ARTICLE

Performance Evaluation Methods and Analysis of Deep Learning Frameworks in Distributed Power Systems

Hao Huang[1,*], **Biao She**[2], **Wei Zhang**[2] **and Ge Pan**[2]

[1] Nanjing Suyi Industry Co., Jiangsu Xinshun Energy Industry Group Co., Ltd: No. 22, West Beijing Road, Nanjing, China
[2] Nanjing Huaqun Energy Group Co., Ltd, No. 251 Zhongshan Road, Gulou District, Nanjing, China

Abstract

This paper presents a vulnerability assessment method for distributed electric power information systems, focusing on the unique characteristics of distributed IP networks and the cumulative risks from exploiting vulnerabilities. To address these challenges, a Hidden Markov Risk Assessment Model (HMRAM) is proposed, incorporating penetration depth and node dependencies to evaluate the cumulative impact of vulnerabilities. This approach provides network managers with a comprehensive understanding of penetration risks and enables prompt actions to enhance system security. The conclusion emphasizes the effectiveness and applicability of the methodology in improving network security, while a technique for ranking vulnerabilities by risk levels is suggested to optimize system maintenance costs and resilience.

Keywords: vulnerability assessment, distributed power information systems, IP networks

Citation
Hao Huang, Biao Shen, Wei Zhang and Ge Pan (2024). Performance Evaluation Methods and Analysis of Deep Learning Frameworks in Distributed Power Systems. Mari Papel Y Corrugado, 2024(1), 159–164.

Submitted: 01 September, 2024
Accepted: 29 October, 2024
Published: 30 December, 2024

Vol. 2024, **No.** 1, 2024.

***Corresponding author:**
✉ Hao Huang
13770348401@163.com

1 Introduction

Electric power information systems are becoming increasingly significant in today's society due to the continuous advancements in information technology [1]. The growth of the social economy and the safety of people's lives heavily rely on the reliable operation of power information systems, which are a vital component of infrastructure. However, as the size and complexity of power information systems grow, system security faces more significant challenges, making vulnerability assessment and risk analysis major concerns [2].

A power information system is a collection of hardware and software tools designed to manage, control, and monitor a power system's performance. With the emergence of new technologies such as renewable energy, electric vehicles, and smart grids, power information systems are expanding in both scope and functionality [3]. Modern power information systems typically employ a distributed architecture composed of numerous nodes and subsystems to enable real-time power system monitoring, scheduling, and control [4].

While the development of power information systems has greatly benefited power generation and delivery, several challenges remain. First, the security of electric power information systems, as critical infrastructure, is paramount. As cyberattack technology evolves, these systems face constant threats from ransomware, phishing, and malware attacks. Second, the distributed architecture of power information systems creates a vast number of nodes and interfaces, which increases system complexity and vulnerability to attacks [5]. Furthermore, the operational stability of electric power information systems is crucial for socio-economic development and public safety. A breach or failure in the system could lead to catastrophic impacts on electricity production and delivery, as well as significant accidents.

Risk analysis and vulnerability assessment are essential components of electric power information system security. Vulnerability assessment aims to identify the existing weaknesses in the system, such as software flaws, configuration issues, and security gaps, establishing a foundation for subsequent security protection and reinforcement efforts [6]. Risk analysis, on the other hand, evaluates the various potential hazards and threats the system may face, including cyberattacks, natural disasters, and human-caused harm. By conducting vulnerability assessment and risk analysis, the security condition of the electric power information system can be thoroughly understood, potential security issues can be identified and addressed promptly, and the system's overall security and stability can be improved [7].

To enhance the security and stability of power information systems, this research investigates vulnerability assessment and risk analysis methods in distributed architectures. Specifically, this study explores the following areas:

1. To examine the security issues and challenges in electric power information systems, such as cyberattacks, vulnerability exploitation, and system failures.
2. To discuss the importance and role of risk analysis and vulnerability assessment in ensuring the security of electric power information systems.
3. To propose a distributed architecture-based approach for risk assessment and vulnerability evaluation in electric power information systems and provide a detailed explanation of its concept and implementation procedures.
4. To evaluate the feasibility and effectiveness of the proposed approach through case studies and experimental verification.

