Received: 25 August 2020 | Revised: 20 January 2021 | Accepted: 5 March 2021 | IET Cyber-Physical Systems: Theory & Applications

DOI: 10.1049/cps2.12010

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH PAPER

# Cyber-physical component ranking for risk sensitivity analysis using betweenness centrality

**Amarachi Umunnakwe**[1] | **Abhijeet Sahu**[1] | **Mohammad Rasoul Narimani**[1] | **Katherine Davis**[1] | **Saman Zonouz**[2]

[1]Electrical and Computer Engineering, Texas A&M University, College Station, Texas, USA

[2]Electrical and Computer Engineering, Rutgers University, New Brunswick, New Jersey, USA

**Correspondence**

Amarachi Umunnakwe, Electrical and Computer Engineering, Texas A&M University, College Station, Texas, USA.
Email: amarachi@tamu.edu

**Funding information**

U.S. Department of Energy, Grant/Award Number: DE-OE0000895

**Abstract**

This article proposes a model for critical component ranking in power system risk analysis using a proposed cyber-physical betweenness centrality (CPBC) index. Risk assessment, as part of the contingency analysis, is a critical activity that can identify and evaluate component outages that lead to system vulnerability, aiding operators to improve resilience. A power system cyber-physical risk assessment model is proposed that calculates and offers an efficient protection strategy to the system operator based on component vulnerability to adversaries and the impact of compromised assets on the system operation. We present the CPBC index, which traverses generated attack graphs to rank components according to their importance in reducing adversary impact on the power system. The CPBC extends upon betweenness centrality and integrates into analysis, the services and security cost of communications between system components, as well as the likelihood of component exploitation as an adversary medium to the target relays. The proposed model recommends actions, taking into account the interconnections between cyber and physical components as well as cyber-induced Common Vulnerabilities and Exposure scores associated with these connections, thus protecting critical components. The proposed model is implemented on the Cyber-Physical Situational Awareness 8-substation and extended IEEE 300-bus cyber-physical power system models, and results are presented on the impacts of the proposed component ranking model on the security-aware operation of the power system.

## 1 | INTRODUCTION

The electric power grid is critical to national security in modern societies and thus should be resilient to adversaries. Increased cyber attack risks come with modern configurations that integrate advancements in communication and control with advancements in devices, thereby introducing novel opportunities for cyber-threats [1–3]. These cyber-threats can lead to data breaches, asset damage and power outages by exploiting control assets in the physical grid. For instance, the 2015 Ukranian attack exploited a phishing email to ultimately enable the adversary to gain control of system circuit breakers, causing 6 hours of power outages for thousands of customers [4]. Similarly, the Stuxnet attack exploited network printer vulnerabilities to stealthily penetrate the control configuration system and tamper with the control logic of the nuclear process [5]. To be prepared for these anomalies, the system operator usually performs contingency analysis as a risk monitoring tool to provide situational awareness of the power grid [6].

The electric power grid is a complex system comprising interdependent networks of control devices, Internet hosts, sensors, data acquisition, communication services and more. These interdependent networks can be broadly grouped into cyber and physical layers, where anomalies in one layer can have repercussions in the other layers [7]. As interactions between cyber and physical layers increase, the potential paths which a system adversary can exploit to reach critical devices also increases, making comprehensive monitoring more intractable for the system operator; this can have the unintended consequence of the grid becoming a 'honey pot' for

cyber attacks [8]. Hence in the case of an adversary attack, cyber-physical risk analysis as a fundamental power system monitoring tool would allow the operator to gain knowledge of the expected system performance and thus, can aid in preparing the system operators for possible scenarios by ranking equipment according to access, vulnerability, and impact.

This article investigates the effect of component vulnerability on the impact (exploitable paths) of the adversary on the power system. More specifically, we develop here a cyber-physical operation model, CRSA, that ranks cyber and physical components in the order of importance towards minimising the impact of the adversary on the power system. In the proposed cyber-physical model, from a physical perspective, we introduce the cyber-physical betweenness centrality (CPBC) index that effectively ranks power system components based on cyber network configuration, whereas from a cyber perspective, control network vulnerabilities are also integrated according to the underlying power system topology. The main contributions of this article are as follows.

- We propose a *component ranking and risk sensitivity analysis model (CRSA),* which integrates the cyber-physical network topology and standard industry level vulnerabilities to model attack and defense from adversary and system operator perspectives simultaneously.
- We propose a cyber-physical component-ranking metric, the CPBC, which aids this security-oriented risk awareness by ranking system assets according to their security tiers. We compare our proposed CPBC index with the existing BC index to further illustrate the improvements attainable by the proposed risk sensitivity analysis model.
- We develop an algorithm to protect critical components to demonstrate the efficiency of the proposed model, while the model scalability is also illustrated using the Cyber-physical Situational Awareness (CyPSA) 8 substation and IEEE 300-bus test systems, respectively.

The rest of this article proceeds as follows. Section 2 provides a literature review of different methods and models developed for vulnerability analysis of the electric power grid. In Section 3, we discuss the modelling and formulation of system vulnerability including generation of adversary attack graphs. Sections 4, 5 and 6 discuss core facets on the proposed CRSA which are the component ranking assessment, the development of the CPBC index, and the risk analysis process. The simulation results are presented in Section 7, and conclusions are drawn in Section 8.

## 2 | LITERATURE REVIEW

Different methods have been proposed to analyse power system risk to adversarial attacks. Researchers in [9, 10] initially presented the concept of cyber-physical contingency analysis to identify high-risk elements using techniques based on Markov Decision Processes as well as reachability analysis of the attack paths [11] and quantifying physical impact in power systems. Graph theory-based analysis can be utilised to improve this cyber-physical contingency analysis by analysing the system as a weighted graph, where priority can be assigned to edges/vertices with the most connection paths passing through [12]. A graph can represent topology, where the vertices are assets/components such as Internet hosts and relays, while the edges are the communication links between the vertices. Example usages include betweenness centrality (BC) measures [13], which generally seek the relative importance of a vertex or an edge in a graph.

Using the graph theory approach, [14] estimates the impact of the cyber layer on the physical system through cost-effect analysis. Furthermore, [15] proposes probabilistic capturing of data packets for cyber traffic monitoring in software defined networks using the concept of betweenness centrality. In [16], principal component analysis and dictionary learning graph decomposition methods are proposed based on graph multi-centrality features which can reflect structural perturbations in graph symmetry and edge weight and direction, and hence, they can be utilised to detect attacks on the network. These measures can also be adapted to power system networks, using graph topology to detect anomalies in electric power grids [17]. In [18], centrality and electrical characteristics are utilised to identify critical vertices. In [12], parallel BC is applied to power grid contingency selection, validated using a model of the western US power grid to help operators identify and mitigate potential widespread cascading failures in real time.

With these topological methods, a variety of metrics have also been developed in order to identify the most critical components in an electric grid [19–22]. In [23], effective graph resistance is utilised as a metric to assess the robustness of power grids against cascading failure, identifying the best pair of connectivity vertices towards optimising the metric. In [24], systematic investigation of topological and electrical characteristics is performed for power grid networks based on real and synthetic grid data, while in [25], the authors rank the importance of the grid vertices and lines based on centrality measures and other characteristics.

