Needle-in-the-Haystack: On Automatic Risk-based Searching to Stop Cyber-Physical Threats in Large-Scale Power Systems

Katherine Davis, *Senior Member, IEEE*, Shining Sun, *Student Member, IEEE*
Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA
Email: sshh2@tamu.edu, katedavis@tamu.edu

***Abstract*—Power grid infrastructure is cyber-physical in nature where a geographically distributed electrical network is monitored and controlled using a vast network of computers or cyber infrastructure. Understanding this cyber-physical interconnectedness is critical to assessing and improving the risk posture of the power grid with respect to cyber threats. We present an analysis of cyber-physical risk approaches to help utilities improve grid defense: cyber-physical situational awareness and control under adversarial presence. A 200-bus cyber-physical power system is analyzed as a case study. Finally, we discuss opportunities for improved automated risk analysis in decision-making problems, for trustworthy response in emerging cyber-physical-human system-of-systems.**

I. Introduction

Power system defenders need to observe and understand what the assets are and how they are connected before they can start to reason about 'normal' behavior. This work examines how models and data impact the ability to automatically assess risk in large-scale power systems. Specifically, this paper is about risk, and it explores methods and use cases for calculating cyber-physical risk in power systems. The idea is to enhance understanding of different ways to measure risk, the data inputs and outputs for those methods, the key variables and how they can be constructed under different assumptions, and related graph-theoretic and deep learning constructs. We thus aim to elucidate benefits and shortcomings and identify opportunities for resilience-preserving automated risk-based searching to stop cyber-physical threats in large-scale power systems, by finding the *needle-in-the-haystack*.

II. Overview and Related Work

The contributions of this paper are as follows. We explore the risk analysis methods that build upon cyber-physical models and data of power systems, and we consider the content of these systems as test cases for cyber-physical modeling, planning, and threat analysis. Then, we present an example with a synthetic cyber model and full topology expansion performed on top of the synthetic electric grid 200-bus model from the test case archive [1]. The details of the cyber-physical model construction are outside the scope of this paper. Several related works discuss the requirements and contents of these models for large-scale power systems, such as [2] and [3]. Here, we focus only on the risk question, and highlight an example in such a cyber-physical system.

Risk searching is important because highly-connected power systems can inadvertently expose vulnerabilities to intruders that can disrupt grid resilience. Cyber threats and propagation of their impacts can be analyzed by performing cyber-security risk assessment [4], which includes the use of attack graphs. Attack graphs are used to analyze network and host vulnerabilities as well as potential access paths adopted by adversaries to exploit these vulnerabilities to compromise their target [5]. These methods are a type of risk analysis. They require updating the attack graph templates based on real-time alerts from Intrusion Detection Systems (IDS).

Because an adversary's behavior is uncertain, risk can be studied using Bayesian Networks (BNs) which are a type of Probabilistic Graphical Model (PGM) beneficial for modeling attack graphs. There is evidence that BNs are useful in power system environments to perform causal reasoning about an adversary's trajectory. For example, [6] uses BNs to study the impact of cascading effects in coupled infrastructure. Risk analysis research is ongoing to examine and validate the use of Bayesian structural learning and inference (e.g., Cooper and Herskovits K2 [7], Monte Carlo Markov Chain (MCMC) [8], Chow Liu [9], Partially Directed Acyclic Graph (PDAG) [10]) for improved estimation of probabilities for how an adversary can exploit a sequence of vulnerabilities and propagate through the network to cause physical impact. Authors in [11] develop a technique to construct Bayesian Attack Graphs (BAG) for power systems and evaluate structural learning algorithms to update the BAG structure based on real-time alerts, on scalability, data dependency, time complexity and accuracy criteria [11]. A multi-level anomaly detector for Supervisory Control and Data Acquisition (SCADA) is proposed in [12] and extended with a causal polytree-based anomaly reasoning engine in [13] to estimate the security state.

Nodal criticality plays an important role in vulnerability assessment. Hence, critical asset ranking and risk evaluation in cyber-physical power systems have been addressed over the past decade, with works including Security-Oriented Cyber-Physical Contingency Analysis (SOCCA) [14], Cyber-Physical Modeling and Analysis (CPMA) [3], and Cyber-Physical Sit-

uational Awareness or Cyber-Physical Security Assessment (CyPSA) [15]. A framework for method comparisons, risk analysis, with results, is given in [16].

