A Cyber-secure Framework for Power Grids Based on Federated Learning

Shutang You
University of Tennessee, Knoxville, TN, USA
Email: syou3@utk.edu

***Abstract*— Cyber security is important of power grids to ensure secure and reliable power supply. This paper presented a cyber-secure framework for power grids based on federated learning. In this framework, each entity, which may be a distribution/transmission/generation service provider or even a customer, can contribute to the overall system immunity and robustness to cyber-attacks, while not required to share local data, which may have privacy, legal and property concerns. The main idea is to use the federated learning framework to share the knowledge learned from local data instead of sharing power grid data itself. With complete knowledge learned from all data from the power grid, each entity is better positioned to defence the cyber-attacks and improve power grid resiliency. Future work on applying this federated learning based framework in power systems is also discussed.**

***Index Terms*— Federated learning, power grid cyber security, distributed machine learning, data authentication.**

I. Introduction

Power grid reliability and security is important for societies [1-10]. The increase of renewable generation, such as wind and solar, and other resources including energy storage and controllable load, are changing how power grids are being operated [3, 11-30]. The development of advanced measurement and control systems has significantly improved power system reliability, resiliency, and efficiency [31-53]. As more information is integrated into the cyber network of power systems, cyber security in power grids is becoming more important.

Existing cyber security technologies are mainly based on encryption or dedicate and physically-isolated communication networks [54]. However, these techniques are either expensive on investment or vulnerable in practice due to recent development of de-encryption computation technologies such as quantum computing.

Recent studies have found that leveraging large amounts of data and data-driven technologies can help improve the safety and performance of many cyber-physical systems [55]. The power industry is increasingly looking at how to use data-driven or machine learning technologies to defense cyber-attacks. However, due to the data privacy and proprietary consideration, one obvious hurdle of data-driven cyber-attack defense is the lack of data availability of the whole system [56]. Some progress has been made to improve the data sharing using some desensitization technologies, such as differential privacy [57], fog computing [58], and data aggregation[59] etc. These improvement has improved the data privacy issue to some extent.

Focusing on power grid cyber security, this paper studied the cyber-attack defense technology based on federated learning. A cyber-secure framework for power grids based on federated learning is proposed. Different parts of the framework and its working principle are described. Future work on using this framework in actual systems is also discussed.

II. The Privacy-preserving and Cyber-secure Framework for Power Grids Based on Federated Learning Framework

A. Background

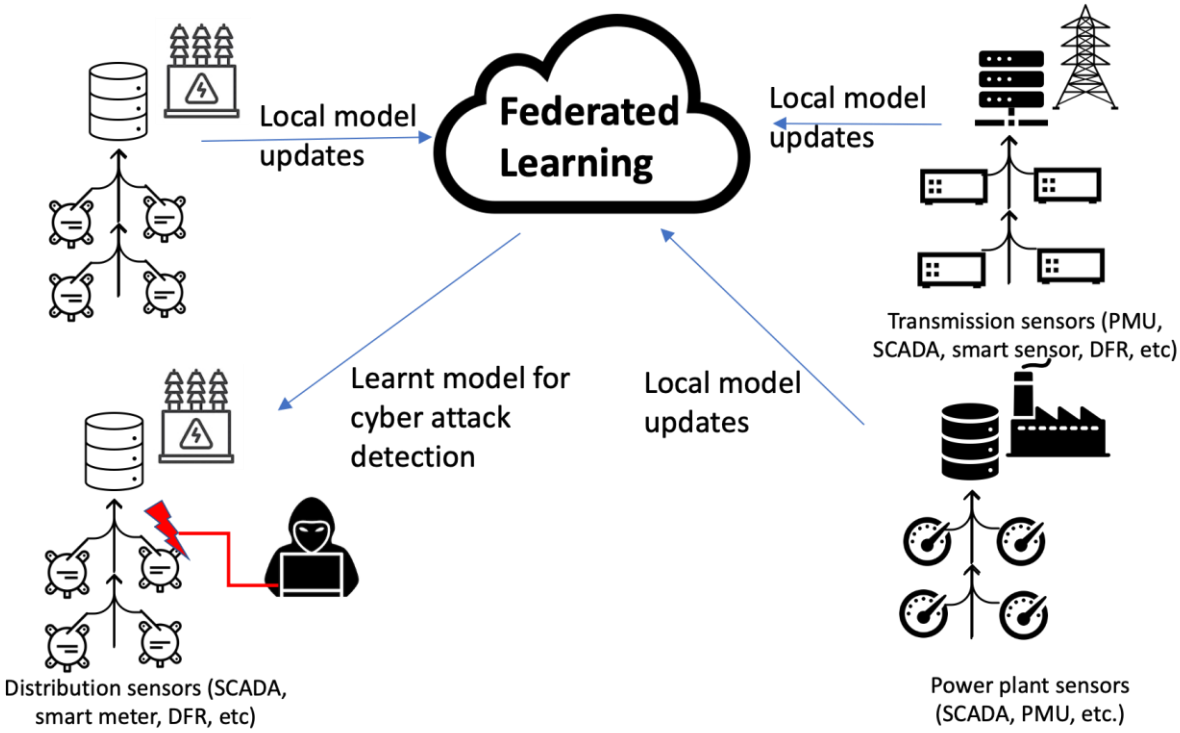
Power System Cyber Security: To provide reliable and economic power supply, power grids heavily rely on various IT systems, such as communication systems, distributed, and centralized computational and control systems. This dependency becomes a concern as they may easily become targets for malicious attacks on power grids. For example, Ukrainian experienced a power grid cyber-attack in 2015, which resulted in 73 MWh of electricity not supplied [60].