In most of these studies, physical/electrical characteristics are investigated, while cyber vulnerabilities are not integrated. However, researchers have been extending test systems to include cyber characteristics that emulate real systems, featuring communication networks and cyber-physical interconnections that are salient in control of power systems [9, 26, 27]. For example, some intelligent electronic devices, such as relays could be reached via TCP/UDP ports from secured control room computers utilising firewall rules to limit access of vendors, customers, or corporate offices in and out of the control perimeter, as illustrated in Figure 1. These communication networks could potentially be penetrated through external connections, internal Internet hosts, virus penetration, and more. Specifically, although the OT (operational technology) network is isolated from the IT (information technology) network with the use of firewalls and DMZs, a collection of vulnerable web and remote access services can

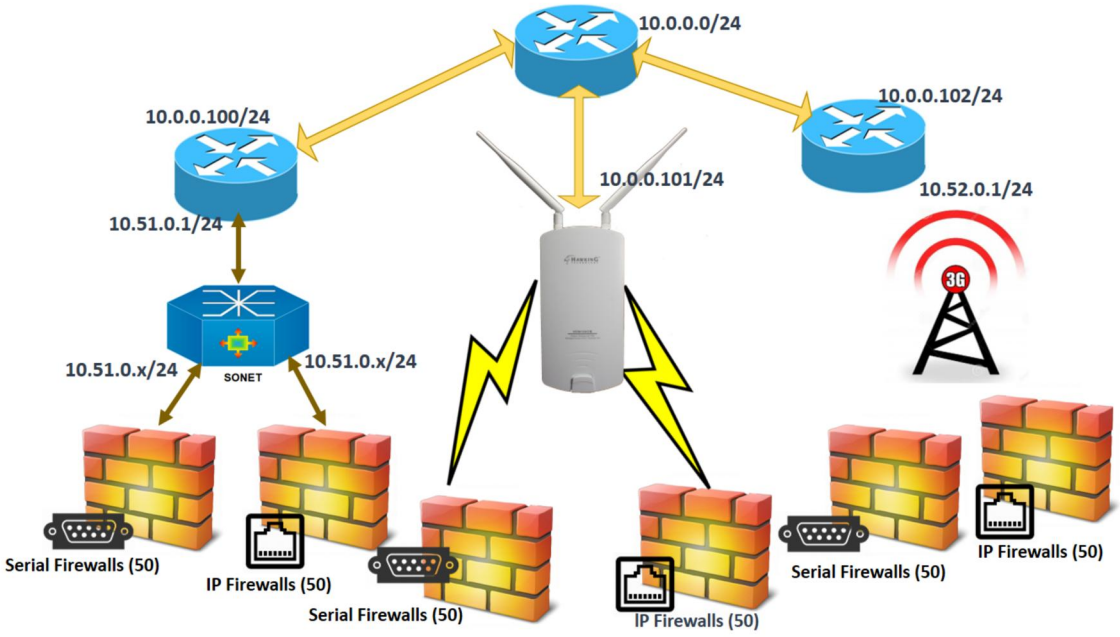


**FIGURE 1** The cyber topology model for 300 substation cases where each substation has dedicated IP or serial interface firewalls and utility control center communication is via long-range communication

still be exploited to plant malware or worms to bridge this isolation [28]. Hence, the main role of the proposed CRSA approach is to assess the component importance towards reducing the security impact (adversary-accessible paths) of adversary intrusion to overall system vulnerability. In particular, we are motivated by the nature of attacks such as the Ukrainian and the Stuxnet, which compromise control assets to create adversarial havoc. Therefore, in this work, we focus on the adversarial process from the operators' host computers (e.g. via phishing emails) to the control network (relays). Thus, as illustrated in Figure 3, the relays form the boundary devices of the cyber-physical network in this article.

Rather than ranking discovered vulnerability by severity [29], the CRSA considers that the operator wants to rank the system components by importance towards reducing the total system vulnerability. Furthermore, CRSA integrates the likelihood and cost of adversary exploitation [30] into cyber-physical risk analysis.

As shown in Figure 2, CRSA utilises the system connectivity, topology information, and user defined adversary and target component lists to generate attack graphs. Given the attack graphs, component ranking follows with the CPBC index consisting of detailed vulnerability scores (cost) of network communication links and the BC of components (vertices), thus demonstrating the relative ease of compromising a communication link and the ease of reaching target assets from unique vertices. The proposed model makes use of information flow, such as services and processes among system components, where the information flow and connectivity of the network are traced at a time when the system is in normal operation. Points of adversary intrusion are then modelled as hosts through which target relays may be reached after a series of vulnerability exploitation.

In other words, using the generated attack graph, potential points of intrusion and potential targets, the CRSA model evaluates the importance of components to the system state based on component services, the security cost of communications between system components, and the likelihood of component exploitation as an adversary medium to the target relays. Based on the results of the proposed approach, ranked components are in turn protected to demonstrate the ability of protecting these components to reduce the overall system vulnerability.

## 3 | CYBER VULNERABILITY MODELING

It is vastly improbable that an adversary will have access to all the information required to carry out an attack on the power system, however, as with all high impact low probability events, the event probability is 0 until the event occurs and then the probability is 1. Hence, in our model we expect that the adversary will inevitably gain system access while the system operator takes contingency measures to minimise adversary impact. We assume that the adversary will prioritise easily accessible paths which pose high vulnerability impact on the system (i.e., access to more targets). Therefore, the adversary has access to the power grid topology information [13] and can carry out an attack based on component vulnerability and graph theory [31]. In this section, we explain how the system information is used to determine the state of the cyber-physical

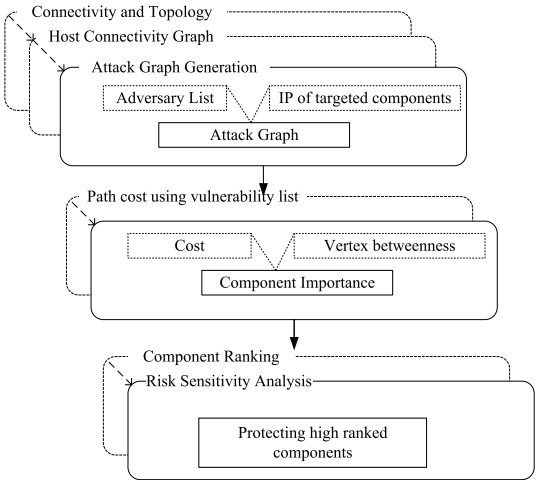


**FIGURE 2** The proposed component ranking and risk sensitivity analysis model

network. We begin by discussing how attack graphs are modelled and generated and continue with discussing the evaluation of inter-component vulnerability.

## 3.1 | Attack graph generation

The goal of the attack graph is to provide details about the cyber-physical power network through dependencies among the system components. The attack graph informs the current state of the system as well as the potential paths an adversary could take to reach target components, given the possible points of intrusion as adapted from Algorithms 2 and 3 of our previous work [32], and is generated from the system connectivity and topology information as follows[1].

### 3.1.1 | The connectivity matrix

The attack graph is generated using the system connectivity matrix (CM) with a pre-defined list of attack vertices and target assets. To aid realistic analysis, this article incorporates the interconnections between the physical (electrical) and cyber networks of synthetic power system models, such as CyPSA 8-substation model [26], which capture normal communications and operational services, for example, remote or secured shell access, between components of the unattacked system and develops a CM. Given the system CM, the security state of the system is evaluated by assigning security scores (cyber costs [CC]) to the communication links between connected components.

### 3.1.2 | Cyber topology and host connectivity generation

Here, we obtain the system topology and host connectivity. Specifically, NMap is used to generate a network mapping

[1]Based on our past work in security assessment [9, 32] and available online [33].