The algorithm(s), data structures, and functions that we refer to as a risk model include physical impact modeled through detailed power system simulations, e.g., [17], and likelihood which can be modeled as ease of access from vulnerability and connectivity information and attack graph analyses: these are the essential components of CyPSA [15].

Techniques to develop and utilize risk models also include a stochastic Bayesian model to calculate cyber-physical security index [18], an expected load curtailment index for protection devices [19], a boolean logic driven MDP leveraging estimated values of a step's success [20], and Bayesian attack graphs (BAGs) [11]. These risk methods produce a model that helps understand and explore static vulnerabilities in the system, e.g., to rank assets and contingencies. Other related risk methods are Bayesian [21], [22], [23], and utilize graph theory [21], [22], [23], [24].

Current techniques for locating critical nodes are mostly dependent on physical topologies [25], like betweenness centrality [24], k-shell [26] and information entropy [27]. In [24], the risk method uses betweenness centrality (BC) and vulnerability scores to formulate a cyber-physical between centrality (CPBC) index to assess vulnerability and risk levels. This approach provides a network-centric perspective for nodal risk assessment, however, it could underestimate risk associated with low-centrality nodes.

III. Case Study of 200-bus Model

The importance of a cyber-physical systems (CPS) modeling approach for power system operation and control has helped to bridge gaps in power systems toward improving the system control [28]. Previous research has focused on creating cyber-physical models to manage power system contingencies and improve grid operation reliability [14], [15]. A detailed cyber-physical model was created in [2] utilizing the Texas 2000 synthetic grid. Recently, the relationships between the system and the information flow are then examined using graph theory [29], [30], [31]. Here, the synthetic 200-bus power system based on the Illinois footprint is used for the risk assessment process. As shown in Fig. 1, the synthetic network contains 200 buses, 245 transmission lines and substations [1]. The network topology is constructed for monitoring and communication, as shown in Fig. 2.

Equipment such as Ethernet switches (SW), protection relays (R), and human-machine interfaces (HMIs) are found at the substation level. The loads, buses, and shunts are protected by the relays. Data is sent to the Utility Control Center (UCC) and Independent System Operator (ISO) levels via substation firewalls. Additional firewalls are present at the UCC level, enforcing security policies to safeguard the network. The SCADA data is obtained at the ISO level and used for centralized control and monitoring.

To clearly depict research results on cyber-physical risk, and to illustrate how risk relates to the cyber-physical dependencies, we establish a cyber-physical graph. This graph gives a visual representation of the interconnected network that encompasses both the cyber (computational and communication) and physical (electrical components and infrastructure) parts of the system. The cyber-physical graph of the 200-bus system is generated based on the physical and cyber topology, with the 1,613 nodes and 5,995 edges.

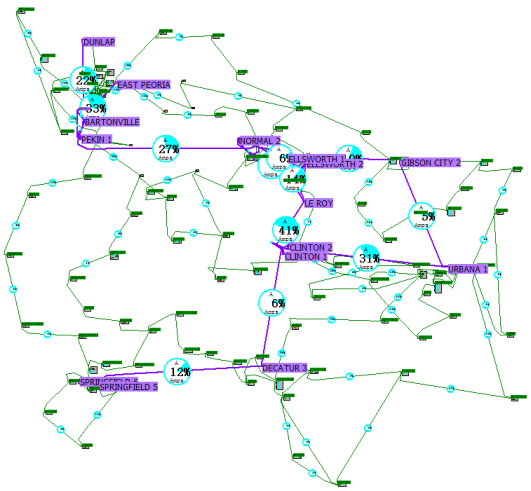

Figure 1. The synthetic 200-bus system [1] one-line diagram.

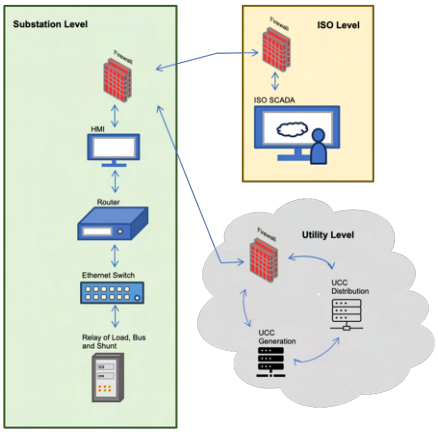

Figure 2. The 200-bus system cyber topology.

