $TP$ | Transition probability |
| $\Delta C_p$ | Line flow difference between $s$ and $s'$ |
| $SI'$ | Security index |
| $R$ | Reward |
| $\theta$ | Action-value parameter |
| $\theta^-$ | Target action-value parameter |

described in previous steps. The generated cyberattack transition policies take into account vulnerabilities in electronic devices that exist in the cyber network layer as well as physical system vulnerabilities related to contingency studies. The DRL algorithm provides a solution known as the optimal attack transition policy that can be used to attack the devices controlling the operations of the critical elements (e.g., microprocessor-based relays controlling circuit breakers, protocol translator converters, etc.) and results in potential power outages in the EPS. Our methodology can also be leveraged by control center operators and stakeholders to identify vulnerable components in the EPS or investigate potential attack strategies.

## IV. CONTINGENCY ANALYSIS SIMULATIONS

In this section, we introduce a number of contingency simulation case studies used to demonstrate the effectiveness of the proposed approach. These case scenarios prove how the most critical contingency pairs of a system vary when wind energy systems are in-place. We provide an analysis of the varying degrees of severity with different contingency scenarios and examine how wind generation impacts critical contingencies. For this validation study, we use a doubly-fed induction generator (DFIG) model for wind power generation modeling and real-time simulation (OPAL-RT) for testing the system in a real-time environment.

### A. CONTINGENCY SCENARIOS

First, we run the assessment process of Section III up to *Step 2* in order to assess multiple contingency scenarios in different test systems. In Table 5, we present the number of critical contingencies for $N - 1$, $N - 1 - 1$, and $N - 2$ scenarios

**TABLE 5.** Number of contingencies for different cases and scenarios.

| Case Name | $N-1$ | $N-1-1$ | $N-2$ |
|---|---|---|---|
| IEEE 30 Bus System | 1 | 2 | 8 |
| IEEE 39 Bus System | 13 | 19 | 71 |
| IEEE 39 Bus System + Wind (Table 6: SC7, $t=800m$) | 16 | 24 | 73 |
| IEEE 39 Bus System + Wind (Table 6: SC7, $t=1400m$) | 17 | 26 | 67 |
| IEEE 39 Bus System + Wind (Table 7: SC5, $t=0m$) | 16 | 23 | 103 |
| UIUC 150 Bus System | 176 | 174 | 442 |
| Polish 2383 Bus System | 2236 | 2234 | 15881 |

in different power system test cases. For example, the IEEE 39 bus system has 13 $N-1$, 19 $N-1-1$, and 71 critical $N-2$ contingencies without any wind power injection, while the number of these contingencies varies with different wind penetration levels.

The $N-1$ contingencies are determined by disconnecting each line and observing system responses. For $N-1-1$, the most severe $N-1$ case is removed from the system, and the process is run again. The $N-2$ pruning algorithm is carried out as described in Section III. In the rest of the section, we focus on the $N-2$ case as the most severe scenario. It should be noted that the proposed approach can be adapted, based on user requirements, for any number of contingencies $k$.
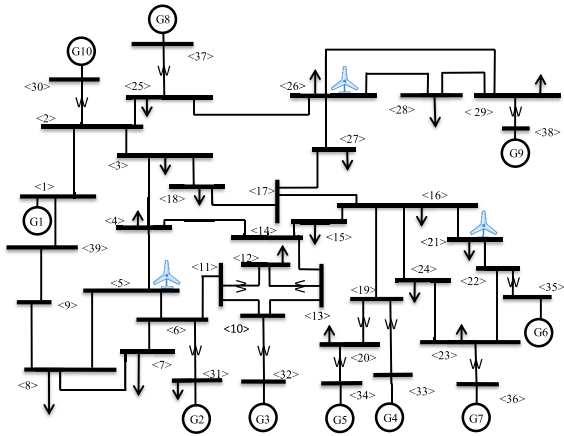


**FIGURE 6.** IEEE 39 bus system with wind power integration.

### B. WIND POWER GENERATION MODELING USING A DFIG MODEL

A DFIG model consists of a wound rotor induction generator driven by wind turbines and an AC/DC/AC insulated-gate bipolar transistor-based pulse width modulated converter. The DFIG model used in our case studies for modeling wind energy systems is developed in MATLAB/Simulink. Using this model, we are able to study the dynamic response of EPS to wind speed variations and investigate the impact of different penetrations. Three DFIGs are modeled and integrated to the IEEE 39 bus system at buses 5, 21, and 26 (Fig. 6) [52].
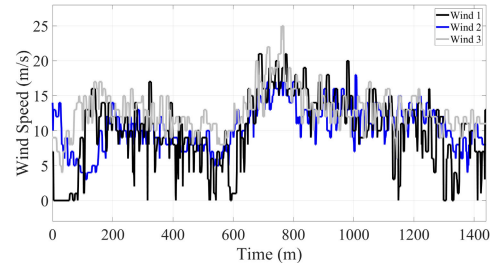


(a)
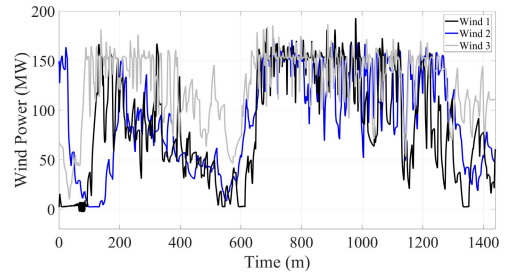


(b)

**FIGURE 7.** (a) Wind speed, and (b) wind power variation for *scenario A*.

The wind speed and wind power data for each wind system are collected at a one-minute resolution on May 14, 2020 (1440 mins = 24 hrs) from [53]. In the rest of the paper, we investigate two scenarios of wind integrated power systems: *scenario A*, in which the wind data is collected from three locations in Tallahassee, FL, with similar variation and power generation levels. The wind speed and corresponding wind power generation information are provided in Fig. 7(a) and Fig. 7(b). In *scenario B*, the wind speed and power data are obtained from three locations in Boston, MA (Wind 1), Dallas, TX (Wind 2), and Tiffin, OH (Wind 3) with different weather characteristics. For this scenario, the wind speed and power generation are shown in Fig. 8(a) and Fig. 8(b), respectively.

### C. CONTINGENCY SCENARIOS WITH WIND POWER INJECTION

The amount of power produced by wind energy systems fluctuates due to wind's intermittent nature. As the generation changes, power flow varies, which may affect contingency analysis results. Therefore, we simulate wind power injection levels at eight distinct timestamps, for the two simulation scenarios (*scenario A* and *scenario B*) throughout one day and observe the changes in reported contingencies with different wind penetration. These tests are performed for the IEEE 39 bus system (Fig. 6). As shown in Table 6 and Table 7, seven wind power integration simulation cases (SC1-SC7) are simulated for *scenario A* and *scenario B*. For each case, we present the amount of power injected by the three DFIG-based wind farms (WF1, WF2, WF3) and the number of identified $N-2$ contingencies.

