PhD Thesis Proposal Form
China Scholarship Council (CSC)/ENS Rennes
Call for projects 2023

FIELD: Artificial intelligence in the field of electrical engineering and cybersecurity

THESIS SUBJECT TITLE

Secure artificial intelligence for the smart grid energy management

Name of French doctoral school: Matisse

Name of research team: SATIE (→ IETR in 2024) at ENS Rennes / IRISA (CIDRE team → PIRAT team in 2024)

Name of French Supervisor, associated supervisors and members of the supervising team:

- SATIE → IETR in 2024: Anne Blavette & Hamid Ben Ahmed (HDR) → electrical engineering and AI
- IRISA: Michel Hurfin[1](HDR), Yufei Han, Gilles Guette → cybersecurity and AI
- Orange : Raphaël Féraud (HDR) → AI

[1] Michel Hurfin holds an HDR (2004), is attached to the Matisse doctoral school and supervises less than 4 PhD students (2 theses in progress). He will be the French thesis director.

Name collaborator in Chinese university (if applicable):

- Name: Jialin LIU
- Position: Assistant Professor at Department of Computer Science and Engineering, Southern University of Science and Technology (SUSTech), Shenzhen, China
- E-mail: liujl@sustech.edu.cn

If previous collaborations with the Chinese co-director/university, please elaborate:

Type of PhD track : please tick the appropriate box:

- Joint PhD/cotutelle (leading to a double diploma) : NO
- Regular PhD (leading to a single French diploma) : Yes
- Research residency visit (for students enrolled at a Chinese institution who will be invited to a French institution to carry out a mobility period): NO

Research proposal abstract (1500 words max.):

Context and challenges

Large-scale fleets of electric vehicles (EVs) are expected to be deployed in power systems in the next few decades. However, optimizing the charging strategy of each vehicle, considering its individual constraints (e.g. mobility needs) and objectives (e.g. minimizing the charging cost) while satisfying grid constraints and maximizing global objectives (e.g. maximizing self-consumption from renewables) constitutes a tremendous challenge. Conventional methods, being centralized, are not sufficiently scalable to ensure the real-time control under uncertainty of millions of flexible entities. Hence, decentralized energy management methods are being developed, as they present a higher potential of scalability. However, a large-scale fleet may, if it is controlled by a malicious attacker, lead to a dramatic impact on the electric network, ranging from local poor power quality to a complete blackout [Reterink, 2021]. Hence, addressing the vulnerabilities of decentralized charging methods is a mandatory step to allow their future deployment in the electrical network, especially in the current context presenting an exploding number of cyber-threats.

Scientific issues addressed

The number of research works on the smart grid cyber-attacks have increased during the last decade, while studies focusing specifically on EVs started to emerge mostly at the end of this decade [Bhusal, 2021], [Acharya, 2020]. EVs present particular characteristics compared to non-EV controllable entities, therefore needing to consider cyber-attacks on them and on their charging infrastructure specifically. They may indeed charge at a greater power (especially when fast charging is considered) compared to usual household-loads, therefore creating a potentially more important impact on the network. Also, their point of connection is dynamic (i.e. EV may recharge/discharge at any point of the network) and it may present some uncertainty, thus increasing the complexity of detecting/mitigating a cyber-attack. Hence, the grid impact of cyber-attacks focusing on EVs needs to be studied specifically.

Some research works started addressing the issue of the EV cybersecurity. However, as most research work do not consider the physical response of the power system [Dey, 2020], [Chandwani, 2020], [Girdhar, 2022], the grid impact of such cyber-attacks has been little addressed [Bharathidasan, 2022]. It is therefore essential to assess this impact.

In particular, the proposed project proposes to focus on the grid impact of cyber-attacks on an artificial intelligence (AI)-based method that has been developed between SATIE, IRIT and Orange based on combinatorial multi-armed bandits and adaptive multi-agent systems [Zafar, 2023]. This method is deemed compatible with the targeted scale of millions of EVs, and has already been tested successfully in numerical simulations on a grid including more than 10 000 EVs. This method, being fully decentralized and based on a selfish approach, presents the additional inherent advantage to be immune to adversarial attacks. Such a type of attacks is characterized by the provision of false data by an agent to the others. As the agents communicate indirectly through their shared environment, our proposed method is immune by design to such attacks [Bonnefoi, 2017]. However, our proposed method remains vulnerable to attacks based on the falsification of environment signals (e.g. electricity prices, grid congestion, etc.). This constitutes a potential threat, whether it leads to the consideration of a false alert/ignorance of a true alert, to data/model poisoning [Wang, 2023], etc. For instance, a model could be corrupted by making it learn a pattern in the input data while preserving its normal behavior. Hence, an input data containing the attack pattern would trigger the malicious behavior of the model, e.g., categorizing the input as a specific class. Although such backdoor attacks have primarily been studied on image data, recent research shows that they can also be used on time series [Yu, 2022]. Therefore, a literature review will be performed at the beginning of the project to identify as exhaustively as possible potential cyber-attack types applicable to the proposed decentralized smart charging strategy.