By giving edges exploitability scores, one can efficiently model complexity of an adversary's path between nodes. More sophisticated analytics can make use of exploitability-weighted graphs to simulate possible intrusion pathways and determine the most likely paths an adversary could take. In this work, we cite the Common Vulnerability Scoring System (CVSS) from National Vulnerability Database (NVD) and assign the network edges exploitability ratings to the graph edge weights based on access complexity (AC) value [32], the degree of difficulty in exploiting a vulnerability, where a lower score denotes a more difficult path. The full CPS graph is shown in Fig. 3.

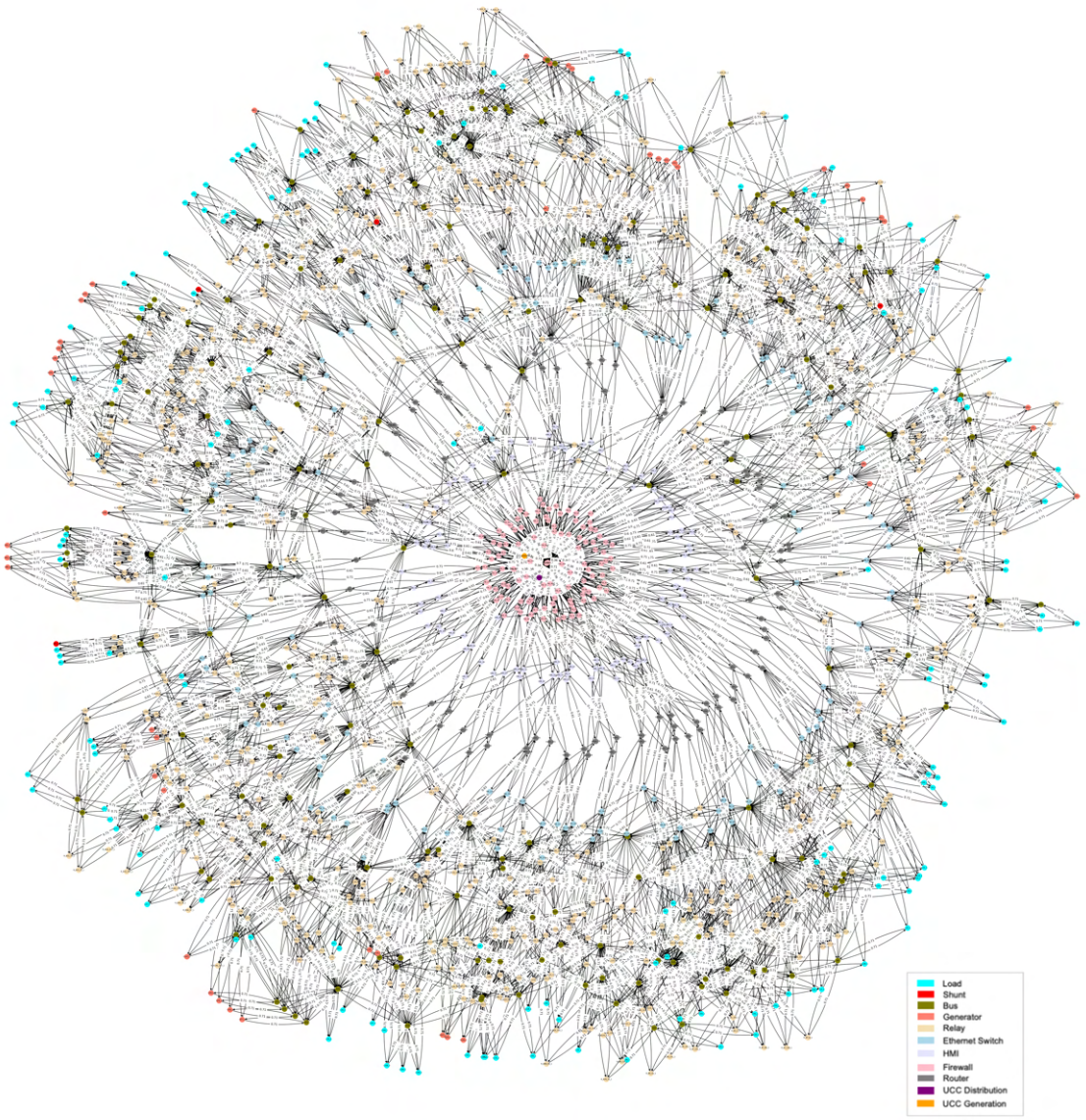

Figure 3. The 200-bus system cyber-physical graph representation.