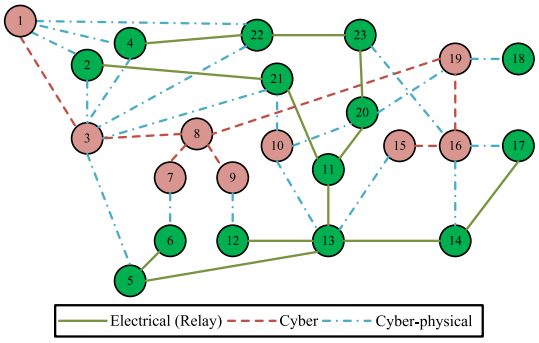


**FIGURE 3** Relay vertices as boundary vertices

report which is spawned from control network hosts and provides host service details. The report is parsed using the NP-View application [34]. Based on the firewall's interface and object group configuration, NP-View generates the cyber topology as a topology dictionary json file having two primary features, namely Device and Network. The feature Device has a list of all the devices such as hosts, relays, gateways with their IP addresses and unique IDs, while the Network feature lists the collection of the model's networks, since there are different networks for the control center, Internet, vendor access and peer utility. These attributes are stored in these dictionaries for accessing unique devices and networks. The connectivity file is generated based on the access control list configured in each firewall [32]. Thus, the NPView parses the Nmap report as well as the firewall rules to generate host connectivity, which along with the physical power grid topology from PowerWorld, is used to model the cyber-physical power system.

## 3.2 | System vulnerability

Given generated attack graphs, the security state of the system can be evaluated. The goal of this section is to explain the system security state. In this work, the adversary gains access to the network and can reach the target relays through Internet hosts. As shown in Figure 2, the attack graph is generated from an input of IPs of the target components, and the adversary list. We assume that an adversary penetrates a utility communication network, and will take a relative path of least resistance to find relays to operate breakers. The adversary penetrates the network which once compromised, ethernet connected relays may be discovered using port scanning tools such as NMap. In our model, discovered relays can be identified using the relay IPs.

Furthermore, connectivity characterisation is stored in three elements: (1) a source object; (2) a sink object; and (3) their security CC. Source and sink are vertices and may have more than one communication link (connectivity edge). For instance, an attack source vertex may leverage knowledge of required username and password to remotely access another sink vertex with hard-coded SSH credentials by exploiting the vulnerability CVE-xxxx-xxxx with a score, hence the path between the two vertices will be weighted on the CCs which are computed based on the Common Vulnerability Scoring System (CVSS) scores obtained from the National Vulnerability Database (NVD). The

vulnerability between vertices can also be depicted by the CIA triad criteria for critical components where vulnerability is modelled, given the information flow between compromised components and those yet to be compromised. Confidentiality and Integrity compromise of an object is captured if a communication link exists with a compromised object which could be in the forward or reverse direction, respectively. Availability is captured, given that a communication link exists and the unavailability of an object is delineated where there is no information flow. For instance, if a critical component frequently communicates another through an *ftp* service, the component once compromised would possibly cause loss of integrity in that ftp service while the confidentiality of the component being written to is also in jeopardy. The component or service availability is threatened if an attack path is *in situ*.

Thus, in this work, vulnerability exploitation through paths is available to adversary to assume relay control. Given these exploitable paths, the component ranking algorithm seeks to identify relatively easy access paths that the adversary can take to get to target assets. Once the paths are ascertained, the vertices most common in these paths have high graph centrality and with consideration of their associated vulnerability types and scores, these vertices are noted as relatively critical for the adversary mission. The critical vertices (important components) are then sent to the system operator to be protected, as a collection of attacks can be prevented by patching system vulnerabilities. For instance, a distributed denial of services (DDoS) can be avoided if vulnerable services or software are patched, uninstalled or filtered. Similarly, a Man-in-The-Middle (MiTM) attack targeting false command or data injection can be avoided, if an intruder is prevented from planting malware or creating botnets. In this work, we assume that once the critical component is protected, the service it provides is deterministically secure and available, that is it becomes 100% secure.

### 3.3 | Dynamic attack graphs for unknown vulnerabilities

Current algorithms to generate attack graphs are based on the CVSS scores of known vulnerabilities. For zero-day attacks or for the source of attacks whose vulnerability are not available in the NVD database, we are exploring dynamic attack graph generation methods such as dynamic Bayesian network (DBN) [35], based on IDS alerts. Hence, probabilistic graphical models such as Bayesian networks (directed) or Markov random fields (undirected) are used to make inference and compute scores or posterior probabilities based on an alert, starting with a prior dummy probability at every node in the attack graph.

## 4 | COMPONENT RANKING ASSESSMENT

In this section, we formulate the component ranking model that integrates cyber topology and vulnerabilities into power system risk sensitivity analysis. Mathematically, the ranking model is formulated given the cyber-physical attack graph $G$. Specifically, the cyber-physical network is a set of components that connect (communicate/control) to one another and hence can be mathematically represented as a graph [36]. The graph vertices represent the system components such as hosts, routers, and relays. The edges represent links between the vertices, for example, service (ssh, tcp) running between two vertices. In particular, if data flows from object $v_i$ to $v_j$, then object $v_j$ becomes dependent on $v_i$ and the dependency is represented by the network edge $e_{ij} = v_i \rightarrow v_j$. To capture this, we represent G as a pair of vertex and edge sets (V, $E$), with vertices, $V = \{v_1, v_2, v_3, \ldots, v_n\}$, and edges, $E = \{e_1, e_2, e_3, \ldots, e_m\}$ with individual weights $CC(e) \rightarrow \mathbf{R}^+$.

### 4.1 | Cyber-physical interdependencies

The nature of historical attacks that is compromises of operator computers to access system control devices, as cited in the introduction, motivates our focus on adversarial analysis between hosts and physical control devices. Hence, it is important to highlight the cyber-physical interdependencies considered in this article:

- From one cyber vertex to another, for example, host–host, host–router link. This interdependency is the data flow or service between cyber vertices.
- From a cyber vertex to a physical vertex (relays) used to send information/commands (control) to the relay.

Given these interdependency types, as illustrated in Figure 3, the physical components are mostly boundary vertices in the network. Figure 3 presents the sample graph where the green vertices are the electrical relay vertices and the pink vertices are the cyber vertices, for example, host vertices, Internet vertices and routers. Similarly, the green edges represent electrical/physical connections, the red edges, ICT/cyber links, and the blue edges represent the interdependency (communication/control) of cyber to electrical vertices. Hence, the goal is to discover the most critical vertices in the system that, if the adversary compromises, will cause higher overall system vulnerability which is measured by the number of attack paths accessible to the adversary. Towards this end, we obtain possible attack scenarios through attack graphs which are analysed for component importance.

### 4.2 | Vertex betweenness centrality

Towards risk assessment, vertex BC assigns ranking coefficients to vertices in a graph through which important components can be identified as those represented by vertices with high coefficient values [37]. It gives insight into the influence of a vertex over the data flow between other vertices. Given the graph $G(V, E)$, the betweenness of a vertex $v$ is the count of the shortest paths between pairs of other vertices that run through $v$ as below:

$$BC(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}. \quad (1)$$

Equation (1) relies on the use of the shortest path distance between the vertices which is computed using the Dijkstra shortest path algorithm, where $\sigma_{st}$ is the number of shortest paths from a source vertex $s$ to a target vertex $t$, $\sigma_{st}(v)$ is the total number from the mentioned paths that pass through vertex $v$, and $n$ is the number of vertices. Hence, the vertices that occur on many shortest paths have relatively higher betweenness [25]. Different studies on cyber-physical vulnerability analysis using graph-theoretic algorithms including BC have been proposed towards contingency analysis [38]. However, utilising just the BC index for critical asset ranking in cyber-physical systems only takes into consideration the centrality positioning of a component in the network graph towards component importance, and hence less accurate results are often obtained. For instance, the BC index may consider a network switch as a critical asset because of its high centrality in the network, even if there are other more critical components with less centrality but with higher potential to cause increased system failure for example, cascade failures.