2 Dependencies Between Nodes

An attacker must perform a series of actions to effectively breach the target node during the vulnerability-penetration process, which requires various strategies. There is a clear sequential logical relationship between these actions since the penetration of the target node is contingent upon the penetration of preceding nodes [8]. In the penetration path, this relationship typically refers to the dependency of later nodes on earlier ones. While every node in the penetration path directly or indirectly contributes to achieving the final objective, the penetration of a preceding node does not necessarily imply that the final goal has also been achieved. The magnitude of this contribution can be quantified using the dependency relationship between nodes [9, 10]. Although network node correlation is mentioned in the literature to describe this relationship, a quantitative approach based solely on subjective evaluation does not accurately depict the dependency between nodes. Therefore, it is necessary to recalculate the dependency between nodes.

Definition 1. *A target node e_j is said to be directly dependent on node e_i if the completion of e_j directly depends on e_i. This means that the successful penetration of e_i is a necessary condition for the completion of e_j. This direct dependency is denoted as $e_j \Rightarrow e_i$. The direct dependency index indicates the degree of dependency between these nodes. Direct dependencies are typically represented in a percolation network by adjacent nodes. In the percolation graph, the direct dependency between nodes is represented by $RF_{j\Rightarrow i}$. These direct dependencies can be expressed using an H-dimensional matrix RF. The matrix element $RF(k,i,j)$ indicates the direct dependency index of node e_i on e_j in the kth path, represented as $RF(k,i,j) = RF_{j\Rightarrow i}$.*

Definition 2. *Indirect dependency occurs when the completion of target node e_j indirectly depends on the penetration of node e_i. In this case, e_i serves as an indirect prerequisite for the penetration of e_j. This indirect dependency is denoted as $e_j \rightarrow e_i$ and is expressed as a non-adjacent node relationship. In the percolation graph, all nodes' indirect dependencies can be represented using an S-dimensional matrix RP. The existence of an indirect dependency index of node e_i on e_j in the kth path is indicated by the matrix element $RP(k,i,j) = RP_{j\rightarrow i}$.*

The dependency of e_j on e_i can fluctuate based on direct dependencies, which cannot be adequately captured using a simple yes-or-no criterion. The degree to which node e_j is dependent on node e_i in a direct dependency relationship is represented by a dependency index.

Definition 3 (Dependency Index)**.** *The dependency index a_{rf} represents the degree of node e_j's dependence on node e_i in a direct dependency relationship. It takes a value in the range $(0,1)$, where $a_{rf} = 1$ indicates full dependency, and $a_{rf} < 1$ indicates partial dependency.*

3 Steps of the Assessment Based on Hidden Markov Modelling

After completing the node dependencies, the evaluation is performed using Hidden Markov Models (HMM), which solve the issue of determining the maximum penetration path by monitoring the node states.

3.1 Categories of State Transfer Matrices

Weakness penetration path evaluation can be viewed as a time-dependent, left-to-right type of HMM [11, 12]. Unlike the traversal-type HMM, this model does not return to the initial state; instead, it progresses states or remains in a state unmodified, moving from left to right in chronological sequence. The state transfer matrix is represented as an upper triangular matrix, with the termination state indicated in the last row.

If there is no self-transfer in the termination state, the last row consists entirely of zeros. If self-transfer occurs, all values in the last row except the last one are zeros. For vulnerability assessment, the dependencies between n nodes are represented by a square matrix in the HMM state transfer matrix. The matrix takes the following form:

$$A = \begin{bmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1m} \\ 0 & 1 & a_{23} & \cdots & a_{2m} \\ 0 & 0 & 1 & \cdots & a_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad i \in [1..m], \, j \in [1..m]. \tag{1}$$

In this matrix, the state transfer probability is $a_{ij} = 0$ when $i > j$, and $a_{ij} = 1$ when $i = j$.

Weakness penetration typically occurs as a result of a state change process, which begins with the initial state of no permission, progresses through several penetration depth stages, and ends with complete control over the target. The state space of weakness penetration depth is represented as $\theta = \{\theta_0, \theta_1, \theta_2, \theta_3, \theta_4\}$.