IV. Method

According to NIST [28], risk is commonly described as the possibility of a certain threat and is calculated by multiplying the potential vulnerability by the impact of resultant incidents, commonly represented as $Risk = Likelihood \times Impact$.

Previous work [31] employs a DeepWalk-based strategy to categorize nodes according to risk severity. The DeepWalk method operates on the assumption that each node has a homogeneous connection, and that every path, regardless of its characteristics, carries the same probability. This approach captures the structural similarities and relationships between nodes, but it may be insufficient for prompt intrusion detection considering different threat complexities. To address this, a risk analysis framework is proposed in Fig. 4, with a cyber-physical graph as the input. We can determine nodal vulnerabilities within the network topology and the probability of possible access pathways, using a biased random walk step, by utilizing the Node2Vec technique. This framework goes beyond calculating likelihood and incorporates effect evaluation to align with the unique operational and security needs of CPS.

A. Component Relationships and Similarity

The Node2Vec approach may decipher intricate relationships between different system components. To simulate graph exploration, random walks are first created. We use this strategy to reflect the random walks as the adversaries' access pathways. Next, within these walks, the skip-gram model is used to forecast a target node's neighbors. The skip-gram technique is particularly significant in representative learning [33], since it can provide latent variables that can be used to represent

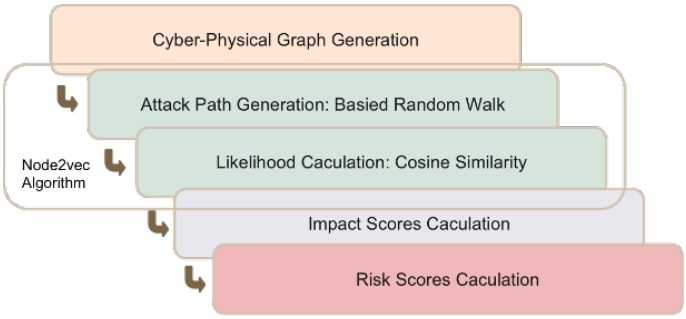

Figure 4. A risk analysis framework.

abstract notions that are not directly measurable and to reveal hidden structures in data. It helps scientists working with data to draw conclusions about hidden or fundamental mechanisms influencing visible occurrences.

A random walk on a weighted CPS graph is defined as a series of steps that start at a vertex, represented as v_i. In accordance with the DeepWalk technique, which applies unbiased random walks, the first- and second-order transition probabilities operate sequentially to integrate the searching algorithm [34]. The definition of the second-order transition probability is as follows: the overall transition probability, when initiating from node *v* and proceeding across the edge (*t,v*), is modulated by the search bias factor.

$$P(u,v) = \frac{\alpha_{pq}(t,u) \times w(u,v)}{\sum_{\hat{v} \in V} \alpha_{pq}(t,u) \times w(u,v)} \tag{1}$$

The quality of the latent variables representing the nodes' neighborhood information and network topology is determined by modifications to the aforementioned parameters. The balance between local and global neighborhood exploration can be adjusted.

Node embeddings are generated to symbolize the graph structure and represent the biased random walks. For each vertex $u \in V$, the likelihood of identifying neighboring nodes of node u is determined through the optimization of the objective function, which is achieved by employing the stochastic gradient descent algorithm.

$$\max_{\mathbf{f}} \sum_{u \in V} \left[-\log Z_u + \sum_{n_i \in N_S(u)} \mathbf{f}(n_i) \cdot \mathbf{f}(u) \right]. \tag{2}$$

The "proximity" nodes in the embedding space could then be defined as cosine similarity of their latent variables. For any two nodes *u* and *v* in the graph, their cosine similarity is calculated using their vector embeddings $\phi(u)$ and $\phi(v)$. The likelihood is estimated by:

$$\text{cosine similarity}(\phi(u), \phi(v)) = \frac{\phi(u) \cdot \phi(v)}{\|\phi(u)\| \|\phi(v)\|} \tag{3}$$

The similarity score produced by the likelihood computation ranges from -1 to 1. A node's tasks and locations within the CPS network are likely comparable when its score is near 1. Significantly different or opposing embeddings are indicated by a number around -1, whereas no discernible similarity is indicated by a score near 0.