**TABLE 6.** Scenario A: Impact of different wind power injections on the number of $N-2$ contingencies in the IEEE 39 bus system.

| Time (m) | WF1 (MW) | WF2 (MW) | WF3 (MW) | SC1 WF1 | SC2 WF2 | SC3 WF3 | SC4 WF1+WF2 | SC5 WF2+WF3 | SC6 WF1+WF3 | SC7 WF1+WF2+WF3 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 2.39 | 146.90 | 67.30 | 71 | 78 | 79 | 78 | 85 | 80 | 84 |
| 200 | 108.10 | 131.10 | 125.60 | 52 | 77 | 83 | 64 | 92 | 65 | 80 |
| 400 | 90.92 | 54.10 | 96.80 | 51 | 74 | 81 | 52 | 83 | 62 | 72 |
| 600 | 2.68 | 54.79 | 73.68 | 71 | 74 | 79 | 76 | 82 | 80 | 83 |
| 800 | 153.70 | 159.20 | 155.70 | 53 | 77 | 84 | 71 | 100 | 78 | 73 |
| 1000 | 93.11 | 76.37 | 148.40 | 52 | 76 | 83 | 51 | 91 | 67 | 82 |
| 1200 | 154.30 | 117.50 | 153.80 | 53 | 77 | 83 | 68 | 96 | 78 | 79 |
| 1400 | 77.52 | 54.52 | 105.70 | 81 | 74 | 81 | 50 | 83 | 63 | 67 |

**TABLE 7.** Scenario B: Impact of different wind power injections on the number of $N-2$ contingencies in the IEEE 39 bus system.

| Time (m) | WF1 (MW) | WF2 (MW) | WF3 (MW) | SC1 WF1 | SC2 WF2 | SC3 WF3 | SC4 WF1+WF2 | SC5 WF2+WF3 | SC6 WF1+WF3 | SC7 WF1+WF2+WF3 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 174.30 | 136.80 | 187.30 | 55 | 77 | 91 | 73 | 103 | 85 | 68 |
| 200 | 30.66 | 178.00 | 66.20 | 74 | 80 | 80 | 81 | 85 | 82 | 87 |
| 400 | 15.39 | 107.00 | 29.09 | 69 | 77 | 72 | 78 | 79 | 77 | 80 |
| 600 | 27.94 | 3.56 | 119.20 | 74 | 71 | 82 | 74 | 82 | 86 | 86 |
| 800 | 179.20 | 182.90 | 134.80 | 56 | 80 | 83 | 80 | 97 | 79 | 67 |
| 1000 | 176.80 | 178.20 | 180.20 | 55 | 80 | 89 | 81 | 101 | 86 | 71 |
| 1200 | 75.87 | 161.20 | 178.40 | 81 | 78 | 89 | 76 | 100 | 70 | 90 |
| 1400 | 126.30 | 163.90 | 69.06 | 52 | 79 | 79 | 70 | 87 | 58 | 81 |



(a)



(b)
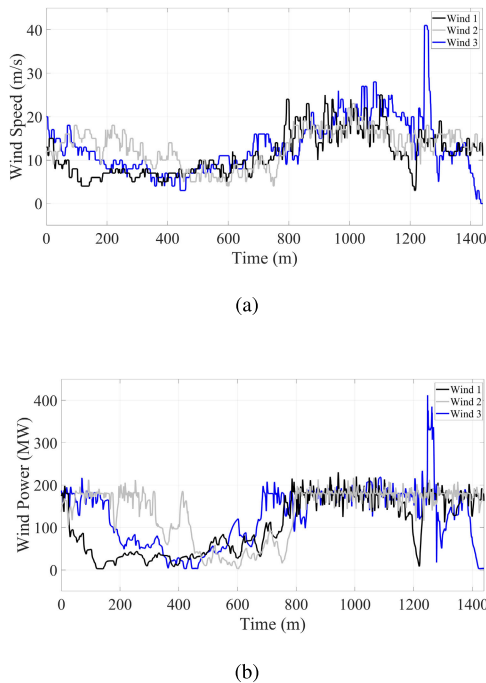
**FIGURE 8.** (a) Wind speed, and (b) wind power variation for *scenario B*.

For *scenario A* in Table 6, the highest number of $N-2$ contingency pairs (100) exists when WF2 and WF3 are integrated to the system (SC5) with generation of 159.20MW and 155.70MW, respectively. The least amount of pairs occurs when WF1 and WF2 turbines are injecting power into the system (SC4), and the wind power injection for WF1 and WF2 are 77.52MW and 54.52MW, respectively. As shown in the results, the number of $N-2$ contingencies change when the same amount of power is injected at different locations. Additionally, injecting varying levels of power in the same location also changes the number of contingencies.

As for *scenario B* in Table 7, the highest number of $N-2$ contingency pairs (103) exists when WF2 and WF3 are integrated into the system (SC5), and the wind power injection for WF2 and WF3 are 136.80MW and 187.30MW, respectively. The least amount of pairs occurs when only WF1 is injecting 126.30MW power into the system (SC1).

Comparing with the normal case of IEEE 39 bus system without wind power injections (71 pairs of $N-2$ contingencies in Table 5), the number of $N-2$ contingencies in 38 cases (out of 56 cases in total in Table 6) of *scenario A* are over 71. For *scenario B*, 44 cases (out of 56 cases in total in Table 7) are more than 71. These results demonstrate how the intermittent behavior of wind energy directly affects the number and location of contingencies in EPS with high penetration of RES. A more specific case that shows how the intermittent behavior of wind can alter the number of contingencies can be observed in Table 5. The number of $N-2$ contingencies can increase or decrease when compared with the case of no wind injection. One scenario that results in a lower number of $N-2$ contingencies is SC7 at $t=1400$ where the number of $N-2$ contingencies decreases from the original 71 to 67; thus making the EPS more secure under contingency conditions. A counterexample of this behavior can be observed in SC7 at $t=0$ where the number of $N-2$ contingencies increases from 71 to 84.