The operation state of the weakness penetration degree is concealed due to the uncertainty surrounding the degree of penetration. It is only possible to infer the evaluation result of the concealed state from observable values; the concealed state cannot be directly obtained. Figure 1 illustrates the state transition diagram for the degree of weakness penetration.

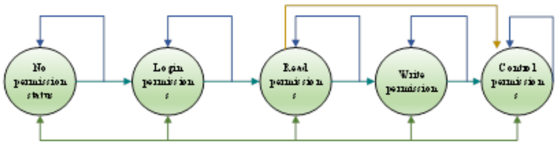

Figure 1. Transition diagram of weakness penetration depth state

Since the penetration process is irreversible, this discussion does not address the dynamic repair of weaknesses during the transition process. Consequently, the state shift involves only a gradually deeper penetration depth and is restricted to a left-to-right state transition. The repair of the penetration depth to revert to the state of no authority or to gradually lessen the situation is not considered. Additionally, each state has the ability to transition to itself.

3.2 Hidden Markov Solution

By monitoring node states, hidden Markov models (HMMs) identify the most probable penetration paths. During the practical assessment, the observed node information is treated as the observation, while the assessment object's state—referred to as its implied state—is the target of the analysis [13, 14]. By dynamically evaluating the cost likelihood of each penetration path, this method connects the observation sequence to the implied state using a series of probability distributions. As a result, the most likely weak penetration path can be determined, and the real penetration path cost can be assessed based on the observed data [15].

First, the parameters used in the vulnerability assessment approach based on HMMs are defined. Let the HMM model for the vulnerability assessment be $\lambda = (\pi, A, B)$, where, W represents the number of states in the HMM and N represents the number of penetration threat levels a node can reach, with $\theta = \{\theta_0, \theta_1, \theta_2, \theta_3, ..., \theta_N\}$.

Each penetration state q_t corresponds to a specific threat level, where $q_t \in \{\theta_0, \theta_1, \theta_2, \theta_3, \theta_4\}$. The penetration depth classification system divides threat levels into five categories. The threat level, represented by $\{\theta_0, \theta_1, \theta_2, \theta_3, \theta_4\}$, progressively increases with the severity of the level.

Let M denote the total number of observations that can be made for each state. In this case, M corresponds to the total number of nodes in the network. Monitoring each node's state is essential for identifying the most likely attack path and associated costs. Assume the observation sequence for each node is represented as $\{V_0, V_1, V_2, V_3, \cdots, V_M\}$, with the observation at time t denoted by $O = \{O_0, O_1, ..., O_m\}$, where each observation value corresponds to $\{V_0, V_1, V_2, V_3, \cdots, V_M\}$.

The observation matrix K can be expressed as:

$$K = \begin{bmatrix} V_{t1}(a_0) & V_{t1}(a_1) & V_{t1}(a_2) & V_{t1}(a_3) & V_{t1}(a_4) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ V_{tm}(a_0) & V_{tm}(a_1) & V_{tm}(a_2) & V_{tm}(a_3) & V_{tm}(a_4) \end{bmatrix}, \tag{2}$$

where $V_{tj}(a_i)$ represents the observed condition of a node at any specific time t, indicating the potential probability of the observation for each node under various penetration paths and penetration depths.

For each observation, the conformity probability of the data series to the given HMM model λ is represented as $p(K|\lambda)$.

4 Testing and Analysis

For testing purposes in this experiment, we deployed a centralized network information simulation to capture power grid faults. The network model was streamlined to simplify computations. The workstation and web server area in the security zone is located on the left side of the network structure, while the production data server area is situated in the Security II zone on the right [16]. Authorized users can access the power grid information network externally via the firewall. It should be noted that the power grid information network extranet, as shown in Figure 2, generally refers to the area outside the power grid information intranet rather than the Internet.