Data Property and Privacy in Power Systems: Power system data are usually collected by the generation, transmission, and distribution system owners or customers. These data are usually proprietary or include private information. Concerns on privacy, commercial interest and legal responsibilities create hurdles for data sharing between different entities. It is technically challenging and inefficient to collaborate without actually sharing data.

Federated Learning: Federated learning is a type of extension of the distributed machine learning technology. Compared with common distributed machine learning, federated learning has the advantage of taking advantage of wide-area multi-entity information while preserving data property and privacy. In addition, it has the following characteristics: the local data can be heterogeneous; the local computation node is more flexible in execution; and the communication burden is further reduced by using more local computation resources [61].

Motivation: The capability of each entity to defense cyber security is limited because of the limited observability of the whole system. A robust defense of cyber security increasingly

This work made use of Engineering Research Center shared facilities supported by the Engineering Research Center Program of the National Science Foundation and the Department of Energy under NSF Award Number EEC-1041877 and the CURENT Industry Partnership Program.



relies on data and knowledge sharing between different entities. In order to preserve the private and proprietary data locally while helping each entity better defense cyber-attacks, federated learning is leveraged to provide a secure cyber network for power grids. ***In this framework, each entity in power grids can help other entities better defense cyber-attacks in power grids without sharing proprietary data. Instead, they share security-relevant knowledge learned from data***.

B. Framework and Working Principles

Overall Framework: The framework of the secure network architecture is shown in Figure 1. Each entity communicate with the central server (or other entities in an alternative configuration without a centralized server) through a network that shares machine learning models and parameters. In this framework, each entity can choose to upload the latest machine learning model parameters that are trained based on local data. It can also request latest machine learning models that is updated based on all available segmental updates from each entity. The learnt model includes knowledge from other entities and therefore can be used to improve the cyber-attack detection capability. The role of each elements in this framework is further explained as follows.

Node: A node is an entity that manages its own data and has some computation resources. Typical nodes in this framework are ISOs, RCs, power plant control centers, and distribution system operation centers, Nodes can also be customer smart phones or smart home control centers that collect smart meter data from household devices. The computation and communication capabilities of nodes can vary in a wide range.

Each node (entity) has its freedom to update its own machine learning parameters based on local data, and then upload the parameters of the trained machine learning model. It can also manage its local resource independently based on its hardware and software configuration and usage needs. Since the trained machine learning model parameters do not contain sensitive data, they are much safer to share with other entities. Data features are removed during feature extraction and selection, and key information are masked by machine learning models.

Communication Network: The communication network that connects nodes can transmit the local model updates and the learnt model from time to time based on needs of each nodes. The communication requirement is much smaller compared with sharing real-time data itself throughout the network, which is becoming more challenging with the increasing number of sensors in power systems. Therefore, the communication can be based on lower-speed, narrower-band communication networks.