To enhance the accuracy in component ranking, we propose the CPBC for ranking system components towards risk assessment in the cyber-physical network. Specifically, the proposed index incorporates the impending likelihood of components being compromised, directly or indirectly, in the attack graph as discussed in Section 2. For instance, a vertex A, for example, an Internet host, is affected directly by an adversary if he/she can successfully access that vertex via for example, malicious emails. Alternatively, vertex B, for example, a router, is indirectly compromised if it gets accessed by the adversary through A. In addition, the CPBC index incorporates security vulnerability scores (CC) calculated as follows, using the lowest cost vulnerability to reach a particular vertex even though the attack graph retains all vulnerability IDs.

$$CC(e) = \sum_{e \in E} \min V_e \quad (2)$$

We obtain the vulnerability scores $V_e$ from the NVD where the cost metric associated with realising an attack edge is obtained from the CVSS with a script that extracts the exploitability sub-score using the access complexity and authentication scores. In this work, the CC represents the severity(operator-side)/vulnerability(adversary-side) of compromising a service between two vertices.

## 5 | CYBER-PHYSICAL BETWEENNESS CENTRALITY INDEX

In this section, we present the proposed cyber-physical security index that the CRSA uses to rank components, given possible attacks against the cyber-physical power system.

The objective of the CPBC index is to rank the cyber-physical power system components in order of importance to the power system operator. This importance stems from the impact the adversary will have on the entire system through the compromise of a component given cyber-originated intrusions that target the introduction of malicious commands to the physical power system control components through several host computers to cause a physical-layer security event.

The CPBC utilises the computed shortest paths containing the vertices as in the BC index, the count, and vulnerability magnitude of communication links, to calculate a unified cyber-physical ranking index for the entire network. In particular, the CPBC index integrates the fact that the important vertices have a greater chance to lie on multiple vulnerability-weighted shortest paths to the target relays, as illustrated in line 10 of Algorithm 1, while the vertices with fewer services and lower CC will have relatively less importance. For instance, as illustrated in Figure 4, the adversary at the red source vertices (with CC = 1) will pass through $v1$ and $v2$ to get to their targets $t1$ and $t2$. As we observe, $v1$ provides about double the number of access paths from which the adversary can take the least cost path to $t1$. In addition, the cost of services associated with $v1$ is higher than $v2$ ($29 > 17.9$), hence it will cost more to the system operator if $v1$ is compromised. Thus, $v1$ will rank higher than $v2$, assuming they have the same centrality in the graph.

In Algorithm 1, the relative importance of a vertex due to its position in the network is obtained by defining the Internet and the relay vertices as inputs. Then, the shortest paths from possible adversary sources (Internet) to targets (relay) are calculated. When these paths are obtained, the number of times a vertex occurs in these paths, $\sigma_{st}(v)$, can be determined. For the BC index, this suffices for calculations as in Equation (1), while the obtained $\sigma_{st}(v)$ is a function of the proposed CPBC index, adequately capturing critical vertices:

$$CPBC(v) = \sum_{s \neq v \neq t \in V} \sigma_{st}(v) \times \varepsilon \times \frac{1}{\sum^{e_v} \left(\frac{1}{CC(e)}\right)}, \quad (3)$$

where $\sigma_{st}(v)$ is the number of shortest paths from source vertex, $s$, to target vertex, $t$, that pass through the vertex, $v$, with edges weighted on the communication link CCs, and $e_v$ is

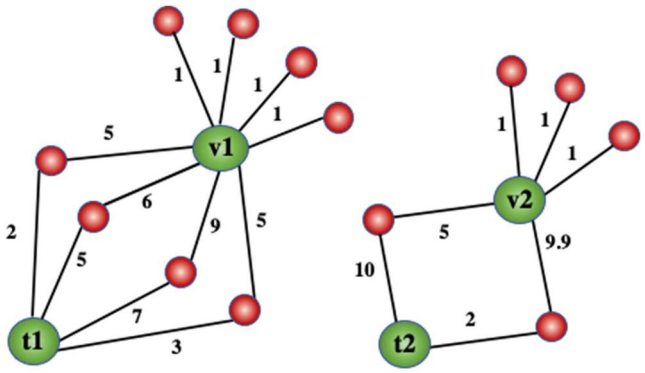


**FIGURE 4** Visual aid for the cyber-physical betweenness centrality

the set of all edges to/from $v$, with cardinality of $\varepsilon$ which is proportional to the vertex density in the network. The reciprocal of the CC is utilised to weigh the vertices in the cyber-physical graph, so that computation is consistent with the cyber vulnerability concept discussed.

**Algorithm 1** Applying Betweenness Centrality

```
Select the IP of targeted relays, Physical_vertices
Select the IP of Internet vertices, Cyber_vertices
function node_importance(vertices)
▷ vertices: Generated Attack graph unique vertices
    for relay in Physical_vertices do
      for host in Cyber_vertices do
        weighted shortest paths      ▷ Get a list of shortest paths, SPL, unless host = relay      ▷ Pass exception if no path
          for short_path S in SPL do
            for node in vertices do
              if vertex is in short_path then
                unique_node_importance += 1
              end if
            end for
          end for
        end for
      end for
    end for
return node_importance, σ_st(v), (for the ranking index)
end function
```

Also worth mentioning is that this index allows for risk analysis where the adversary can compromise a vertex without having access to compromise all the services being provided by that vertex since CC is summed for each compromised $e_v$. For instance, for the ranking of $v2$ in Figure 4, the service represented by the edge with CC of 5 could be compromised with an expected higher probability than that of CC 9.9, the CPBC index is formulated in such a way that this information can be incorporated if so desired. In this case, if granular analysis of the compromised vertex services is required, the CPBC index can be utilised effectively. Another important advantage of this setup is that it allows for the grouping of vertices in security tiers with similar importance, and hence impact, on the overall system vulnerability. This will be further illustrated in the results section.

## 6 | MODEL EVALUATION: RISK SENSITIVITY ANALYSIS

Risk sensitivity analysis proceeds with the prioritised protection of ranked components while the impact of protection towards reducing the system's vulnerability is measured. The objective is to give the system operator enough information about the combination of components that is chosen to protect to have a tractable number of possible adversary accessible paths in case of an attack.

**Algorithm 2** Protecting Important vertices

```
function Generate_Attack_Graph, H(G, L, sel_t)
    create empty attackGraph; H
    Get CC(e) (vuln_list) of x ranked vertices
    for vertex in x do
      v_list = Get(vuln_list - y% of vuln_list)
      new_path = get_path(G, v_list)
      for adversary a in L do
        d, p = djikstra_shortest_path(a, G)
        for target t in d do
          if t in L then
            path = G(t)      (▷) get the path from G
            Add path to attackGraph, H
          end if
        end for
      end for
    end for
return new_attackGraph, H
end function
```

As illustrated in Algorithm 2, the protection of the critical vertices follows with the removal of y% of the unique vertex's associated edges in the attack graph $G$. This generates a new attack graph, $H$, which is a sub graph of $G$, with number of attack paths less than or equal to $G$. In particular, if a vertex is critical, its protection should reduce the number of attack paths $P$ accessible to the adversary. For instance, in Figure 4, assuming the same graph centrality for $v1$ and $v2$, protecting $v1$ with eight immediate communication links will generally reduce attacker access paths in the network than the latter.