A total of 56 vulnerabilities, comprising 1 high-risk vulnerability, 24 medium-risk vulnerabilities, and 31 low-risk vulnerabilities, were identified during this periodic vulnerability self-assessment. To evaluate the degree of danger associated with these vulnerabilities, we referenced the CVE database. Logical penetration diagrams were created, and impact evaluations were conducted for vulnerabilities of varying risk levels, including CVE-2004-0786, CVE-1999-0516, CVE-1999-0517, CVE-2005-2491, CVE-2005-2970, CVE-2004-0747, and CVE-2004-0493. Using the CVE-2004-0786 vulnerability as an example, the logical penetration diagram is demonstrated in the following content. The primary objective of the CVE-2004-0786 vulnerability is to exploit a remote IPv6 buffer weakness in the Apache Web Server to attack the web server.

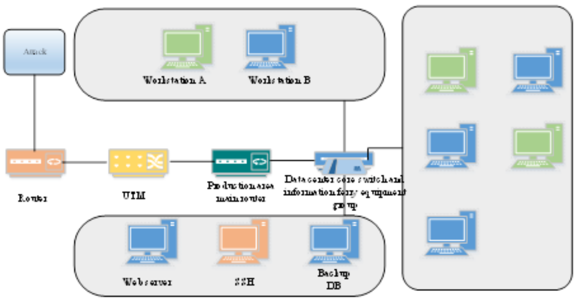
Attack
Workstation A
Workstation B
Router
UTM
Production area main router
Data center core switch and information ferry equipment group
Web server
SSH
Backup DB

Figure 2. Topology structure of production information network

To obtain the minimal logistic permeability map, the logistic permeability map was simplified using the greedy approach. To ease the deduction process, the minimum logic penetration diagram (Figure 3) was constructed using the seven most representative atomic penetrations involved.

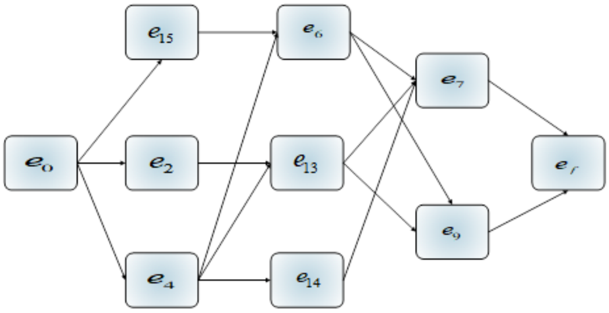

Figure 3. Logical Penetration Relationship Diagram

Let the empirical penetration indices be 1.25, 1.12, 1.05, and 1.22 for Web control vulnerabilities, Windows host rights, Windows root privileges, and Oracle database vulnerabilities, respectively [17].

The formula for the conversion probability matrix r is as follows. When multiple antecedent nodes exist, multiple sets of dependencies between antecedent nodes are obtained. Table 1 lists the three types of dependencies: overall, indirect, and direct. The penetration cost is calculated using penetration depth and control authority as examples.

Finally, a transformation matrix containing the penetration cost was included into the HMM. The starting state was established by combining historical data with network operation $\mathrm{I} = \{0.2, 0.1, 0.15, 0.05, 0.1, 0.1, 0.1, 0.2\}$. To obtain the observation matrix, each node's probability value was configured and modified in accordance with the penetration depth Pe:

Project	First jump	Second jump	Third jump	Fourth jump
Node	e_{14}	e_6	e_7	e_f
Dependency relationship	1,1,2	1.2,2,3,2.5	1.56,3.89,5.66	2.19,6.18,8.32
Project	e_{14}	e_6	e_7	e_f
Node	1,1,2	1.3,2.2,3.5	1.36,3.89,5.26	2.18,6.18,8.22
Dependency relationship	3.425	4.326	6.324	11.982

Table 1. Conversion cost calculation for infiltration nodes under various infiltration pathways

Penetration depth	No permission α_0	Logon rights α_1	Read permissions α_2	Write permission α_3	Control permissions α_4
Penetration cost	6.864	7.885	8.657	10.725	13.114