B. Impact Scores

An essential component of risk analysis is determining the impact score of CPS components. It entails assessing the possible outcomes in the event that a device were compromised. In this research, we present a comprehensive metric to calculate the impact score. As shown by Eq. 4, these dimensions are network influence (NI), operational importance (OI), security and confidentiality (SC), and access complexity (AC). The "sum of all scores" denotes the sum of OI, SC, MR, NI, and AC for each component.

$$\text{Impact Score} = \frac{\text{OI} \cdot \text{SC} \cdot \text{MR} \cdot \text{NI} \cdot \text{AC}}{\sum_{i=1}^{5} \text{all score}_i} \tag{4}$$

Table I demonstrates a structured framework in assessing impact sore by various components. Each category is defined at three levels: *High*, *Medium* and *Low* with scores of 0.9, 0.6 and 0.3 respectively. OI scores devices based on their significance in network operations and integration complexity. SC measures the level of the device's role in network security and its handling of sensitive data. AC varies based on the difficulty of accessing components. MR scores the maintenance needs and reliability of the device. BC is used to indicate the nodal network influence; it measures the frequency at which a node appears on the shortest paths. We define the top 25 nodes of BC scores as the *High* level of network influence, the bottom 25 as *Low* and the rest as *Medium*.

Table I
COMPONENT RISK LEVELS BASED ON OPERATIONAL AND SECURITY FACTORS

Component	Operational Importance	Security and Confidentiality	Maintenance and Reliability	Access Complexity
Bus (B)	Medium (0.6)	Low (0.3)	Medium (0.6)	High (0.9)
Generator (G)	High (0.9)	Low (0.3)	High (0.9)	High (0.9)
Load (L)	Low (0.3)	Low (0.3)	Low (0.3)	High (0.9)
Shunt (S)	Medium (0.6)	Low (0.3)	Low (0.3)	High (0.9)
Relay (R)	Medium (0.6)	Medium (0.6)	Medium (0.6)	Medium (0.6)
Ethernet Switch (SW)	Medium (0.6)	Medium (0.6)	Medium (0.6)	Medium (0.6)
HMI	Medium (0.6)	High (0.9)	Medium (0.6)	Medium (0.3)
Firewall (FW)	Medium (0.6)	High (0.9)	Medium (0.6)	Low (0.3)
Router (r)	High (0.9)	High (0.9)	Medium (0.6)	Low (0.3)
Control Center Devices (CC)	High (0.9)	High (0.9)	High (0.9)	Low (0.3)

V. ANALYSIS AND RESULTS

We assume that Denial of Service (DoS) attacks can occur on any physical and network component in the CPS network and can be recognized. As a result of the intrusion, we assume that one component is in a failed state, and we are interested in the adversary's next target. We consider a situation where a DoS is occurring, and the specific target has been pinpointed: a DoS targeting the Ethernet Switch (SW_4) at Substation 2.

Tables II shows the results of the top 10 nodes from risk assessments, where we focus the results on the substation level. For example, for SW_4 in Substation 2 as target, $router\_2$, R_L4_1, and $R\_B4\_3\_1$, could be directly affected, leading

to operational and security challenges within Substation 2. Moreover, the high-risk scores of physical nodes like *B3*, *B137*, and *B57* suggest an impact on physical operations within substation 2 and neighboring substations. Similarly, the impact of compromising *B_57* and *R_B57_3_1* can have ripple effects throughout the network, affecting other substations and nodes, especially those with high operational importance and interconnected functionalities.

The violin plot displays the distribution of risk scores across each substation. The wider parts of the violins indicate a higher density of nodes at certain risk scores within the substation. For instance, in Fig. 5, Substation 2 shows a wide distribution with a peak around a risk score of 0.6. This indicates a high frequency of nodes around this risk score, suggesting that Substation 2 has a substantial number of nodes at a high risk level when *SW_4* is targeted, which may warrant prioritized risk management strategies. Substations with narrower violins, like 26 and 71, might have risk scores in the middle range and could be classified as moderate risk. The rest of the substations have a higher density of data points at the lower end of the risk score range and lower median risk scores.