### D. REAL-TIME SIMULATION OF IEEE 39 BUS SYSTEM

We further examine the effect of contingency scenarios in a real-time simulation environment. We observe the impact of intermittent wind power injections across the IEEE 39 bus
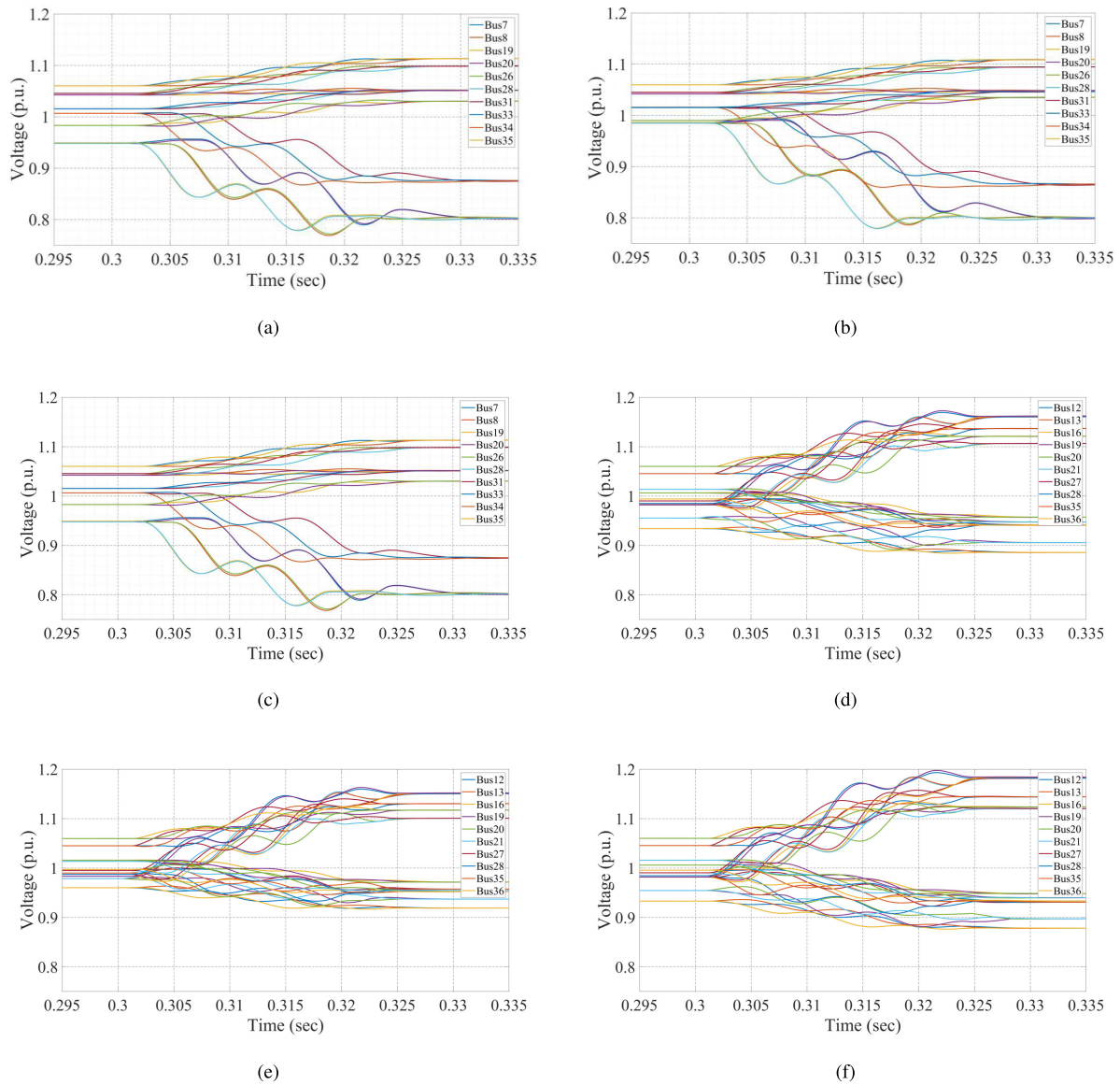
**FIGURE 9.** Contingency scenarios for IEEE 39 bus system: (a) contingency pair at lines $5-8, 6-7$ without wind penetration, (b) contingency pair $5-8, 6-7$ with wind penetration levels as case of Table 6: SC7, $t = 800m$, (c) contingency pair $5-8, 6-7$ with wind penetration as case of Table 7: SC5, $t = 0m$. (d) contingency pair $10-13, 16-21$ without wind penetration, (e) contingency pair $10-13, 16-21$ with wind penetration as case of Table 6: SC7, $t = 800m$, (f) contingency pair $10-13, 16-21$ with wind penetration as case of Table 7: SC5, $t = 0m$.

system by analyzing the variability of all the bus voltages in the system. At $t = 0.3s$, a $N-2$ contingency event is triggered by simultaneously disconnecting two three-phase circuit breakers. To understand the severity of losing critical elements, we disconnect the most critical pair (lines $5-8$ and $6-7$) from the $N-2$ contingency set of the IEEE 39 bus system. Fig. 9(a) presents the $N-2$ effect that disconnecting lines $5-8$ and $6-7$ have in the test system without any wind connected. Fig. 9(b) demonstrates the same contingency scenario (disconnection of lines $5-8$ and $6-7$) with wind power being injected to the system (Table 6: SC7 $t = 800m$). In this scenario, WF1, WF2, and WF3 inject 153.70MW, 159.20MW, and 155.70MW power to the test case system,

respectively. An additional test scenario is run using the same contingency pairs (disconnection of lines $5-8$ and $6-7$) with different wind power injections (Table 7: SC5 $t = 0m$). In this case, WF2 and WF3 inject 136.80MW and 187.30MW, respectively, with the results depicted in Fig. 9(c).

In order to understand the effect that different $N-2$ contingency pairs may have in the EPS, we perform studies using different $N-2$ pairs present in the contingency set. For these studies, we disconnect a less critical contingency pair from the $N-2$ contingency set. At $t = 0.3s$, circuit breakers are tripped at lines $10-13$ and $16-21$, in a test case system without any wind power penetration, and the respective voltage variations can be observed in Fig. 9(d). Fig. 9(e) depicts

how the voltage variations change when wind penetration (Table 6: SC7 $t = 800m$) is considered under the same contingency scenario. An additional case (Table 7: SC5 $t = 0m$) with the contingency pair $10 - 13, 16 - 21$ is demonstrated in Fig. 9(f).

