A Development and Experimental Environment for Privacy-preserving and Secure (DEEPSECURE) Machine Learning

THE UNIVERSITY OF ARIZONA

Hongyi “Michael” Wu, PI, University of Arizona
Chunming Qiao and Hongxin Hu, PIs, University at Buffalo

University at Buffalo

Background and Motivation

While Machine Learning (ML) is embraced as important tools for efficiency and productivity, it is becoming an increasingly attractive target for cybercriminals. It is encouraging to witness the recent development in privacy-preserving and secure ML, which is drawing expertise from both ML and security/privacy and has been making promising progress in multiple fronts to tackle the multi-faceted problem. However, this emerging research community is facing a few fundamental challenges due to its interdisciplinary nature:

1) On the one hand, opensource deep learning frameworks such as Pytorch and Tensorflow have been made widely available, attributing significantly to the rapid advance of this flourishing field. But a critical hurdle faced by ML researchers is the steep learning curve to effectively use security techniques and libraries such as the Homomorphic Encryption (HE), Garbled Circuit (GC), Oblivious Transfer (OT), Secret Share (SS) and Differential Privacy (DP) to just name a few, to tackle security and privacy problems in ML.
2) On the other hand, while the security and privacy community has developed highly efficient techniques and libraries, it remains nontrivial to integrate them into deep learning models to achieve a computation efficiency suited for practical applications.

Overarching Goals

The overarching goal of the proposed project is to close the gap by establishing a Development and Experimental Environment for Privacy-preserving and Secure (DEEPSECURE) machine learning research. It will integrate a wide spectrum of essential functions and building blocks that are ready-to-use to shorten the learning curve for researchers coming from both ML and security/privacy communities, and at the same time fully customizable and scalable, enabling deep and fundamental research toward privacy-preserving and secure deep learning. Under this overarching goal, the specific project objectives include:

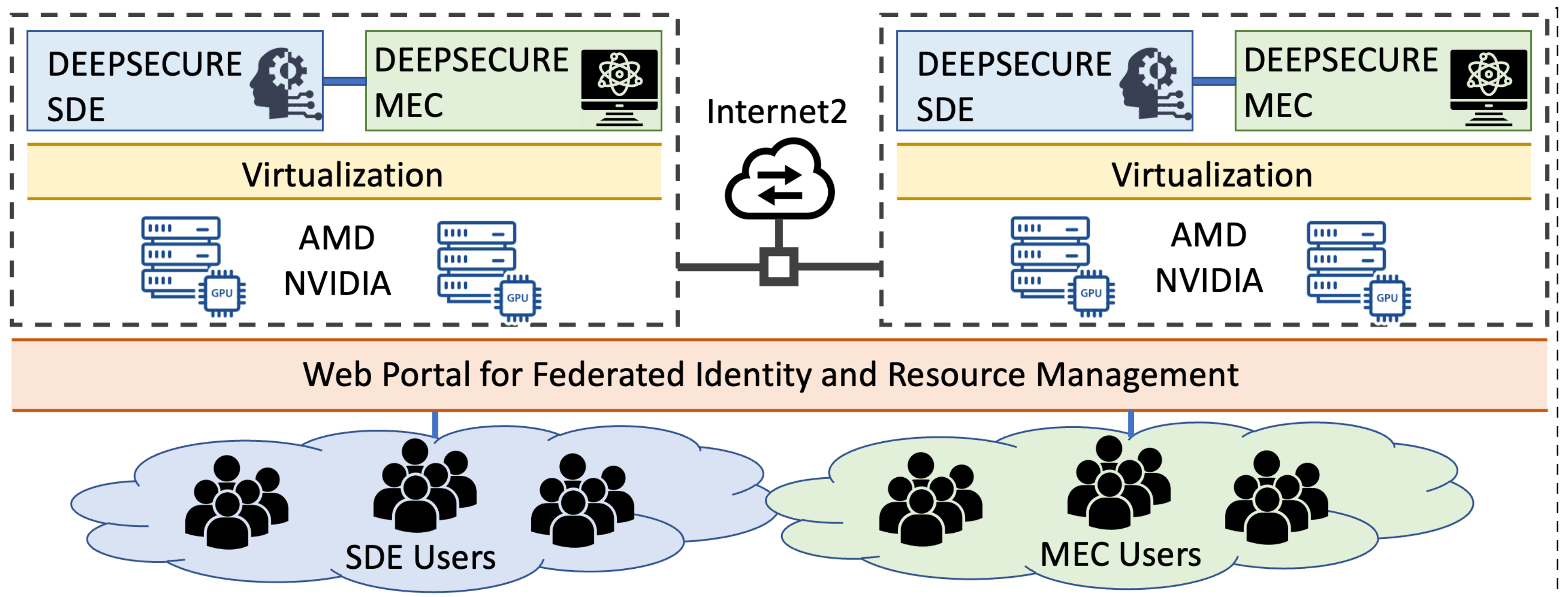
1) acquiring a scalable and re-configurable computing environment based on the latest Dell, AMD and Nvidia GPU technologies to establish the DEEPSECURE hardware infrastructure across the campuses of University of Arizona (UA) and University of Buffalo (UB);
2) developing a new software platform to support DEEPSECURE SDE (Software Development Environment) and MEC (Multi-user Experimental Chamber), integrated with PyTorch to enable great usability for beginners & advanced researchers and feature a scalable and customizable modular framework with seamlessly integrated libraries, function blocks and sample modules;
3) promoting DEEPSECURE across the nation to ensure broad participation and collaboration, serving as a hub for faculty, students and industry collaborators to learn the latest research advancement, share ideas, seek solutions, and solidify their research via testbed experiments;
4) leveraging DEEPSECURE to foster a long-lasting, self-sustainable ML security and privacy research community that engages all stakeholders in a sustained and ongoing way;
5) Educating/training diverse cybersecurity workforce to safeguard future intelligent cyber systems.