Table 2. Infiltration costs at different infiltration depths

Weakness	Seriousness	Key Weaknesses	Path	Penetration Cost	Weight
CVE-2004-0786	High Risk	✓	✓	13.021	8
CVE-2007-3902	High Risk		✓	14.252	8
CVE-2010-0244	High Risk	✓	✓	8.236	8
CVE-2011-0660	High Risk	✓		10.325	8
CVE-2007-2219	High Risk	✓	✓	5.241	8
CVE-M05-2491	Medium Risk		✓	8.352	5
CVE-2004-0747	Risk			1.145	4
CVE-2003-05412	Risk			Impermeable	2

Table 3. Comparison of Green Alliance penetration testing and this paper's methodology

$$K = \begin{bmatrix} 0.42 & 0.5 & 0.04 & 0.02 \\ 0.01 & 0.21 & 0.55 & 0.23 \\ 0.03 & 0.31 & 0.45 & 0.21 \\ 0.22 & 0.27 & 0.1 & 0.5 \\ 0.35 & 0.22 & 0.27 & 0.16 \\ 0.23 & 0.35 & 0.05 & 0.37 \\ 0.34 & 0.18 & 0.35 & 0.13 \\ 0.15 & 0.33 & 0.25 & 0.27 \end{bmatrix}.$$

This was introduced for computation in the HMM model. Considering the infiltration paths $\{1, 4, 5, 6\}$, the ninth path, $\{e_0, e_4, e_{14}, e_7, e_f\}$, was identified as the most likely one. Based on this result, nodes e_4, e_{14}, e_7 were designated as weak points and prioritized for protection. Table 2 displays the penetration costs at various depths.

It was observed that if the penetration path has fewer than three hops, it cannot reach the final target node and, therefore, cannot pose a threat to the final server. Penetrations involving more than three hops must be carefully inspected to determine if the penetration depth boundaries are crossed and represent a threat. Subsequently, the methodology employed in this article was compared with the evaluation outcomes of the Green Alliance's remote security assessment system using selected weak point penetration test results.

Table 3 provides a summary of penetration test results, showing that the evaluation outcomes of this study's approach are more precise than those of the Green Alliance software in identifying attacker vulnerabilities.

Lastly, the smoothness and independence of the penetration process were analyzed using this study's approach. The data from 30 sets of maximum penetration paths were presented, as shown in Figure 4.

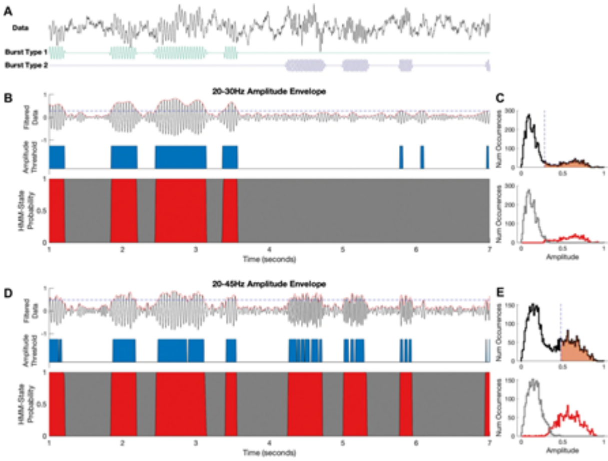

Figure 4. HMM test chart for stationarity and independence of stochastic processes

The analysis revealed that the sequence was reasonably smooth, with the autocorrelation coefficient oscillating around 0. The Q-statistic indicated that the F-value was higher at $K = 10, 11, 12$. Based on these evaluation results, incorporating node dependencies into the Hidden Markov Risk Assessment model allows for the dynamic reevaluation of node relationships, offering a more realistic depiction of vulnerabilities in the network.

5 Conclusion

Cascaded cumulative risk in penetration testing security risk assessment is caused by vulnerabilities that, if exploited, may raise the likelihood of attacks on additional hosts or services. The actual risk that varies with the network's dynamics can be more properly reflected by using the technique of estimating the re-dependencies between nodes. The actual risk outcomes brought about by vulnerabilities being exploited can be more accurately and thoroughly reflected by the WHMM-based penetration cost calculation approach for various penetration depths. Depending on the degree of danger, network administrators can reduce remediation costs by concentrating all of their attention on and safeguarding the vulnerabilities that represent the biggest risk to the system as a whole.