Local Model Update: Each node has its own data repository. For example, data collected in distribution system nodes include distribution SCADA, PMU, smart meter and DFR data, etc. Transmission nodes have PMU, SCADA, smart sensors, DFRs, weather, fire, and alarm data, etc. Power plant nodes have SCADA data of power generation units, associated control systems and PMU data installed at the system integration point. As they are from different types of nodes, these data are usually heterogeneous and have different distributions.

Coordination between Multiple Nodes: The machine learning parameter updates calculated by local nodes are sent to a central processor (or adjacent nodes in its decentralized version) to update iteratively the overall machine learning model for cyber attack defense.

Learnt Model for Cyber Attack Detection: After receiving the updated machine learning model that contains consensus knowledge of all available nodes in the federated learning

framework, each node can perform cyber attack detection independently. Since the machine learning model includes knowledge learnt from data in other nodes that are electrically connected together, the detection robustness and accuracy will be improved.

III. Discussion on Future Work

This paper proposed a framework for cyber security in power grids based on federated learning. The framework can eliminate the concern of data privacy while collaboratively defensing the cyber security of a large power system with many entities. Future work on this framework will include the following aspects:

1. Improve the degree of information masking of power grids data by developing appropriate feature selection techniques and using appropriate machine learning algorithms in the federated learning framework. This will further reduce the privacy and data property concern.
2. Improve the resiliency of the system by mitigating the propagation of the effects of data poisoning attacks from SCADA, PMUs, and smart meters, etc. when the system failed to detect a cyber-attack. Improve Byzantine-robustness in cyber-attack detection.
3. Bridge the heterogeneity gap between different types of data and create synergy in defending cyber-attacks.
4. Deal with data quality issues, such as bad data and missing data, and node availability and failure issues, such as loss of model updates.