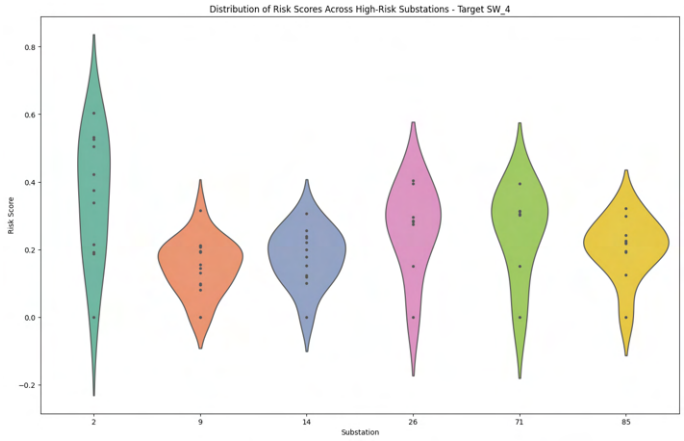

Figure 5. Violin Plot with 'SW_4' in Sub2 as Target

VI. Discussion: Assumptions and How it Works

The operation of a power grid is governed by the laws of physics, and the underlying power system is represented as a set of nonlinear AC equations that include active (MW) and reactive (MVAr) power flows. Power balance is expressed by the power injection mismatch vector $f(x,u) = [\Delta p, \Delta q]^T$ that must equal zero, where x is the power system state, a vector of bus voltage magnitudes V and angles θ, and u is a vector of controls such as generator outputs P_i^g, Q_i^g. The power balance requirement is enforced in the AC power flow solution. The complex power flow into each line terminal $(l,m) \in \mathcal{L}$ is denoted by $P_{ij} + jQ_{ij}$, and $\theta_{ij} = \theta_i - \theta_j$ for $(i,j) \in \mathcal{L}$ (Eqn. 5-6). The number of buses is n, and $G + jB$ is the system admittance matrix which contains the network line parameters from $I = Ybus \cdot V$.

$$P_{ij} = V_i^2 \left[G_{ij}\right] + V_i V_j \left[G_{ij} \cos\left(\theta_{ij}\right) + B_{ij} \sin\left(\theta_{ij}\right)\right] \quad (5)$$

$$Q_{ij} = -V_i^2 \left[-B_{ij}\right] + V_i V_j \left[G_{ij} \sin\left(\theta_{ij}\right) - B_{ij} \cos\left(\theta_{ij}\right)\right] \quad (6)$$

The idea is that a navigation agent can assess and then utilize a risk value function, to take actions that protect the system, dictated by a Bellman equation that captures the risk. The agent starting from s obtains expected return, G, taking action a, and following policy π. The agent finds and takes the sequence of actions that will maximize the reward.

$$Q_\pi(s,a) = E_\pi \left[\sum_{j=0}^{T} \gamma r_{t+1+j}^{j} | S_t = s, A_t = a \right] \quad (7)$$

$$V_\pi(s) = E_\pi \left[G_t | s = s_t\right] = \sum_a (\pi(a|s)) \cdot Q_\pi(s,a) \quad (8)$$

An MDP is a discrete-time stochastic process used to describe the agent and environment interactions. The behavior of the agent in the environment is defined by the MDP's states (S), actions (A), state transition model $P(s_{t+1}|s_t, a_t)$, reward model $R(s_{t+1}|s_t, a_t)$, and discount factor γ. The state transition probabilities model the uncertainty of when the agent performs action a in a state s. Q is the expected value of the reward which captures that fact that it is a stochastic environment (Eq. 7), and $\pi(a|s)$ is the probability that policy π selects action a given current state s; these must sum to one: $\sum_a (\pi(a|s)) = 1$. The value function is an estimate of how good it is for the agent to be in a given state, which is the expected total reward, $V_\pi(s)$ (Eq. 8).

The reward function that controls the risk search captures the worst physical violations (e.g., Eq. 5) and access paths. How to find V is given by the Bellman equation for value function (Eq. 9):

$$V_\pi(s) = \sum_a (\pi(a|s)) \cdot \sum_{s' \in S} p\left(s'|s,a\right) \left(\Delta F(s,s') + \gamma V_\pi(s')\right) \quad (9)$$

The first term $\Delta F(s'|s,a) = F(s') - F(s)$ measures the change in severity of "physical consequence" in going from state s to state s', the immediate expected reward. Then, $V_\pi(s')$ is the value of the next state, and $p\left(s'|s,a\right)$ reflects the difficulty to perform vulnerability exploitations on the path.

Now, if we find a *Needle-in-the-Haystack* using our automatic risk searching, then, next we can also automatically mitigate the risk, if such a mitigation exists and is not too costly. This is *Wayfinder*, which can tune the system according to risk tolerance, and will be detailed in a follow up paper.

VII. Conclusion

This paper investigates improved ways to automatically find and mitigate hidden threats and explores achieving this goal for using a 200-bus test case. Finally, it suggests next steps on CPS risk analysis and control.