The formulation of the protection algorithm is as follows. Let $e_1$ be the set of edges with links to a unique vertex $v1$ in the attack graph $G$, and $e_{v1_c}$ be the set of edges with links to critical vertex $v1_c$ in the attack graph $H$. Then, the list of edges $e_{v1_c}$, associated with critical vertex $v1_c$, is defined as unique row entries with all but y% of the edges of the original set $e_1$, where $e_{v1_c} \in e_1 \in E$. Hence, within a row $e_*$ (e.g., $e_1, e_2,\ldots$) of $E$, the set of edges $e_{y*}$, from vertex $v*$ (e.g., $v1$, $v2,\ldots$), not in $e_{v*_c}$ is defined as:

$$e_{y*} = \frac{y}{100} \quad of \quad e_*. \tag{4}$$

This means that $e_{v1_c}$ is a subset of $e_1$, where $y\%$ of the edges in $e_1$ are removed. Hence for a vertex, $v1$,

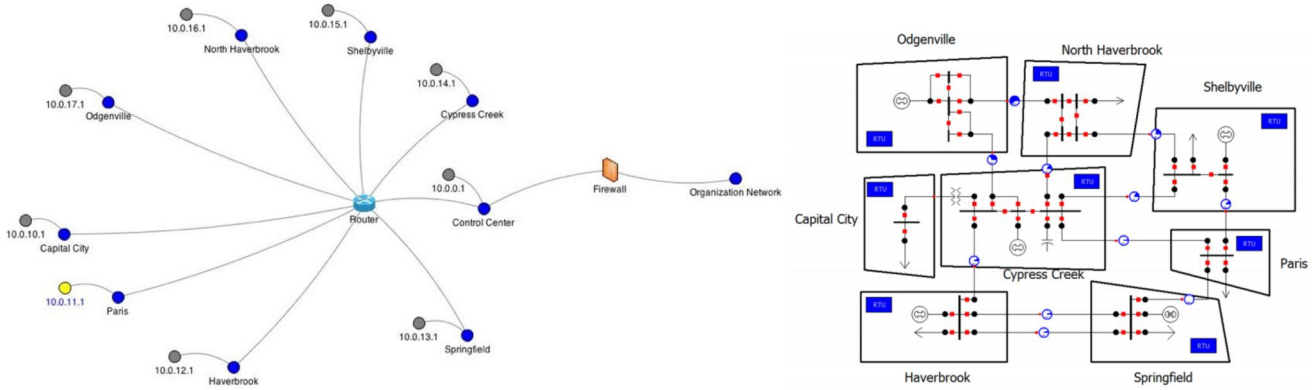


**FIGURE 5** The cyber-physical 8 substation model [26] from NPView and PowerWorld simulator, respectively

**TABLE 1** Component ranking: 8 substation test case

| Rank | BC | Vertex ID | Component type | CPBC | Vertex ID | Component type |
|---|---|---|---|---|---|---|
| 1 | 0.1393 | 1896 | Host PC | 0.0652 | 1896 | Host PC |
| 2 | 0.0837 | 2010 | Distance relay (SEL_421_*) | 0.0583 | [2018,2020,2004,2006] | Host PC |
| 3 | 0.0686 | 1894 | Host PC | 0.0528 | [2014,2016,1998,2000, 2002,2008,1996] | [Overcurrent relay x2, distance relay x4, host PC] |
| 4 | 0.0490 | 1930 | Overcurrent relay | 0.0476 | 2012 | Overcurrent relay |
| 5 | 0.0460 | [1875,1892,2026] | [Router/Switch, Host PC x2] | 0.0304 | 1930 | Overcurrent relay |
| 6 | 0.0301 | 1881 | Router/Switch | 0.0282 | [1920,1922,1924,1926,1928] | [Overcurrent relay x5] |
| 7 | 0.0213 | [1882,1898,1900, 1902,2030] | [Local machine gateway, Host PC x3, Router/Switch] | 0.0175 | 2024 | Distance relay |
| 8 | 0.0210 | 2029 | Router/Switch | 0.0105 | [1938,1940,1942, 1934,1936,1932] | Overcurrent relay x3, distance relay x2, host PC] |
| 9 | 0.0178 | [1920,1922,1924,1926,1928] | [Overcurrent relay x5] | 0.0067 | 2022 | Host PC |
| 10 | 0.0156 | [1916,1918,1910,1912,1914] | [Overcurrent relay x2, distance relay x3] | 0.0061 | [2010,1877] | [Distance relay, router/switch] |
| 11 | 0.0142 | [2012,2014,2016,2018,2020, 1998,2000,2002, 2004,2006,2008] | [Overcurrent relay (SEL_451_*) x5, distance relay x6] | 0.0015 | [1916,1918,1910, 1912,1914,1870] | [Overcurrent relay x2, distance relay x3, router/switch] |
| 12 | 0.0070 | 1996 | Host PC | 0.0013 | 1871 | [Router/switch] |
| 13 | 0.0054 | [1938,1940,1942, 1934,1936] | [Overcurrent relay x3, distance relay x2] | 0.0007 | 1878 | [Router/switch] |
| 14 | 0.0049 | 2024 | Distance relay | 0.0005 | 1894 | Host PC |
| 15 | 0.0015 | 2022 | Host PC | 0.0003 | [1898,1900,1902,2030] | [Host PC x3, router/switch] |

Abbreviations: BC, betweenness centrality; CPBC, cyber-physical betweenness centrality.

$$e_{v1_c} = e_1 - e_{y1}. \tag{5}$$

Analysis for the new generated attack graph advances by calculating the impact of increased protection of important components on overall system attack paths as follows:

$$P_{Total} = \sum P(e_{v*_c}). \tag{6}$$

Hence, Equation (6) measures the improvement, that is, reduction in paths accessible to the adversary, that increased protection of critical vertices provides the system operator. This implies that protection of more critical vertices should relatively provide a higher improvement in the overall system vulnerability with a reduced number of attack paths accessible to the adversary.

Also worth mentioning is that the total system vulnerability cost (CC[P]) could also be used as a measure of overall system vulnerability with some modification to the third term of Equation (3) where instead of considering the CC of edges linked to $v$, the CC of the path from that edge to the target is utilized. For example, for the CPBC of $v2$ in

**TABLE 2** Component ranking: 300 bus test case

| Rank | BC | Vertex ID | Component type | CPBC | Vertex ID | Component type |
|---|---|---|---|---|---|---|
| 1 | 0.2196 | 79,377 | Host PC | 0.9356 | 79,373 | Host PC |
| 2 | 0.2191 | 79,373 | Host PC | 0.0164 | 86,051 | Branch breaker |
| 3 | 0.0439 | [80961,80963,88795,79115] | [Host PC x2, router/switch x2] | 0.0138 | 85,751 | Communications processor |
| 4 | 0.0025 | [87565,87567,87569,87671] | [Bus differential relay, terminal relay x3] | 0.0005 | 79,377 | Host PC |
| 5 | 0.0439 | [81639,82137,82635,83133,...] | [Host PC x4, …] | 0.0001 | [86053,86055,86057, 86059,…] | [Branch breaker, terminal relay x3,…] |

Abbreviations: BC, betweenness centrality; CPBC, cyber-physical betweenness centrality.