References

1. Wang, R., et al. *A Novel Transmission Planning Method for Integrating Large-Scale Wind Power*. in *2012 Asia-Pacific Power and Energy Engineering Conference*. 2012. IEEE.
2. You, S., et al. *Ring-down oscillation mode identification using multivariate empirical mode decomposition*. in *2016 IEEE Power and Energy Society General Meeting (PESGM)*. 2016. IEEE.
3. Hadley, S.W. and S. You, *Influence Analysis of Wind Power Variation on Generation and Transmission Expansion in US Eastern Interconnection.*
4. Lv, C., et al., *Short‐term transmission maintenance scheduling based on the benders decomposition.* International Transactions on Electrical Energy Systems, 2015. **25**(4): p. 697-712.
5. Till, A., S. You, and Y. Liu, *Impact of High PV Penetration on Transient Stability—a Case Study on the US ERCOT System.*
6. Wang, J., et al., *Long-term maintenance scheduling of smart distribution system through a PSO-TS algorithm.* Journal of Applied Mathematics, 2014. **2014**.
7. Wang, J., et al., *Flexible transmission expansion planning for integrating wind power based on wind power distribution characteristics.* J. Electr. Eng. Technol, 2015. **10**: p. 709-718.
8. You, S., et al., *Oscillation mode identification based on wide-area ambient measurements using multivariate empirical mode decomposition.* Electric Power Systems Research, 2016. **134**: p. 158-166.
9. You, S., et al., *Impact of high PV penetration on the inter-area oscillations in the US eastern interconnection.* IEEE Access, 2017. **5**: p. 4361-4369.
10. Xiao, H., et al., *Data-Driven Security Assessment of Power Grids Based on Machine Learning Approach*. 2020, National Renewable Energy Lab.(NREL), Golden, CO (United States).
11. Guo, J., et al. *An ensemble solar power output forecasting model through statistical learning of historical weather dataset*. in *2016 IEEE Power and Energy Society General Meeting (PESGM)*. 2016. IEEE.
12. Hadley, S., et al., *Electric grid expansion planning with high levels of variable generation.* ORNL/TM-2015/515, Oak Ridge National Laboratory, 2015.
13. Sun, K., et al., *A Review of Clean Electricity Policies—From Countries to Utilities.* Sustainability, 2020. **12**(19): p. 7946.
14. You, S., *Electromechanical Dynamics of High Photovoltaic Power Grids.* 2017.
15. You, S., et al., *Co-optimizing generation and transmission expansion with wind power in large-scale power grids—Implementation in the US Eastern Interconnection.* Electric Power Systems Research, 2016. **133**: p. 209-218.
16. Tan, J., et al. *Developing High PV Penetration Case for Frequency Response study of U.S. Western Interconnection*. in *The 9th Annual IEEE Green Technologies Conference*. March 2017. Denver, Colorado: IEEE.
17. Tan, J., et al. *Frequency Response Study of US Western Interconnection under Extra-High Photovoltaic Generation Penetrations*. in *2018 IEEE Power & Energy Society General Meeting (PESGM)*. 2018. IEEE.
18. You, S., et al. *Energy Storage for Frequency Control in High Photovoltaic Power Grids*. in *IEEE EUROCON 2019-18th International Conference on Smart Technologies*. 2019. IEEE.
19. You, S., et al. *US Eastern Interconnection (EI) Electromechanical Wave Propagation and the Impact of High PV Penetration on Its Speed*. in *2018 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*. 2018. IEEE.
20. You, S., et al. *Impact of high PV penetration on US eastern interconnection frequency response*. in *2017 IEEE Power & Energy Society General Meeting*. 2017. IEEE.
21. You, S., et al. *Impact of High PV Penetration on U.S. Eastern Interconnection Frequency Response*. in *Power and Energy Society General Meeting (PESGM), 2017*. 2017. IEEE.
22. Yuan, H., et al. *Machine Learning-Based PV Reserve Determination Strategy for Frequency Control on the WECC System*. in *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. 2020. IEEE.
23. Yuan, Z., et al. *Frequency control capability of Vsc-Hvdc for large power systems*. in *2017 IEEE Power & Energy Society General Meeting*. 2017. IEEE.
24. Li, H., et al., *Analytic analysis for dynamic system frequency in power systems under uncertain variability*. IEEE Transactions on Power Systems, 2018. **34**(2): p. 982-993.
25. Liu, Y., S. You, and Y. Liu, *Study of wind and PV frequency control in US power grids—EI and TI case studies.* IEEE power and energy technology systems journal, 2017. **4**(3): p. 65-73.
26. Liu, Y., et al., *Frequency response assessment and enhancement of the US power grids toward extra-high photovoltaic generation penetrations—An industry perspective.* IEEE Transactions on Power Systems, 2018. **33**(3): p. 3438-3449.
27. Sun, K., et al., *Frequency secure control strategy for power grid with large-scale wind farms through HVDC links.* International Journal of Electrical Power & Energy Systems, 2020. **117**: p. 105706.
28. Yao, W., et al., *A fast load control system based on mobile distribution-level phasor measurement unit.* IEEE Transactions on Smart Grid, 2019. **11**(1): p. 895-904.
29. You, S., et al., *Non-invasive identification of inertia distribution change in high renewable systems using distribution level PMU.* IEEE Transactions on Power Systems, 2017. **33**(1): p. 1110-1112.
30. You, S., et al., *Comparative assessment of tactics to improve primary frequency response without curtailing solar output in high photovoltaic interconnection grids.* IEEE Transactions on Sustainable Energy, 2018. **10**(2): p. 718-728.
31. Liu, Y., et al. *Recent application examples of FNET/GridEye*. in *2018 IEEE 12th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG 2018)*. 2018. IEEE.
32. Wu, L., et al. *Statistical analysis of the FNET/grideye-detected inter-area oscillations in Eastern Interconnection (EI)*. in *2017 IEEE Power & Energy Society General Meeting*. 2017. IEEE.
33. You, S., et al. *A survey on next-generation power grid data architecture*. in *2015 IEEE Power & Energy Society General Meeting*. 2015. IEEE.
34. You, S., et al. *Data architecture for the next-generation power grid: Concept, framework, and use case*. in *2015 2nd International*