Then, a grid impact assessment will be performed, followed by the design of optimal preventive means, as well as mitigation ones. While preventive means are intended to prevent any impact from occurring on the electrical network, mitigation means are intended to reduce this impact once a cyber-attack has been launched successfully. Preventive means may be of an algorithmic nature and/or of a physical nature. Preventive algorithmic means may consist for instance of a protection layer between the decision-making algorithm and the environment, for example in the form of a generative model of the environment. This model could be learnt by the agents in a decentralized manner and their decision based on such a model, rather than relying directly on the environment signal. This provides therefore a "filter" to cyber-attacks, but may lead to reduced performance under normal conditions. Other potential solutions will also be investigated. Mitigation means are algorithmic and consist in ensuring that EVs not targeted by a cyber-attack can function relatively satisfactorily under degraded grid conditions due to such an event, and that they can even help in reducing its grid impact. The cost of both preventive and mitigation means will be quantified and compared to the cost of physical preventing means in the form of grid reinforcement.

Main stages

1. Literature review on attacks targeting electric vehicles and AI-based algorithms
2. Identification of attacks types applicable to our proposed algorithm: data/backdoor poisoning, etc.
3. Identification (and if necessary, development) of metrics for assessing the impact
4. Development of realistic cyber-attacks scenarios
5. Grid impact assessment (performed iteratively with Steps 6 and 7)
6. Development of preventive and mitigation countermeasures (algorithmic and physical).

Methodology

A bi-level attacker/defender (AD) approach will be used to define the optimal countermeasures to be adopted by the non-infected EVs and grid operators (defenders). A power system model will be developed. As power system simulations (called load flows and solved with the Newton-Raphson algorithm) are usually time-intensive, a faster power system model may be developed. Physics-informed neural network (PINNs) are considered for the development of such a model [Huang, 2023].

The development of a generative model of the environment will be done through deep learning. For learning faster such models [Sohn, 2015],[Xu, 2019],[Kotelnikov, 2023], a federated learning approach [McMahan, 2017] may be used. The federated learning approach may also be considered to correct the model subjected to model poisoning. This allows to correct a model by un-learning only the erroneous data, compared to re-training it from scratch which can be very costly [Xu, 2024].