**TABLE 3** Actual component rank: 8 substation case. The protected vertex ID is associated with the components as mapped in Table 1

| Protected vertex ID | Final_no_of_attack_paths | Decrease_attack_paths(%) |
|---|---|---|
| 1896 | 68,398 | 12.960 |
| [2018,2020,2004,2006] | 70,469 | 10.324 |
| [2014,2016,1998,2000, 2002,2008,1996] | 70,860 | 9.827 |
| 2012 | 71,256 | 9.323 |
| 2024 | 74,991 | 4.570 |
| [1920,1922,1924,1926,1928] | 75,063 | 4.478 |
| 1930 | 75,097 | 4.435 |
| [1938,1940,1942,1934, 1936,1932] | 75,267 | 4.219 |
| 2022 | 76,080 | 3.184 |
| [2010,1894,1875,1892, 1877,1870,1871,1916, 1910, ...] | 78,582 | 0.000 |

**TABLE 4** Actual component rank: 300 bus case. The protected vertex ID is associated with the components as mapped in Table 2

| Protected vertex ID | Final_no_of_attack_paths | Decrease_attack_paths(%) |
|---|---|---|
| 79,373 | 86,322 | 67.762 |
| 86,051 | 262,242 | 2.062 |
| 85,751 | 262,482 | 1.972 |
| 79,377 | 264,462 | 1.232 |
| [86053,86055,86057,86059, 88343,80961,82365,…] | 267,762 | 0.000 |

Figure 4, instead of utilizing [5; 9.9], the CC of the paths [5, 10; 9.9, 2] is used.

## 7 | SIMULATION AND NUMERICAL RESULTS

The goal of this section is to demonstrate how the CPBC index can aid system operators and administrators in calculating the component ranks using realistic case studies. The proposed CRSA model is implemented on an 8-substation test case [26], as shown in Figure 5, with 78,582 attack paths in $G$, and an extended cyber-physical IEEE 300 bus test case with 267,762 attack paths in $G$ [39]. The 300 bus test case is utilised to illustrate the computational complexity of the proposed model, as the case consists 4500 IP addressable devices with 1301 operational devices, that is, relays and 2384 non-operational devices, for example, fault recorder, alarm systems, batteries. The cases are further described in [26, 32] and are publicly available for download [33]. To illustrate the effectiveness of the proposed model, we consider the improvements offered by using the CPBC index in the risk sensitivity analysis compared with the BC index. The results are computed using a computer with an i7 1.80 GHz processor and 16 GB of RAM.

### 7.1 | Cyber-physical component ranking

We implemented the proposed cyber-physical ranking model on the test cases with results as illustrated in Tables 1 and 2. The tables show the calculated and normalised values for the indices that is, the CPBC and the BC index. The first column in the table shows the rank of the vertices until such a rank where the decrease in overall system vulnerability is negligible for the test system. The second and fifth columns furnish the calculated and normalised values for the BC and the CPBC indices, respectively. The third and sixth columns furnish the unique identification (ID) for the vertices as ranked by the BC and the CPBC. respectively, while, the fourth and seventh columns present the component type. For instance, Host PC with ID 1896, ranked 1 (most critical) by both the BC and the CPBC in Table 1, when protected, drastically reduces the adversary security impact on the system by 12.95% as observed from Table 3.

### 7.2 | Cyber-physical risk sensitivity analysis

Here, we evaluate the proposed model to assess the impact of the ranked components on decreasing the overall system

**TABLE 5** Computational complexity

| Test case | Internet hosts | Relays | Attack paths | CPBC time (s) | BC time (s) |
|---|---|---|---|---|---|
| 8 substation | 11 | 54 | 78,582 | 7.6697 | 7.6 |
| 300 bus | 5 | 1300 | 267,762 | 2024.62 | 1712 |

Abbreviations: BC, betweenness centrality; CPBC, cyber-physical betweenness centrality.

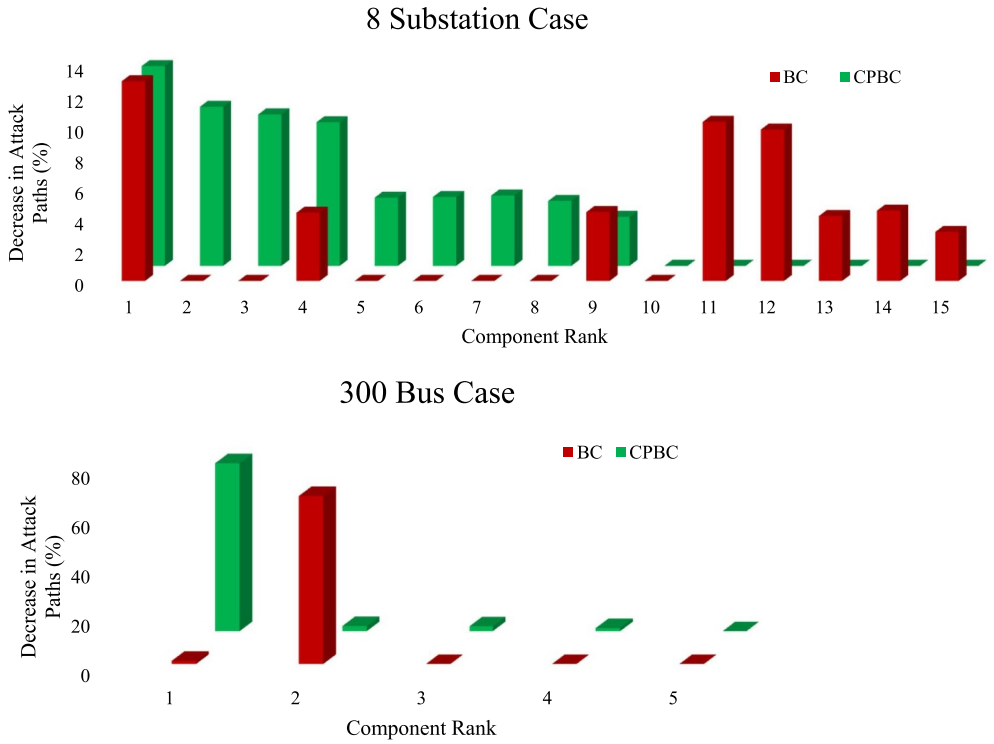


**FIGURE 6** 8 Substation and 300 bus test cases: Visualising the decrease in attack paths illustrated in Tables 1 and 2. BC, betweenness centrality; CPBC, cyber-physical betweenness centrality

vulnerability by reducing the total number of adversary-accessible attack paths. After component ranking, as illustrated in Tables 1 and 2, the vertices are in-turn protected as in Algorithm 2, by reducing the vulnerabilities associated with that vertex by 100%, hence deterministically patching the vulnerabilities. We choose 100% for the purpose of this evaluation to eliminate bias that can occur in the results due to randomly choosing different vulnerability types to be removed. This leads to a new system attack graph with total adversary-accessible attack paths less than or equal to that of the original attack graph. Tables 3 and 4 furnish the decrease in attack paths that the protection of each of the power system components provides. The second column represents the total number of attack paths present in $H$. The third column furnishes the total percentage decrease in the attack paths present in $H$, from the number of attack paths present in the original attack graph $G$, before component protection. In Figure 6, the accuracy of the CRSA model is observed in the decreasing slope of percentage adversary-accessible attack paths as the component ranks progress from 1 to 15 and one to five, for the 8 substation and 300 bus test cases, respectively. This sustained reduction, as opposed to the random decrease in ranking attained by using the BC index, is preferable since [component_importance $\propto$ percentage_decrease_in_attack_paths]. Hence the decrease in attack paths is attained by protecting a component of Rank 1 > Rank 2 > Rank 3 > …as illustrated in Figure 6. Thus, reduction in system vulnerability is expected to be higher with the protection of highly ranked components.