When comparing the real-time simulation results, we can observe that the most critical contingency pair (lines $5 - 8$ and $6 - 7$) causes higher voltage variations when compared to a less critical contingency pair (lines $10 - 13$ and $16 - 21$). Several buses in the EPS reach under and over-voltage values of around 0.87 p.u. and 1.1 p.u. for the most critical contingency pair (lines $5 - 8$ and $6 - 7$) and under and over-voltage values of around 0.92 p.u. and 1.18 p.u. for the less critical contingency pair (lines $10 - 13$ and $16 - 21$). Also, besides observing the voltage variations different contingency pairs can produce, we can also observe, in some cases, how the intermittent behavior of wind power helps to mitigate the severity of line overloads. Figs. 9(d) – 9(f) demonstrate this behavior. For instance, in Fig. 9(e), most buses of the power system have voltage measurements that are closer to the nominal 1.0 p.u value. On the other hand, the case in Fig. 9(f) shows the opposite, since some voltage values measured at some buses are farther apart from 1.0 p.u when compared with the case where no wind power injection is included, i.e., Fig. 9(d). Our results demonstrate how important is to coordinate the amount of wind power as it penetrates the system. For example, the authors in [24] proposed a scheme for power systems to maintain $N - 1$ security within different levels of wind power injection. In addition, a dynamic reserve allocation of DFIG wind farms is presented in [54] to sustain system frequency stability. In our case, the results not only demonstrate the variation of $N - 2$ contingency numbers but also how these results can be used to control the penetration level of wind farms to increase the $N - 2$ secure operational range of power systems.

## V. RESULTS: THE EFFECTIVENESS OF THE PROPOSED CYBERSECURITY ASSESSMENT

This section presents our experimental results that demonstrate the effectiveness of the proposed cybersecurity assessment approach. We evaluate the efficacy of the process according to the optimal attack transition policies given as outputs. In this part, we provide the experimental setup for the presented test cases, the DQN agent model implementation details, and its corresponding hyperparameters. Six test case systems are used to demonstrate the number of transitions needed to identify the optimal attack path for the corresponding case. Furthermore, the performance of the DQN model is evaluated according to the obtained rewards and losses, i.e., convergence for each test case. Finally, the effectiveness of the DQN, used to solve the transition model, is verified by comparing it to other transition-path policy-finding methods, and specifically to the: *(i) random policy search*, *(ii) depth-first search (DFS)*, *(iii) Dijkstra's shortest path* algorithm, and an *(iv) IVSS-based DQN* model.

**TABLE 8.** DQN hyperparameters.

| DQN Hyperparameters | |
|---|---|
| Num. of hidden layers | System size |
| Num. of hidden neurons | 1000 |
| Learning rate ($\alpha$) | 0.005 |
| Activation fcn. | ReLu |
| Optimizer | Adam |
| Max. steps | 1000 |
| Max. episodes | 100 |
| Replay memory buffer (samples) | 1000 |
| Exploration policy | $\epsilon$-greedy linear decay |
| Exploration rate ($\epsilon$) | 0.01 |

### A. EXPERIMENTAL SETUP AND DQN HYPERPARAMETERS

The RL DQN model is trained and tested on a 64-bit machine with an Intel Core i7-7600U, 2.8GHz, and 16.00GB of memory. The proposed algorithm is implemented in `Julia`, a high-level, high-performance, dynamic programming language. The DQN solver for POMDP is provided in [55]. The source files and models associated with this work can be found at [56]. The DQN hyperparameters are presented in Table 8.

**TABLE 9.** Average number of transitions and timing for cyberphysical attacks.

| Case Name | # Trans | # PV | # PQ | $T_{Tr}$ (sec) | $T_{To}$ (sec) |
|---|---|---|---|---|---|
| IEEE 30 Bus System | 9.8 | 2.6 | 9.8 | 76.2 | 83.6 |
| IEEE 39 Bus System | 8.8 | 0.4 | 4.6 | 89.3 | 96.9 |
| IEEE 39 Bus System + Wind W1 (Table 6: SC7, $t = 800m$) | 11.6 | 0.6 | 5.8 | 90.1 | 97.6 |
| IEEE 39 Bus System + Wind W2 (Table 7: SC5, $t = 0m$) | 14.8 | 2.8 | 7.6 | 87.8 | 95.5 |
| UIUC 150 Bus System | 34.0 | 2.4 | 28.4 | 237.9 | 245.6 |
| Polish 2383 Bus System | 35.0 | 5.4 | 30.8 | 1459.5 | 1506.7 |

### B. CYBERSECURITY ASSESSMENT: ATTACK-PATH TRANSITION RESULTS

In order to demonstrate the efficacy of the proposed cybersecurity assessment process, we use six test case power systems related with the contingency studies in Section IV (Table 5): *(a)* IEEE 30 bus system, *(b)* IEEE 39 bus system, *(c)* IEEE 39 bus system with wind W1 (Table 6 SC7 at $t = 800m$), *(d)* IEEE 39 bus system with wind W2 (Table 7 SC5 at $t = 0m$), *(e)* UIUC 150 bus system, and the *(f)* Polish 2383 bus system. Based on the identified critical $N - 2$ pairs, the malicious agent begins at a random initial state and finds the optimal attack-path transition policy to the existing and most critical $N - 2$ contingencies. A contingency is identified when one of the two buses has been visited by the agent. In Table 9, we show the number of transitions required to reach both critical contingencies as well as the number of *PV* and *PQ* buses visited by the agent. For each comparison, five

random initial states are selected for each test system, and the average results are presented. For example, the IEEE 39 bus system requires an average of 8.8 transitions to correctly identify the most critical contingency pair. During the transitions, an average number of 0.4 generation ($PV$) and 4.6 load ($PQ$) buses need to be visited, i.e., compromised, by the agent. $T_{Tr}$ is the training and evaluation time (in seconds) needed for the D$Q$N to 'learn' the optimal attack path for different cases, and $T_{To}$ is the total time (in seconds) required to complete the process. The utilization of the Polish 2383 bus system in our experimental results aids in the evaluation of our proposed process with a realistic large-scale EPS. As seen in Table 9, the training and evaluation process of the D$Q$N in a typical computer with 2.8GHz CPU and 16.00GB RAM requires an average of 1459.5 seconds (approximately 24 minutes), and a total running time of 1506.7 seconds (approximately 25 minutes). Note that the running time of the entire process could be reduced by decreasing the number of hidden neurons in the D$Q$N model. The results demonstrate that the proposed cybersecurity process can be used in tandem with medium and long-term control and planning applications. On the other hand, the proposed approach would require high computing power in order to be integrated into very short-term decision making processes [57].
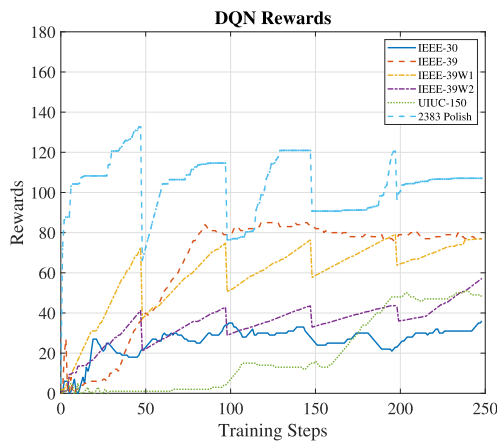
**FIGURE 10.** DQN rewards for bus test systems.