References

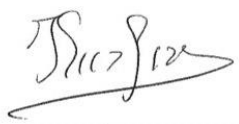
[Acharya, 2020]	S. Acharya, et al", "Cybersecurity of Smart Electric Vehicle Charging: A Power Grid Perspective", IEEE Access, 2020.
[Bharathidasan, 2022]	A review on electric vehicle: Technologies, energy trading, and cyber security, M. Bharathidasan et al., Energy Reports, Volume 8, 2022.
[Bhusal, 2021]	Narayan Bhusal et al., Cybersecurity of Electric Vehicle Smart Charging Management Systems, 52nd North American Power Symposium (NAPS), Tempe, AZ, USA, 2021.
[Bonnefoi, 2017]	R. Bonnefoi, L. Besson, C. Moy, E. Kaufmann, and J. Palicot, "Multi Armed Bandit Learning in IoT Networks: Learning helps even in non-stationary settings, CrownCom, 2017.
[Chandwani, 2020]	Ashwin Chandwani , "Cybersecurity of Onboard Charging Systems for Electric Vehicles—Review, Challenges and Countermeasures", IEEE Access (Volume: 8), 2020.
[Dey, 2020]	Satadru Dey et al., "Cybersecurity of Plug-In Electric Vehicles: Cyberattack Detection During Charging", IEEE Transactions on Industrial Electronics, Volume: 68, Issue: 1, January 2021.
[Huang, 2023]	B. Huang, J. Wang, "Applications of Physics-Informed Neural Networks in Power Systems - A Review", IEEE Transactions on Power Systems, 2023
[Girdhar, 2022]	M. Girdhar et al., "Machine Learning-Enabled Cyber Attack Prediction and Mitigation for EV Charging Stations", IEEE Power & Energy Society General Meeting (PESGM), Denver, CO, USA, 2022
[Kotelnikov, 2023]	Akim Kotelnikov, Dmitry Baranchuk, Ivan Rubachev, Artem Babenko , "TabDDPM: Modelling Tabular Data with Diffusion Models", in Proc. Proc. of the 40th International Conference on Machine Learning (ICML), Honolulu, Hawaii, USA, 2023.
[McMahan, 2017]	H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, Blaise Agüera y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data", Proc. 20th Int. Conf. on Artificial Intelligence and Statistics (AISTATS), Fort Lauderdale, FL, USA, 2017.
[Reterink, 2021]	Hans Reterink et al. "Impact of cybersecurity risks on the Dutch national charge point infrastructure", Technical report, 2021.
[Sohn, 2015]	Kihyuk Sohn, Honglak Lee, Xinchen Yan, "Learning Structured Output Representation using Deep Conditional Generative Models", Neural Information Processing Systems (NeurIPS), Montreal, Canada, 2015.
[Xu, 2019]	Lei Xu, Maria Skoularidou, Alfredo Cuesta-Infante, Kalyan Veeramachaneni, "Modeling Tabular Data using Conditional GAN", Neural Information Processing Systems (NeurIPS), Vancouver, Canada 2019.
[Xu, 2024]	Heng Xu, Tianqing Zhu, Lefeng Zhang, Wanlei Zhou, et Philip S. Yu, "Machine Unlearning: A Survey", ACM Computing Surveys 56, n° 1, 2024.
[Yu, 2022]	Yinbo Yu, Jiajia Liu, Shouqing Li, Kepu Huang, Xudong Feng, "A Temporal-Pattern Backdoor Attack to Deep Reinforcement Learning", In Proc. IEEE Global Communications Conference (GLOBECOM), Rio de Janeiro,Brazil, Dec. 2022.
[Wang, 2023]	Weizheng Wang et al. « Secure-Enhanced Federated Learning for AI-Empowered Electric Vehicle Energy Prediction », IEEE Consumer Electronics magazine, 2023
[Zafar, 2023]	S.Zafar, R. Féraud, A. Blavette, G. Camilleri, H. Ben Ahmed, "Decentralized Smart Charging of Large-Scale EVs using Adaptive Multi-Agent Multi-Armed Bandits", In Proc. Int. Conf. on Electricity Distribution (CIRED), Rome, Italy, 2023.

Publications of the laboratory in the field (max 5):

[Zafar, 2023]	S. Zafar, R. Féraud, A. Blavette, Guy Camilleri, H. Ben Ahmed, "Decentralized Smart Charging of Large-Scale EVs using Adaptive Multi-Agent Multi-Armed Bandits", In Proc. International Conference on Electricity Distribution (CIRED), Rome, Italy, 2023.
[Naseri, 2024]	Mohammad Naseri, Yufei Han and Emiliano De Cristofaro, "BadVFL: Backdoor Attacks in Vertical Federated Learning", to be presented at the 45th IEEE Symposium on Security and Privacy (S&P Oakland), 2024.
[Xu, 2023]	Xuran Xu, Yufei Han, and WangWei, "CGIR: Conditional generative instance reconstruction attacks against federated learning", in IEEE Transactions on Dependable and Secure Computing, vol. 20, no. 6, pp. 4551-4563, Nov.-Dec. 2023.
[Bao, 2023]	Bao Hongyan, Yufei Han, Yujun Zhou, Xin Gao, and Xiangliang Zhang, "Towards efficient and domain agnostic evasion attack with high-dimensional categorical inputs", In Proceedings of the 37th AAAI Conference on Artificial Intelligence, AAAI '23. Association for the Advancement of Artificial Intelligence (AAAI), 2023.
[Lyu, 2023]	Xiaoting Lyu, Yufei Han, Wei Wang, and Xiangliang Zhang, "Poisoning with cerberus: Stealthy and colluded backdoor attack against federated learning", In Proceedings of the 37th AAAI Conference on Artificial Intelligence, AAAI '23. Association for the Advancement of Artificial Intelligence (AAAI), 2023.

Date: 17/11/2023

Signature of the French Supervisor:



Michel HURFIN