Furthermore, we observe that the proposed CPBC ranking, as shown in Tables 1 and 2, calculates the same rank for the vertices with an equal decrease in the number of attack paths accessible to the adversary. This is due to the comprehensiveness of the CPBC index with the incorporation of criticality (communication link vulnerability costs and cardinality) with vertex betweenness, in the proposed CPBC index. Hence, this additional component grouping functionality, not provided by the traditional BC index, aids in simplifying and reducing the computational burden during cyber-physical risk analysis as illustrated in Tables 3 and 4, where the set of components with equal importance is provided to the system operator.

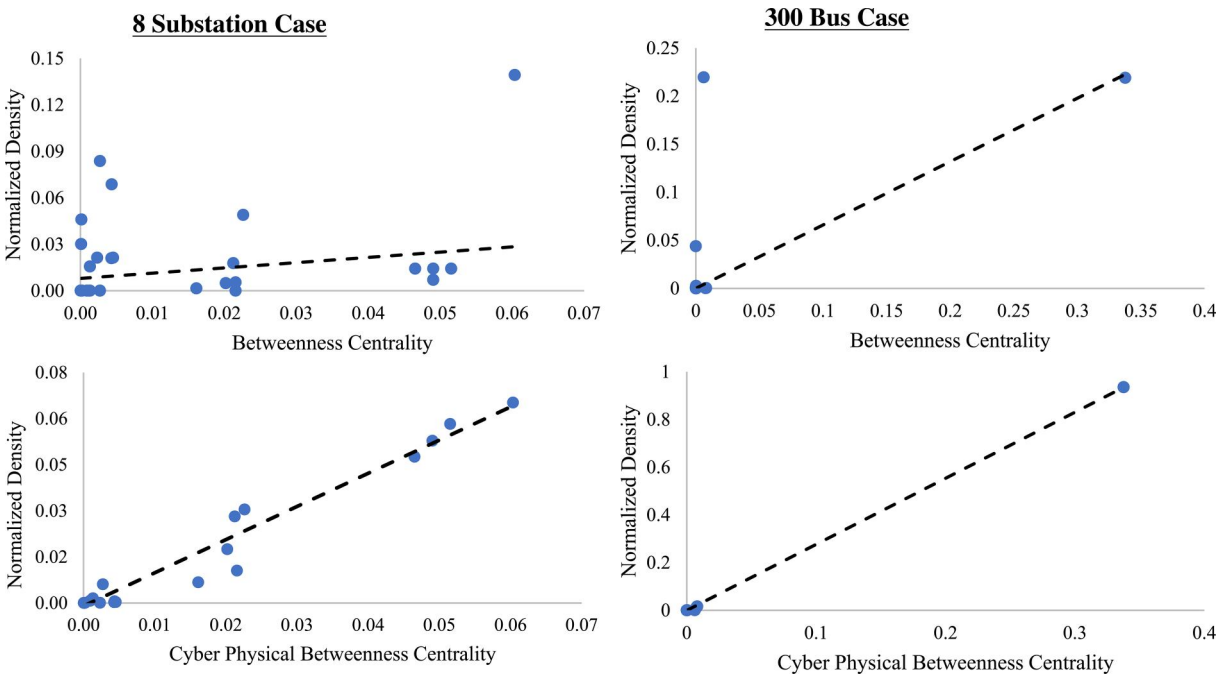


**FIGURE 7** Vertex density analysis: Comparing the correlation of using in the ranking model, the betweenness centrality versus the proposed cyber-physical betweenness centrality index with the node density of the 8 substation and IEEE 300 bus test cases

## 7.3 | Complexity and computational efficiency

From Algorithm 1, we can compute the time complexity of the component ranking algorithm to be of the order of $O(I \times R \times Avg_{PL} \times N)$, where $I$ is the number of Internet vertices, $R$ is the number of relay vertices, $Avg_{PL}$ is the average shortest path length which will depend on the graph density, and $N$ is the total number of vertices. With approximation, we can consider the time complexity of the BC algorithm to be $O(N^4)$. The number of the Internet and relay vertices as shown in Table 5 also influence the computation time, as the CPBC index traverses, the attack graph starting from Internet hosts and terminating in the relay vertices, hence adding to the time complexity of the CPBC ranking. Note that the time for the attack graph generation, an input to the proposed ranking model, increases with larger connected networks (9 min for the IEEE 300 test case) as detailed in our previous work [32], while in this article, we focus on the time complexity of the proposed ranking model.

## 7.4 | Vertex density analysis

Vertex density is the relationship between the number of edges associated with a vertex and the total number of possible edges in the attack graph [40]. Hence, the vertex density holds information on the importance of a vertex [41]. Here, we show the improvements attained by the proposed CPBC index as opposed to the traditional BC index using their correlations with vertex density as shown in Figure 7 where we observe approximately linear relationships, however, with higher correlation between vertex densities and the CBPC index as opposed to the traditional BC index.

# 8 | CONCLUSION

This article proposed a model for critically ranking system components, which integrates cyber-layer industry-standard security vulnerabilities into the risk sensitivity analysis of the cyber-physical power system. The proposed model includes three main stages, where the first stage leverages the subjective adversary vertices and the targeted components to generate an attack graph which estimates the potential adversary attack paths, using the system connectivity and topology. The second stage integrates criticality and target reachability of components, via a proposed CPBC index, to determine the component importance which is passed on to the system operator to analyse the system risk. By prioritising protection of the critical components, the system operator analyses the impact of protection on the overall system vulnerability. The proposed model is implemented on a test 8 substation cyber-physical power system, in addition to the cyber-physical IEEE 300-bus test system, and compared with the BC index to illustrate the advantages of the proposed CPBC index. The simulation results demonstrate that using the proposed index promises improved determination of the important system components. Future works may include expanding the proposed model to a dynamic risk assessment, which would account for changes in the cyber-physical power system with time.

## ACKNOWLEDGEMENT

The work described in this article was supported by funds from the US Department of Energy under award DE-OE0000895 and the National Science Foundation under Grant 1,916,142.