### C. DQN REWARDS AND LOSS CONVERGENCE

As mentioned in Section III, the D$Q$N aims to minimize the loss between the target value and the predicted value. The D$Q$N agent learns the optimal policy as this loss is minimized. Here, we verify and evaluate the performance of our proposed approach by examining the convergence of the D$Q$N loss during the training process. We also show how the average reward gradually increases at each step, for each test case, up to 250 training steps. It should be noted that the total number of training steps used is 500 while the update frequency of the plot is set to 2, thus only 250 steps can be observed in the graph. Fig. 10 shows the rewards for each test case system and Fig. 11 shows the corresponding loss for each case. As shown in Fig. 10, the D$Q$N agent progressively 'learns'
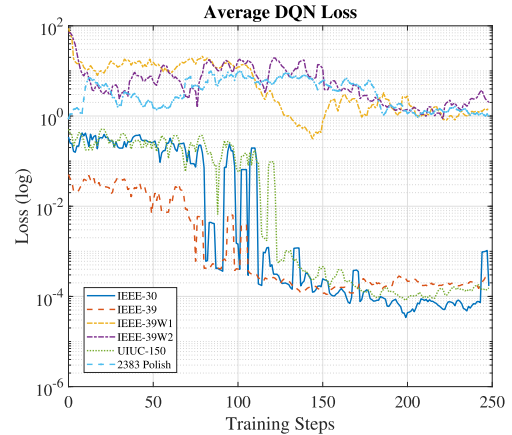
**FIGURE 11.** Average DQN loss for bus test systems.

how to maximize the cumulative rewards in each test case system. At the same time, as the agent 'learns', the loss keeps decreasing until it converges to a minimum value as depicted in Fig. 11. These results showcase the training process of the D$Q$N agent and its performance on all bus test case systems.

### D. EFFECTIVENESS OF DQN: COMPARISON WITH OTHER TRANSITION TECHNIQUES

The effectiveness of using a D$Q$N model in our cybersecurity assessment process is demonstrated by comparing our D$Q$N agent based on the CVSS scoring system with different techniques that could be used to find the optimal attack transition policy in a graph. The techniques used to compare the performance of the proposed D$Q$N are: *(i) random policy search, (ii) DFS, (iii) Dijkstra's shortest path* algorithm, and *(iv) IVSS-based DQN* model. The *random* transition technique provides a baseline, or naive case, where transitions are performed randomly, i.e., without any intelligent control mechanisms. *DFS* is a searching technique for traversing a tree structure by starting from an arbitrary root node and exploring each branch as far as possible before going back to the root node and continuing to the next branch. *Dijkstra's* algorithm is a more sophisticated way of finding an optimal path through a graph structure. *Dijkstra's* algorithm is used to solve shortest-path problems in non-negative weighted graphs by finding an acyclic path between a source and a target node with the minimum transition cost. Both *DFS* and *Dijkstra's* search policies need full observability of the network, hence, for testing purposes in those two cases, we assume full observability of the system and its corresponding contingency pair. Finally, the *IVSS-based* D$Q$N model is designed to evaluate the differences between the CVSS and IVSS vulnerability assessment criteria.

The tests conducted are run using the power system test cases presented in Table 9. For each case, five random initial states are selected and the average number of transitions is calculated. The maximum, minimum, and average number of transitions for each case are shown in the box plots presented in Fig. 12. From Figs. 12(a) – 12(f), we can observe that, in general, the results of the D$Q$Ns-based transition
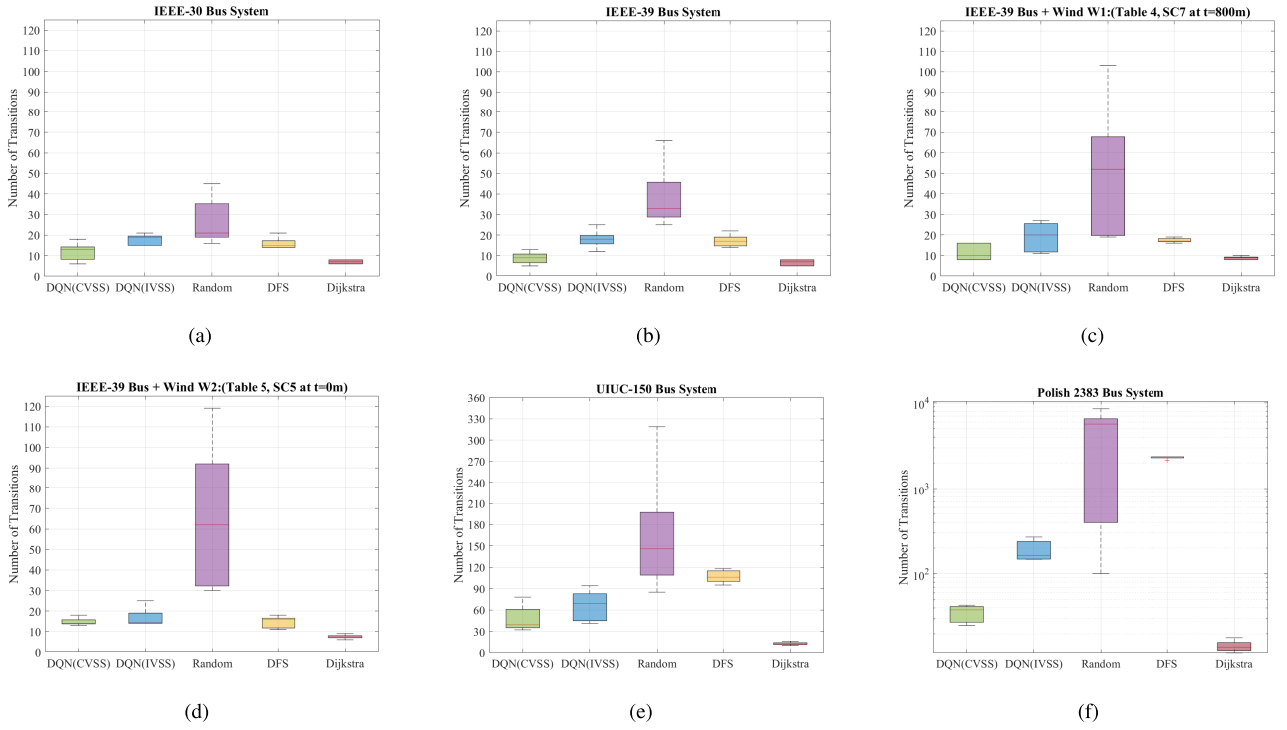
**FIGURE 12.** Number of transitions needed for *DQN (CVSS-based), DQN (IVSS-based), DFS, random,* and *Dijkstra* transition policies for (a) IEEE 30 bus system, (b) IEEE 39 bus system, (c) IEEE 39 bus system + Wind W1 (Table 6: SC7, $t = 800m$), (d) IEEE 39 bus system + Wind W2 (Table 7: SC5, $t = 0m$), (e) UIUC 150 bus system, and the (f) Polish 2383 bus system.
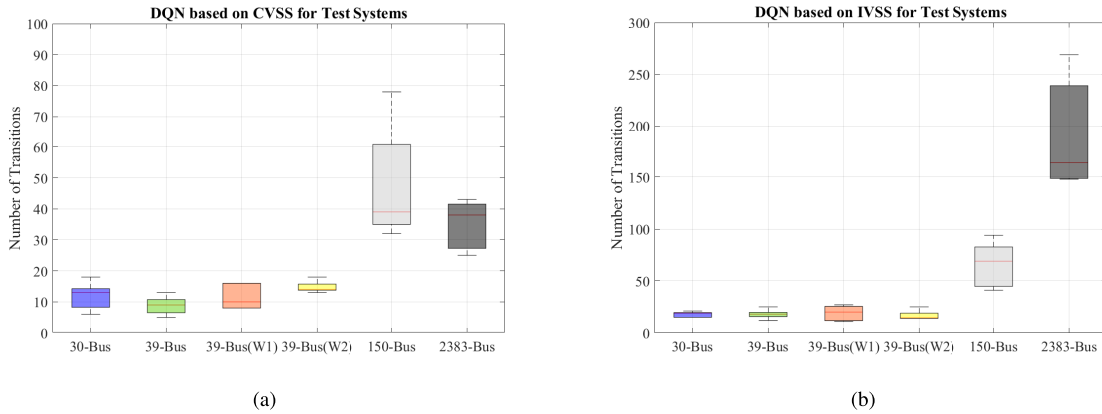