Broader Impacts

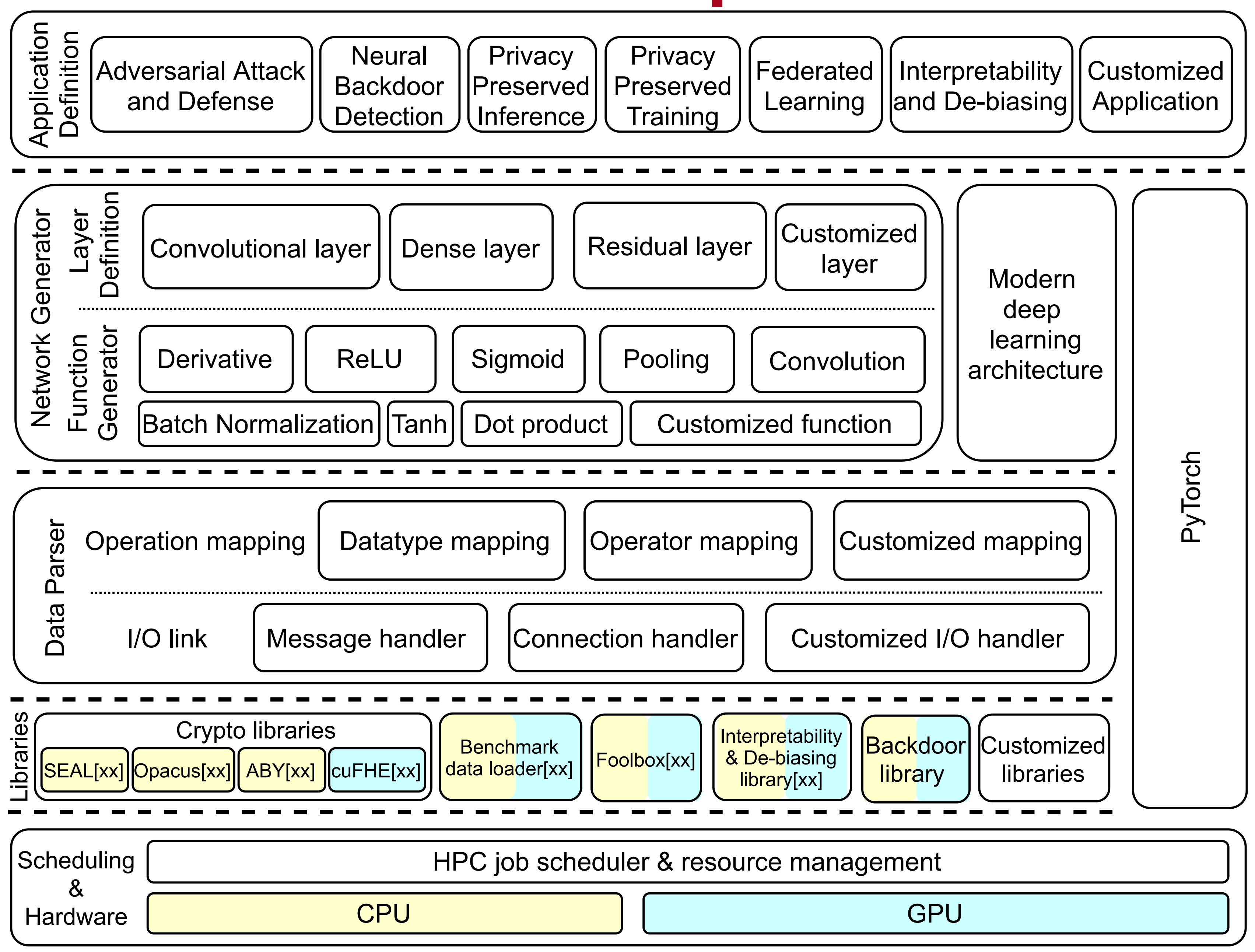
1) *Community building*: The project has received strong community support, as evidenced by the letters of support from 5 key stakeholders and letters of collaboration from 17 institutions across the country. The project includes significant efforts for fostering and sustaining an ML security and privacy research community, including regular virtual meetings (open forums) to provide regular updates to and seek feedback from the community, advisory board meetings, annual symposiums, and a training workshop series.
2) *Diversity and Inclusion*: A Diversity & Inclusion Committee will be formed to develop specific measures and plans for inspiring the participation of underrepresented groups and infusing diversity and inclusion in all DEEPSECURE events and activities such as the open forums, newsletters, board meetings, training workshops, symposiums, student recruiting, and research and development efforts. DEEPSECURE will be leveraged as an open-source and easy-to-use learning platform for curriculum development and workforce training.
3) *Outreach*: In addition to offering training opportunities to university researchers and industry practitioners, the project will participate in the existing NSF REU site program, NSA/NSF GenCyber summer camps, and a Cyber Saturday series to introduce career paths and educational resources to K-12 counselors, teachers, students, and parents.