## REFERENCES

1. Petit, F., et al.: Analysis of Critical Infrastructure Dependencies and Interdependencies. Argonne National Lab.(ANL), Argonne. Tech. Rep. (2015)
2. Mell, P., Scarfone, K., Romanosky, S.: Common vulnerability scoring system. IEEE Secur. Priv. Mag. 4(6), 85–89 (2006)
3. O'Donnell, L.: Attackers Exploiting High-Severity Network Security Flaw, Cisco Warns. https://threatpost.com/attackers-exploiting-high-severity-network-security-flaw-cisco-warns/157756/ (2020). Accessed 20 Jan 2021
4. Sullivan, J.E., Kamensky, D.: How cyber-attacks in Ukraine show the vulnerability of the U.S. power grid. Electr. J. 30(3), 30–35 (2017)
5. Langner, R.: Stuxnet: dissecting a cyberwarfare weapon. IEEE Secur. Priv. Mag. 9(3), 49–51 (2011)
6. Grainger, J.J., et al.: Power System Analysis McGraw-Hill Inc, Singapore (2003). https://www.mheducation.com/highered/product/power-system-analysis-grainger-stevenson/9780070612938.html
7. Amin, A.M.: Electricity infrastructure security: Toward reliable, resilient and secure cyber-physical power and energy systems. IEEE PES General Meeting, Minneapolis, 25–29 July 2010. https://doi.org/10.1109/PES.2010.5589488
8. Li, X., et al.: Securing smart grid: cyber attacks, countermeasures, and challenges. IEEE Commun. Mag. 50(8), 38–45 (2012)
9. Zonouz, S., et al.: A security-oriented cyber-physical contingency analysis in power infrastructures. IEEE Trans. Smart Grid. 5(1), 3–13 (2013)
10. Davis, K.R., et al.: A cyber-physical modeling and assessment framework for power grid infrastructures. IEEE Trans. Smart Grid. 6(5), 2464–2475 (2015)
11. Davis, K., et al.: Cyber-physical security assessment (CYPSA) for electric power systems. IEEE-HKN: The Bridge. 112(2), 8–19 (2016). http://hkn.ieee.org/wp-content/uploads/2017/10/Bridge_2_2016.pdf
12. Jin, S., et al: A novel application of parallel betweenness centrality to power grid contingency analysis. IEEE International Symposium on Parallel & Distributed Processing (IPDPS), Atlanta, 19–23 April 2010. https://doi.org/10.1109/IPDPS.2010.5470400
13. Vellaithurai, C., et al.: Cpindex: cyber-physical vulnerability assessment for power-grid infrastructures. IEEE Trans. Smart Grid. 6(2), 566–575 (2014)
14. Kundur, D., et al.: Towards modelling the impact of cyber attacks on a smart grid. Int. J. Secur. Network. 6(1), 2–13 (2011)
15. Yoon, S., et al.: Scalable traffic sampling using centrality measure on software-defined networks. IEEE Commun. Mag. 55(7), 43–49 (2017)
16. Chen, P., Choudhury, S., Hero, A.O.: Multi-centrality graph spectral decompositions and their application to cyber intrusion detection. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, 20–25 March 2016. https://doi.org/10.1109/ICASSP.2016.7472539
17. Kersulis, J.A., et al.: Topological graph metrics for detecting grid anomalies and improving algorithms. Power Systems Computation Conference (PSCC), Dublin, 11–15 June 2018. https://doi.org/10.23919/PSCC.2018.8442682
18. Liu, B., et al.: Recognition and vulnerability analysis of key nodes in power grid based on complex network centrality. IEEE Trans. Circuits and Syst. II: Express Briefs. 65(3), 346–350 (2017)
19. Yan, J., He, H., Sun, Y.: Integrated security analysis on cascading failure in complex networks. IEEE Trans. Inf. Forensics Secur. 9(3), 451–463 (2014)
20. Bompard, E., Pons, E., Di Wu, D.: Extended topological metrics for the analysis of power grid vulnerability. IEEE Syst. J. 6(3), 481–487 (2012)
21. Bompard, E., Napoli, R., Xue, F.: Extended topological approach for the assessment of structural vulnerability in transmission networks. IET Gener., Transm. Distrib. 4(6), 716–724 (2010)
22. Bompard, E., Wu, D., Xue, F.: Structural vulnerability of power systems: a topological approach. Elec. Power Syst. Res. 81(7), 1334–1340 (2011)
23. Wang, X., et al: A network approach for power grid robustness against cascading failures. 7th International Workshop on Reliable Networks Design and Modeling (RNDM), Munich, 5–7 October 2015. https://doi.org/10.1109/RNDM.2015.7325231
24. Wang, Z., Scaglione, A., Thomas, R.J.: Generating statistically correct random topologies for testing smart grid communication and control networks. IEEE Trans. Smart Grid. 1(1), 28–39 (2010)
25. Wang, Z., Scaglione, A., Thomas, R.J.: Electrical centrality measures for electric power grid vulnerability analysis. In: Proceedings of the 49th IEEE Conference on Decision and Control (CDC), pp. 5792–5797. IEEE (2010)
26. Weaver, G.A., et al.: Cyber-physical models for power grid security analysis: 8-substation case. In: Proceedings of the 2016 IEEE International Conference on Smart Grid Communications (SmartGridComm), pp. 140–146. IEEE (2016)
27. Wlazlo, P., et al.: A cyber topology model for the Texas 2000 synthetic electric power grid. In: Proceedings of the 2019 Principles, Systems and Applications of IP Telecommunications (IPTComm), pp. 1–8 (2019)
28. Gaudet, N., et al.: Firewall configuration and path analysis for smartgrid networks. In: Proceedings of the IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR), 1–6 (2020)
29. Morris, T., Vaughn, R., Dandass, Y.: A retrofit network intrusion detection system for modbus rtu and ascii industrial control systems. In: Proceedings of the 2012 45th Hawaii International Conference on System Sciences, pp. 2338–2345. IEEE (2012)
30. Srivastava, A., et al.: Modeling cyber-physical vulnerability of the smart grid with incomplete information. IEEE Trans. Smart Grid. 4(1), 235–244 (2013)
31. Ernster, T.A., Srivastava, A.K.: Power system vulnerability analysis-towards validation of centrality measures. In: Proceedings of the Power and Energy Society Transmission and Distribution (PES T&D) 2012, pp. 1–6. IEEE (2012)
32. Sahu, A., et al.: A framework for cyber-physical model creation and evaluation. In: Proceedings of the 2019 20th International Conference on Intelligent System Application to Power Systems, pp. 1–8. ISAP) (2019)
33. Cyber Physical Resilient Energy Systems. https://cypres.engr.tamu.edu/test-cases (2021). Accessed 20 Jan 2021
34. NP-View. https://www.network-perception.com/np-view (2021). Accessed 20 Jan 2021
35. Wang, J., et al.: A method for information security risk assessment based on the dynamic bayesian network. In: Proceedings of the 2016 International Conference on Networking and Network Applications (NaNA), pp. 279–283 (2016)
36. Cohen, R., Havlin, S.: Complex Networks: Structure, Robustness and Function. Cambridge University Press, New York (2010). https://doi.org/10.1017/CBO9780511780356
37. Freeman, L.C.: A set of measures of centrality based on betweenness. Sociometry. 40, 35–41 (1977)
38. Srivastava, A.K., et al.: Graph-theoretic algorithms for cyber-physical vulnerability analysis of power grid with incomplete information. J. Mod. Power Syst. Clean Energy. 6(5), 887–899 (2018)
39. [Online]. http://publish.illinois.edu/iti-cypsa/cypsa-analysis/
40. Zahariev, P.Z., Hristov, G.V., Iliev, T.B.: Study on the impact of node density and sink location in wsn. In: Technological Developments in networking, Education and Automation, pp. 539–542. Springer (2010)
41. Callaway, E.H., Jr.: Wireless Sensor Networks: Architectures and Protocols. CRC press (2003)

**How to cite this article:** Umunnakwe, A., et al.: Cyber-physical component ranking for risk sensitivity analysis using betweenness centrality. IET Cyber-Phys. Syst., Theory Appl. 6(3), 139–150 (2021). https://doi.org/10.1049/cps2.12010