**FIGURE 13.** Number of transitions needed for *DQN (CVSS-based)* and *DQN (IVSS-based)* on all test systems evaluated. These figures demonstrate the scalability of the proposed approach as the number of buses increases.

techniques tend to require fewer number of transitions, i.e., are more efficient, when compared with the *random* and the *DFS* transition techniques. When compared with *Dijkstra's* algorithm, our DQN implementation performs slightly worse due to its iterative learning process. However, *Dijkstra's shortest path* algorithm has the major disadvantage of requiring full system observability. The results demonstrate the advantages of using DQN as the main solver technique for our proposed process. Finally, it can also be observed from Fig. 12 that using CVSS v.3.1 has major advantages when compared to the IVSS scoring system. The CVSS-based DQN consistently requires fewer number of transitions in all evaluated test cases.

To understand the scalability of the proposed DQN approach based on CVSS and IVSS, the number of transitions for the different test systems evaluated are plotted in the box plots shown in Figs. 13(a) and 13(b). In these figures, we can observe that as the number of buses increase, the number of transitions also increases but not in an exponential fashion. Additionally, the results depicted in the figures demonstrate that in almost all test systems the DQN based on CVSS requires a smaller number of transitions than the DQN approach based on IVSS. An example of the improved performance when using the CVSS-based approach can be observed when comparing the number of transitions required for the Polish 2383 test system.

## VI. CONCLUSION AND FUTURE WORK

In this article, we present a cybersecurity assessment approach designed to assess the cyberphysical security of EPS with high penetration of wind. The proposed process assumes that adversaries could leverage OSINT to perform contingency analysis. Based on the contingency results and identified exploitable cyberphysical vulnerabilities via an adapted CVSS metric, an optimal attack transition policy is generated that can be potentially leveraged to cause major outages in an EPS. The results provided by the proposed process are also critical for improving cybersecurity visibility for system operators and stakeholders; it provides information regarding the most critical attack-path an adversary must follow to severely compromise the system alongside with information about the most vulnerable elements in the EPS at a particular time. The proposed approach is tested using real-time simulation, realistic data from various actual wind energy systems, and various test case power systems. Additionally, results regarding the training and convergence of the D$Q$N agent, proposed as the main optimal attack-path transition technique, are presented and compared with other competing techniques. These results demonstrate the applicability of the cybersecurity assessment approach in modern EPS.

The limitations of the cybersecurity assessment approach presented in this work include mostly the assumptions related to the threat model: *(i)* The contingency analysis can only be performed when the attacker has sufficient power system data acquired using OSINT techniques. Without the necessary information, the set of contingencies cannot be correctly identified. *(ii)* The cyber system network graph is assumed to be isomorphic with the physical system graph, indicating that the topology of the communication network is mapped one-to-one with the topology of the physical system. This assumption may not be necessarily true on some systems, since the physical and communication networks could have different network topologies.

Based on the limitations discussed, future work will focus on: *(i)* Exploring potential defense strategies based on moving target defense methods that could be used to enhance the overall system security and resilience by dynamically updating time-varying parameters within the control system of EPS (act as a moving target), thus, limiting adversaries understanding of the cyberphysical EPS model. *(ii)* Analyzing and investigating other DRL solvers that can be adapted into the proposed cybersecurity assessment approach (e.g., UCB, A3C, or TRPO). *(iii)* Investigating how transitions, in the proposed cybersecurity assessment approach, are affected in scenarios where the cyber and physical networks are not assumed isomorphic; examined by using real-time co-simulation testbeds.

## REFERENCES

[1] U. S. Energy Information Administration. (2019). *International Energy Outlook 2019*. [Online]. Available: https://www.eia.gov/outlooks/ieo/pdf/ieo2019.pdf

[2] S. Muyeen and S. Rahman, *Communication, Control and Security Challenges for the Smart Grid*. London, U.K.: Institution Engineering Technology, 2017.

[3] (Jul. 2019). *Generation From Renewable Sources: Total (All Sectors)*. [Online]. Available: https://www.eia.gov/energyexplained/electricity/electricity-in-the-us.php

[4] N. Balu, T. Bertram, A. Bose, V. Brandwajn, G. Cauley, D. Curtice, A. Fouad, L. Fink, M. G. Lauby, B. F. Wollenberg, and J. N. Wrubel, "On-line power system security analysis," *Proc. IEEE*, vol. 80, no. 2, pp. 262–282, Feb. 1992.

[5] *N-1-1 Contingency-Constrained Grid Operations*. Accessed: Nov. 18, 2020. [Online]. Available: https://cms.ferc.gov/sites/default/files/2020-05/20140411131820-T4-A%2520-%2520Fan.pdf

[6] BBC News. (Aug. 9, 2019). *Major Power Failure Affects Homes and Transport*. [Online]. Available: https://www.bbc.com/news/uk-49300025

[7] J. Stright, P. Cheetham, and C. Konstantinou, "Defensive cost-benefit analysis of smart grid digital functionalities," 2020, *arXiv:2008.12843*. [Online]. Available: http://arxiv.org/abs/2008.12843

[8] C. Konstantinou and M. Maniatakos, "Hardware-layer intelligence collection for smart grid embedded systems," *J. Hardw. Syst. Secur.*, vol. 3, no. 2, pp. 132–146, Jun. 2019.