Acknowledgements

The authors acknowledge the TEES Smart Grid Center, US Department of Energy DE-OE0000895 and DE-CR0000018, the National Science Foundation 2220347, and those project teams.

Table II
RISK ASSESSMENT RESULT FOR HIGH-RISK NODES WITH 'SW_4' AS TARGET

Node ID	Cosine Sim.	BC	BC Class.	Network Risk	Operational Importance	Security & Confid.	Maint. & Reliab.	Access Complex.	Impact Score	Risk Score	Substation No#
router_2	0.934	0.0054	Medium	0.6	0.9	0.9	0.6	0.3	0.647	0.604	2
R_L4_1	0.935	0.0014	Medium	0.6	0.6	0.6	0.6	0.6	0.569	0.532	2
R_B4_3_1	0.926	0.0014	Medium	0.6	0.6	0.6	0.6	0.6	0.569	0.527	2
B3	0.779	0.0098	High	0.9	0.6	0.3	0.6	0.9	0.647	0.504	2
B137	0.652	0.0091	High	0.9	0.6	0.3	0.6	0.9	0.647	0.422	2
B57	0.624	0.0084	High	0.9	0.6	0.3	0.6	0.9	0.647	0.404	26
R_B57_3_1	0.694	0.0010	Medium	0.6	0.6	0.6	0.6	0.6	0.569	0.395	26
R_B137_3_1	0.693	0.0011	Medium	0.6	0.6	0.6	0.6	0.6	0.569	0.394	71
B4	0.912	0.0009	Medium	0.6	0.6	0.3	0.6	0.9	0.412	0.375	2
HMI_2	0.822	0.0065	Medium	0.6	0.6	0.9	0.6	0.3	0.412	0.339	2

REFERENCES

[1] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Grid structural characteristics as validation criteria for synthetic networks," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3258–3265, 2017.

[2] P. Wlazlo, K. Price, C. Veloz, A. Sahu, H. Huang, A. Goulart, K. Davis, and S. Zounouz, "A cyber topology model for the texas 2000 synthetic electric power grid," in *Proceedings of 2019 Principles, Systems and Applications of IP Telecommunications (IPTComm)*. IEEE, 2019, pp. 1–8.

[3] 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 Transactions on Smart Grid*, vol. 6, no. 5, pp. 2464–2475, 2015.

[4] J. Wang, K. Fan, W. Mo, and D. Xu, "A method for information security risk assessment based on the dynamic bayesian network," in *2016 International Conference on Networking and Network Applications*, 2016, pp. 279–283.

[5] S. Jha, O. Sheyner, and J. Wing, "Two formal analyses of attack graphs," in *Proceedings 15th IEEE Computer Security Foundations Workshop. CSFW-15*, 2002, pp. 49–63.

[6] N. HadjSaid, C. Tranchita, B. Rozel, M. Viziteu, and R. Caire, "Modeling cyber and physical interdependencies - application in ict and power grids," in *2009 IEEE/PES Power Systems Conference and Exposition*.

[7] P. Spirtes, C. Glymour, and R. Scheines, *Causation, Prediction, and Search*, 01 1993, vol. 81.

[8] G. F. Cooper and E. Herskovits, "A bayesian method for the induction of probabilistic networks from data," *Machine Learning*, vol. 9, 1992.

[9] C. Chow and C. Liu, "Approximating discrete probability distributions with dependence trees," *IEEE Transactions on Information Theory*, vol. 14, no. 3, pp. 462–467, 1968.

[10] P. Giudici and R. Castelo, "Improving markov chain monte carlo model search for data mining," *Machine Learning*, pp. 127–158, 01 2003.

[11] A. Sahu and K. Davis, "Structural learning techniques for bayesian attack graphs in cyber physical power systems," in *2021 IEEE Texas Power and Energy Conference (TPEC)*, 2021, pp. 1–6.

[12] W. Ren, T. Yardley, and K. Nahrstedt, "Edmand: Edge-based multi-level anomaly detection for scada networks," 10 2018, pp. 1–7.

[13] W. Ren, T. Yu, T. Yardley, and K. Nahrstedt, "Captar: Causal-polytree-based anomaly reasoning for scada networks," in *2019 IEEE SmartGridComm*, 2019, pp. 1–7.