Conference on Information Science and Control Engineering. 2015. IEEE.
35. Zhang, X., et al. *Measurement-based power system dynamic model reductions*. in *2017 North American Power Symposium (NAPS)*. 2017. IEEE.
36. Alharbi, A., S. You, and Y. Liu. *Correlation between Generator Trips and Locational Marginal Prices (LMPs)*. in *2020 17th International Conference on the European Energy Market (EEM)*. 2020. IEEE.
37. Wu, L., et al. *Multiple Linear Regression Based Disturbance Magnitude Estimations for Bulk Power Systems*. in *2018 IEEE Power & Energy Society General Meeting (PESGM)*. 2018. IEEE.
38. You, S., et al. *FNET/GridEye for Future High Renewable Power Grids—Applications Overview*. in *2018 IEEE PES Transmission & Distribution Conference and Exhibition-Latin America (T&D-LA)*. 2018. IEEE.
39. Zhang, X., et al. *US eastern interconnection (EI) model reductions using a measurement-based approach*. in *2018 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*. 2018. IEEE.
40. Zhao, J., et al. *Data quality analysis and solutions for distribution-level PMUs*. in *2019 IEEE Power & Energy Society General Meeting (PESGM)*. 2019. IEEE.
41. Liu, Y., et al., *Recent developments of FNET/GridEye—A situational awareness tool for smart grid*. CSEE Journal of Power and Energy Systems, 2016. **2**(3): p. 19-27.
42. Tong, N., et al. *Dynamic Equivalence of Large-Scale Power Systems Based on Boundary Measurements*. in *2020 American Control Conference (ACC)*. 2020. IEEE.
43. Wang, W., et al. *Advanced synchrophasor-based application for potential distributed energy resources management: key technology, challenge and vision*. in *2020 IEEE/IAS Industrial and Commercial Power System Asia (I&CPS Asia)*. 2020. IEEE.
44. Li, J., et al., *A fast power grid frequency estimation approach using frequency-shift filtering*. IEEE Transactions on Power Systems, 2019. **34**(3): p. 2461-2464.
45. Liu, Y., et al., *A distribution level wide area monitoring system for the electric power grid–FNET/GridEye*. IEEE Access, 2017. **5**: p. 2329-2338.
46. Yao, W., et al., *Source location identification of distribution-level electric network frequency signals at multiple geographic scales*. IEEE Access, 2017. **5**: p. 11166-11175.
47. Yao, W., et al., *GPS signal loss in the wide area monitoring system: Prevalence, impact, and solution*. Electric Power Systems Research, 2017. **147**: p. 254-262.
48. Liu, S., et al., *Model-free Data Authentication for Cyber Security in Power Systems*. IEEE Transactions on Smart Grid, 2020.
49. Wang, W., et al., *Information and Communication Infrastructures in Modern Wide-Area Systems*. Wide Area Power Systems Stability, Protection, and Security: p. 71-104.
50. You, S., et al., *Disturbance location determination based on electromechanical wave propagation in FNET/GridEye: a distribution-level wide-area measurement system*. IET Generation, Transmission & Distribution, 2017. **11**(18): p. 4436-4443.
51. Cui, Y., S. You, and Y. Liu. *Ambient Synchrophasor Measurement Based System Inertia Estimation*. in *2020 IEEE Power Engineering Society General Meeting*. 2020. IEEE.
52. You, S., et al., *Power system disturbance location determination based on rate of change of frequency*. 2019, Google Patents.
53. You, S., et al., *Wide-area monitoring and anomaly analysis based on synchrophasor measurement*, in *New Technologies for Power System Operation and Analysis*. Elsevier. p. 143-161.
54. Sridhar, S., A. Hahn, and M. Govindarasu, *Cyber–physical system security for the electric power grid*. Proceedings of the IEEE, 2011. **100**(1): p. 210-224.
55. Jiang, Y., S. Yin, and O. Kaynak, *Data-driven monitoring and safety control of industrial cyber-physical systems: Basics and beyond*. IEEE Access, 2018. **6**: p. 47374-47384.
56. Guan, Z., et al., *Privacy-preserving and efficient aggregation based on blockchain for power grid communications in smart communities*. IEEE Communications Magazine, 2018. **56**(7): p. 82-88.
57. Fioretto, F., T.W. Mak, and P. Van Hentenryck, *Differential privacy for power grid obfuscation*. IEEE Transactions on Smart Grid, 2019. **11**(2): p. 1356-1366.
58. Zhu, L., et al., *Privacy-preserving authentication and data aggregation for fog-based smart grid*. IEEE Communications Magazine, 2019. **57**(6): p. 80-85.
59. Lu, R., et al., *EPPA: An efficient and privacy-preserving aggregation scheme for secure smart grid communications*. IEEE Transactions on Parallel and Distributed Systems, 2012. **23**(9): p. 1621-1631.
60. Case, D.U., *Analysis of the cyber attack on the Ukrainian power grid*. Electricity Information Sharing and Analysis Center (E-ISAC), 2016. **388**.
61. Konečný, J., et al., *Federated learning: Strategies for improving communication efficiency*. arXiv preprint arXiv:1610.05492, 2016.