[9] J. Ospina, I. Zografopoulos, X. Liu, and C. Konstantinou, "DEMO: Trustworthy cyberphysical energy systems: Time-delay attacks in a real-time co-simulation environment," in *Proc. Joint Workshop CPS IoT Secur. Privacy*, Nov. 2020, p. 69.

[10] C. Konstantinou and M. Maniatakos, "Security analysis of smart grid," in *Communication, Control and Security Challenges for the Smart Grid*. London, U.K.: The Institution of Engineering and Technology, 2017, pp. 451–487. [Online]. Available: https://digital-library.theiet.org/content/books/10.1049/pbpo095e_ch15

[11] S. McLaughlin, C. Konstantinou, X. Wang, L. Davi, A.-R. Sadeghi, M. Maniatakos, and R. Karri, "The cybersecurity landscape in industrial control systems," *Proc. IEEE*, vol. 104, no. 5, pp. 1039–1057, May 2016.

[12] D. Gritzalis, M. Theocharidou, and G. Stergiopoulos, *Critical Infrastructure Security and Resilience*. Cham, Switzerland: Springer, 2019.

[13] A. Keliris, C. Konstantinou, M. Sazos, and M. Maniatakos, "Low-budget energy sector cyberattacks via open source exploitation," in *Proc. IFIP/IEEE Int. Conf. Very Large Scale Integr. (VLSI-SoC)*, Oct. 2018, pp. 101–106.

[14] A. Keliris, C. Konstantinou, M. Sazos, and M. Maniatakos, "Open source intelligence for energy sector cyberattacks," in *Critical Infrastructure Security and Resilience*. Cham, Switzerland: Springer, 2019, pp. 261–281.

[15] X. Liu, A. Keliris, C. Konstantinou, M. Sazos, and M. Maniatakos, "Assessment of low-budget targeted cyberattacks against power systems," in *Proc. IFIP/IEEE Int. Conf. Very Large Scale Integr.-Syst. Chip*. Cham, Switzerland: Springer, 2018, pp. 232–256.

[16] D. U. Case, "Analysis of the cyber attack on the ukrainian power grid," *Electr. Inf. Sharing Anal. Center (E-ISAC)*, vol. 388, p. 29, Mar. 2016. Accessed: Nov. 18, 2020. [Online]. Available: https://ics.sans.org/media/E-ISAC_SANS_Ukraine_DUC_5.pdf

[17] G. C. Ejebe, H. P. Van Meeteren, and B. F. Wollenberg, "Fast contingency screening and evaluation for voltage security analysis," *IEEE Trans. Power Syst.*, vol. 3, no. 4, pp. 1582–1590, Nov. 1988.

[18] S. Burada, D. Joshi, and K. D. Mistry, "Contingency analysis of power system by using voltage and active power performance index," in *Proc. IEEE 1st Int. Conf. Power Electron., Intell. Control Energy Syst. (ICPEICES)*, Jul. 2016, pp. 1–5.

[19] Y. Zhao, C. Yuan, G. Liu, and I. Grinberg, "Graph-based preconditioning conjugate gradient algorithm for 'N-1' contingency analysis," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2018, pp. 1–5.

[20] S. Hasan, A. Ghafouri, A. Dubey, G. Karsai, and X. Koutsoukos, "Heuristics-based approach for identifying critical n—K contingencies in power systems," in *Proc. Resilience Week (RWS)*, Sep. 2017, pp. 191–197.

[21] C. Vellaithurai, A. Srivastava, S. Zonouz, and R. Berthier, "CPIndex: Cyber-physical vulnerability assessment for power-grid infrastructures," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 566–575, Mar. 2015.

[22] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, "A probabilistic framework for reserve scheduling and $N - 1$ security assessment of systems with high wind power penetration," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3885–3896, Nov. 2013.

[23] K. Sundar, H. Nagarajan, M. Lubin, L. Roald, S. Misra, R. Bent, and D. Bienstock, "Unit commitment with N-1 security and wind uncertainty," in *Proc. Power Syst. Comput. Conf. (PSCC)*, Jun. 2016, pp. 1–7.

[24] L. Bai, F. Li, T. Jiang, and H. Jia, "Robust scheduling for wind integrated energy systems considering gas pipeline and power transmission N–1 contingencies," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1582–1584, Mar. 2017.

[25] J. M. Morales, L. Baringo, A. J. Conejo, and R. Mínguez, "Probabilistic power flow with correlated wind sources," *IET Generat., Transmiss. Distrib.*, vol. 4, pp. 641–651, May 2010.

[26] C. Konstantinou, M. Sazos, and M. Maniatakos, "Attacking the smart grid using public information," in *Proc. 17th Latin-American Test Symp. (LATS)*, Apr. 2016, pp. 105–110.

[27] A. Keliris, C. Konstantinou, and M. Maniatakos, "White paper: GE multilin SR protective relays passcode vulnerability," Black Hat USA, Las Vegas, NV, USA, Tech. Rep., 2017. Accessed: Nov. 18, 2020. [Online]. Available: https://www.blackhat.com/docs/us-17/thursday/us-17-Keliris-And-Then-The-Script-Kiddie-Said-Let-There-Be-No-Light-Are-Cyberattacks-On-The-Power-Grid-Limited-To-Nation-State-Actors-wp.pdf

[28] *Common Vulnerability Scoring System Version 3.1, Specification Document, Revision 1*. Accessed: Jun. 16, 2020. [Online]. Available: https://www.first.org/cvss/v3-1/cvss-v31-specification_r1.pdf

[29] Y. Zhang, L. Wang, Y. Xiang, and C.-W. Ten, "Power system reliability evaluation with SCADA cybersecurity considerations," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1707–1721, Jul. 2015.

[30] V. Venkataramanan, A. Hahn, and A. Srivastava, "CyPhyR: A cyber-physical analysis tool for measuring and enabling resiliency in micro-grids," *IET Cyber-Phys. Syst., Theory Appl.*, vol. 4, no. 4, pp. 313–321, Dec. 2019.

[31] Q. Huang, R. Huang, W. Hao, J. Tan, R. Fan, and Z. Huang, "Adaptive power system emergency control using deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1171–1182, Mar. 2020.

[32] Z. Wang, H. He, Z. Wan, and Y. L. Sun, "Coordinated topology attacks in smart grid using deep reinforcement learning," *IEEE Trans. Ind. Informat.*, early access, May 15, 2020, doi: 10.1109/TII.2020.2994977.