[14] S. Zonouz, C. Davis, K. Davis, R. Berthier, R. Bobba, and W. Sanders, "Socca: A security-oriented cyber-physical contingency analysis in power infrastructures," *Smart Grid, IEEE Transactions on*, vol. 5, no. 1, pp. 3–13, Jan 2014.

[15] K. Davis, R. Berthier, S. Zonouz, G. Weaver, R. Bobba, E. Rogers, P. Sauer, and D. Nicol, "Cyber-physical security assessment (cypsa) for electric power systems," *IEEE-HKN: THE BRIDGE*, 2016.

[16] A. Sahu, H. Huang, K. Davis, and S. Zonouz, "A framework for cyber-physical model creation and evaluation," in *2019 20th International Conference on Intelligent System Application to Power Systems (ISAP)*, 2019, pp. 1–8.

[17] C.-W. Ten, A. Ginter, and R. Bulbul, "Cyber-based contingency analysis," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 3040–3050, 2016.

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

[19] X. Liu, M. Shahidehpour, Z. Li, X. Liu, Y. Cao, and Z. Li, "Power system risk assessment in cyber attacks considering the role of protection systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 572–580, 2017.

[20] S. Kriaa, M. Bouissou, and L. Pièvre-Cambacédès, "Modeling the stuxnet attack with bdmp: Towards more formal risk assessments," in *2012 7th International Conference on Risks and Security of Internet and Systems (CRiSIS)*, 2012, pp. 1–8.

[21] A. Sahu and K. Davis, "Inferring adversarial behaviour in cyber-physical power systems using a bayesian attack graph approach," *IET Cyber-Physical Systems: Theory & Applications*, 2023.

[22] X. Lyu, Y. Ding, and S.-H. Yang, "Bayesian network based c2p risk assessment for cyber-physical systems," *IEEE Access*, vol. 8, pp. 88 506–88 517, 2020.

[23] X. Zhang and D. Zhang, "Quantitative risk assessment of cyber physical power system using bayesian based on petri net," in *2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS)*, 2018, pp. 988–992.

[24] A. Umunnakwe, A. Sahu, M. R. Narimani, K. Davis, and S. Zonouz, "Cyber-physical component ranking for risk sensitivity analysis using betweenness centrality," *IET Cyber-Physical Systems: Theory & Applications*, vol. 6, no. 3, pp. 139–150, 2021.

[25] B. Fan, C.-X. Zheng, L.-R. Tang, and R.-Z. Wu, "Critical nodes identification for vulnerability analysis of power communication networks," *IET Communications*, vol. 14, no. 4, pp. 703–713, 2020.

[26] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse, "Identification of influential spreaders in complex networks," *Nature physics*, vol. 6, no. 11, pp. 888–893, 2010.

[27] A. Zareie, A. Sheikhahmadi, and A. Fatemi, "Influential nodes ranking in complex networks: An entropy-based approach," *Chaos, Solitons & Fractals*, vol. 104, pp. 485–494, 2017.

[28] "Multiyear plan for energy sector cybersecurity," Department of Energy, Office of Electricity Delivery and Energy Reliability, Tech. Rep., Mar. 2018. [Online]. Available: https://www.energy.gov/ceser/articles/doe-multiyear-plan-energy-cybersecurity

[29] N. Jacobs, S. Hossain-McKenzie, A. Summers, C. B. Jones, B. Wright, and A. Chavez, "Cyber-physical observability for the electric grid," in *2020 IEEE Texas Power and Energy Conference (TPEC)*, 2020, pp. 1–6.

[30] S. Hossain-McKenzie, N. Jacobs, A. Summers, R. Adams, A. Chatterjee, A. Layton, K. Davis, and H. Huang, "Towards the characterization of cyber-physical system interdependencies in the electric grid," 2023.

[31] S. Sun, E. Payne, A. Layton, K. Davis, S. Hossain-McKenzie, and N. Jacobs, "Bio-inspired and ai deepwalk based approach to understand cyber-physical interdependencies of power grid infrastructure," in *IEEE 55th North American Power Symposium (NAPS)*, Oct 2023.

[32] National Vulnerability Database. (2023) Common vulnerability scoring system v2 calculator. Accessed: 2023-12-16. [Online]. Available: https://nvd.nist.gov/vuln-metrics/cvss/v2-calculator

[33] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," *CoRR*, vol. abs/1607.00653, 2016. [Online]. Available: http://arxiv.org/abs/1607.00653

[34] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deepwalk: Online learning of social representations," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014, pp. 701–710.