References

[1] Tuyen, N. D., Quan, N. S., Linh, V. B., Van Tuyen, V., & Fujita, G. (2022). A comprehensive review of cybersecurity in inverter-based smart power system amid the boom of renewable energy. *IEEE Access, 10*, 35846-35875.

[2] Yohanandhan, R. V., Elavarasan, R. M., Manoharan, P., & Mihet-Popa, L. (2020). Cyber-physical power system (CPPS): A review on modeling, simulation, and analysis with cyber security applications. *IEEE Access, 8*, 151019-151064.

[3] Wang, S., Hui, H., Ding, Y., Ye, C., & Zheng, M. (2022). Operational reliability evaluation of urban multi-energy systems with equivalent energy storage. *IEEE Transactions on Industry Applications, 59*(2), 2186-2201.

[4] Ahsan, F., Dana, N. H., Sarker, S. K., Li, L., Muyeen, S. M., Ali, M. F., ... & Das, P. (2023). Data-driven next-generation smart grid towards sustainable energy evolution: techniques and technology review. *Protection and Control of Modern Power Systems, 8*(3), 1-42.

[5] Shawly, T. (2023). A Detection and Response Architecture for Stealthy Attacks on Cyber-Physical Systems. *JOIV: International Journal on Informatics Visualization, 7*(3), 801-807.

[6] Malik, F. H., Khan, M. W., Rahman, T. U., Ehtisham, M., Faheem, M., Haider, Z. M., & Lehtonen, M. (2024). A Comprehensive Review on Voltage Stability in Wind-Integrated Power Systems. *Energies, 17*(3), 644.

[7] Matijašević, T., Antić, T., & Capuder, T. (2022). A systematic review of machine learning applications in the operation of smart distribution systems. *Energy Reports, 8*, 12379-12407.

[8] Tan, S., Guerrero, J. M., Xie, P., Han, R., & Vasquez, J. C. (2020). Brief survey on attack detection methods for cyber-physical systems. *IEEE Systems Journal, 14*(4), 5329-5339.

[9] Moudoud, H., Mlika, Z., Khoukhi, L., & Cherkaoui, S. (2022). Detection and prediction of fdi attacks in iot systems via hidden markov model. *IEEE Transactions on Network Science and Engineering, 9*(5), 2978-2990.

[10] Kyeremeh, F., Fang, Z., Yi, Y., & Peprah, F. (2023). Segmentation of a conventional medium voltage network into solar PV powered microgrid. *Energy Reports, 9*, 5183-5195.

[11] Diaba, S. Y., & Elmusrati, M. (2023). Proposed algorithm for smart grid DDoS detection based on deep learning. *Neural Networks, 159*, 175-184.

[12] AlHaddad, U., Basuhail, A., Khemakhem, M., Eassa, F. E., & Jambi, K. (2023). Ensemble model based on hybrid deep learning for intrusion detection in smart grid networks. *Sensors, 23*(17), 7464.

[13] Moustafa, N., Keshk, M., Choo, K. K. R., Lynar, T., Camtepe, S., & Whitty, M. (2021). DAD: A Distributed Anomaly Detection system using ensemble one-class statistical learning in edge networks. *Future Generation Computer Systems, 118*, 240-251.

[14] Muhati, E., & Rawat, D. B. (2021). Hidden-Markov-model-enabled prediction and visualization of cyber agility in IoT era. *IEEE Internet of Things Journal, 9*(12), 9117-9127.

[15] Thamilarasu, G., Odesile, A., & Hoang, A. (2020). An intrusion detection system for internet of medical things. *IEEE Access, 8*, 181560-181576.

[16] Abdul, D., & Wenqi, J. (2022). Evaluating appropriate communication technology for smart grid by using a comprehensive decision-making approach fuzzy TOPSIS. *Soft Computing, 26*(17), 8521-8536.

[17] Zhang, Z., Zhang, C., Li, M., & Xie, T. (2020). Target positioning based on particle centroid drift in large-scale WSNs. *IEEE Access, 8*, 127709-127719.