[33] A. Ghasemkhani, A. Darvishi, I. Niazazari, A. Darvishi, H. Livani, and L. Yang, "DeepGrid: Robust deep reinforcement learning-based contingency management," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT)*, Feb. 2020, pp. 1–5.

[34] K. R. Davis, C. M. Davis, S. A. Zonouz, R. B. Bobba, R. Berthier, L. Garcia, and P. W. Sauer, "A cyber-physical modeling and assessment framework for power grid infrastructures," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2464–2475, Sep. 2015.

[35] M. Mehdizadeh, R. Ghazi, and M. Ghayeni, "Power system security assessment with high wind penetration using the farms models based on their correlation," *IET Renew. Power Gener.*, vol. 12, no. 8, pp. 893–900, Jun. 2018.

[36] X. Liu and C. Konstantinou, "Reinforcement learning for cyber-physical security assessment of power systems," in *Proc. IEEE Milan PowerTech*, Jun. 2019, pp. 1–6.

[37] D. Hayes, "Using open source intelligence for risk assessment to the U.S. power grid," in *Proc. E-Soc.*, 2017, pp. 1–10.

[38] *Intelligence Gathering on U.S. Critical Infrastructure*. Accessed: Jul. 15, 2020. [Online]. Available: https://www.icscybersecurityconference.com/intelligence-gathering-on-u-s-critical-infrastructure/

[39] *Bayesian Optimisation for Solving Continuous State-Action-Observation POMDPs*. Accessed: Nov. 18, 2020. [Online]. Available: https://bayesopt.github.io/papers/2016/Morere.pdf

[40] K. S. Turitsyn and P. A. Kaplunovich, "Fast algorithm for N-2 contingency problem," in *Proc. 46th Hawaii Int. Conf. Syst. Sci.*, Jan. 2013, pp. 2161–2166.

[41] H. F. Habib, C. R. Lashway, and O. A. Mohammed, "On the adaptive protection of microgrids: A review on how to mitigate cyber attacks and communication failures," in *Proc. IEEE Ind. Appl. Soc. Annu. Meeting*, Oct. 2017, pp. 1–8.

[42] *Security Guideline for the Electricity Sector*. Accessed: Jun. 20, 2020. [Online]. Available: http://www.iwar.org.uk/cip/resources/nerc/Security%20Guidelines%20for%%20the%20Electricity%20Sector%20-%20Version%201.pdf

[43] P. Mell, K. Scarfone, and S. Romanosky, "A complete guide to the common vulnerability scoring system version 2.0," in *Proc. 1st-Forum Incident Response Secur. Teams*, vol. 1, 2007, p. 23.

[44] C. Bogunden. (2019). *Industrial Vulnerability Scoring System (IVSS)*. [Online]. Available: http://securingics.com/IVSS/IVSS.html

[45] K. Scarfone and P. Mell, "The common configuration scoring system (CCSS): Metrics for software security configuration vulnerabilities," Nat. Inst. Standards Technol. (NIST), Gaithersburg, MD, USA, NIST Interagency Rep. 7502, 2010.

[46] E. Van Ruitenbeek and K. A. Kent, "The common misuse scoring system (CMSS): Metrics for software feature misuse vulnerabilities," US Dept. Commerce, NIST, Gaithersburg, MD, USA, Tech. Rep. 7864, 2009.

[47] S. Amini, F. Pasqualetti, and H. Mohsenian-Rad, "Dynamic load altering attacks against power system stability: Attack models and protection schemes," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2862–2872, Jul. 2018.

[48] B. Huang, A. A. Cardenas, and R. Baldick, "Not everything is dark and gloomy: Power grid protections against IoT demand attacks," in *Proc. 28th USENIX Secur. Symp.*, 2019, pp. 1115–1132.

[49] Y. Cao, Y. Li, X. Liu, and C. Rehtanz, *Cyber-Physical Energy and Power System*. Singapore: Springer, 2020.

[50] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," 2013, *arXiv:1312.5602*. [Online]. Available: http://arxiv.org/abs/1312.5602

[51] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, M. G. Bellemare, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.

[52] H. Bevrani, F. Daneshfar, and R. Daneshmand, "Intelligent power system frequency regulations concerning the integration of wind power units," in *Wind Power Systems*. Berlin, Germany: Springer, 2010, pp. 407–437.

[53] *Weatherstem*. Accessed: Jun. 20, 2020. [Online]. Available: https://lucas.weatherstem.com/data

[54] L.-R. Chang-Chien, C.-M. Hung, and Y.-C. Yin, "Dynamic reserve allocation for system contingency by DFIG wind farms," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 729–736, May 2008.

[55] *DQN Solver for POMDPS.jl*. Accessed: Nov. 18, 2020. [Online]. Available: https://github.com/chsin/DQN.jl

[56] *Github/dss-lab/desaw*. Accessed: Aug. 30, 2020. [Online]. Available: https://github.com/DSS-lab/DRLCyberAssessment_DQNCode

[57] *Optimization and Control of Electric Power Systems*. Accessed: Jul. 10, 2020. [Online]. Available: https://www2.lbl.gov/CS/html/exascale4energy/optimization.html

**XIAORUI LIU** (Student Member, IEEE) received the M.S. degree in electrical engineering from Florida State University, Tallahassee, FL, USA, in 2017, where she is currently pursuing the Ph.D. degree in electrical engineering. Her research interests include real-time simulation of power systems, cybersecurity, and machine learning.

**JUAN OSPINA** (Member, IEEE) received the dual B.Sc. degree in electrical and computer engineering and the M.S. and Ph.D. degrees in electrical engineering from Florida State University, Tallahassee, FL, USA, in 2016, 2018, and 2019, respectively. He is currently a Postdoctoral Research Associate with Florida State University and the Center for Advanced Power Systems, Tallahassee. His research interests include the development of intelligent systems for electric power systems (EPS) and smart-grid applications, machine learning and reinforcement learning models for DER control, renewable energy integration, cybersecurity, and real-time simulation. He is a member of IEEE PES.

**CHARALAMBOS KONSTANTINOU** (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from New York University, NY, in 2018. He is currently an Assistant Professor of electrical and computer engineering with Florida A&M University, and the FAMU-FSU College of Engineering and the Center for Advanced Power Systems, Florida State University, Tallahassee, FL. His research interests include cyberphysical and embedded systems security with a focus on power systems. He was a recipient of the 2020 Myron Zucker Student-Faculty Grant Award from the IEEE Foundation, the Southeastern Center for Electrical Engineering Education (SCEEE) Young Faculty Development Award 2019, and the Best Paper Award at the International Conference on Very Large Scale Integration (VLSI-SoC) 2018.

• • •