*** Dr. Chunsheng Xin, Old Dominion University, was a Co-PI before he joined NSF as a PD.**

Hardware and Network Infrastructure



Software Dev and Exp Environment



Tools, Resources, and Data Sets

1) DEEPSECURE employs a spectrum of tools to support great portability, expandability, and shareability. Lightweight, portable, and self-sufficient Docker containers are used for development, enabling relocatability to new hardware and downloadability to local machines.
2) Data sharing is critical to the DEEPSECURE community. Such data include generic and specialized training and testing datasets as well as pretrained models (e.g., benign and malicious models for attack & defense experiments). Integrated data management is enabled by Pachyderm.
3) DEEPSECURE is modular and expandable. Its cloud infrastructure is built on the scalable, open-source OpenStack platform. Compute nodes can be added as users' needs grow. The federation of the DEEPSECURE's OpenStack clouds will be enabled using the Aristotle Cloud Federation.

Sample Projects

1) *Privacy Preserving in Machine Learning as a Service (MLaaS)*. While MLaaS is emerging as a promising cloud service, the interaction between clients and cloud servers leads to new privacy concerns. DEEPSECURE can effectively support the research in privacy-preserving MLaaS by leveraging its built-in libraries and Network Generator's function/layer modules.
2) *Adversarial Attacks and Defence on Multimodal Learning*. Leveraging the adversarial attack and defense libraries provided by DEEPSECURE, this research will investigate multimodal adversarial attacks, which perturb multiple modalities together, and corresponding defense.
3) *Hidden Neural Backdoor Detection and Mitigation*. Neural backdoor is becoming a real threat to ML. This research includes a series of investigations ranging from gaining deep understand of neural backdoor to its detection, eradication, and prevention.
4) *Interpreting Learning-based Network Intrusion Detection Systems*. Learning-based Network Intrusion Detection Systems (NIDSes) have been explored with great success. Using the interpretability library in DEEPSECURE, this research proposes xNIDS, a new framework that facilitates active intrusion response with explainable learning-based NIDSes.
5) *Attack and Defense in Federated Learning*. DEEPSECURE supports this emerging research by offering a development and experimentation environment that integrates crypto libraries and orchestrates them through automated network generation and data transformation.
6) *Privacy-Preserved ML Model Training*. This research will design accurate and efficient privacy preserving ML training protocols by capitalizing on different training modules. They will leverage DEEPSECURE network generator to improve training performance and utilize the data parser to handle the communication among different parties.
7) *Privacy-preserving Neural Architecture Search*. This research aims to advance neural architecture search (NAS) on privacy-preserving models by jointly optimizing privacy preserving computation and model hyperparameters. It will capitalize on DEEPSECURE by considering crypto libraries, data parsing, function generators, and secure computation in the search space.
8) *Secure and Privacy Preserving Machine Learning in Healthcare*. DEEPSECURE will be used to build secure and privacy-preserving ML protocols to enable models to be trained on vast amount of private health data from partners such as Sentara Healthcare and EVMS and used for privacy-preserving inference to return diagnosis results on encrypted data.
9) *Hands-on Lab Development*: Adversarial Attacks and Defenses in ML Models. The DEEPSECURE system can be leveraged as an open-source and easy-to-use learning platform to support AI-security education and training.

References

1) Q. Zhang, T. Xiang, C. Xin, and H. Wu, "From Individual Computation to Allied Optimization: Remodeling Privacy-Preserving Neural Inference with Function Input Tuning", in *IEEE Symposium on Security and Privacy (**S&P**)*, 2024. (Acceptance ratio: 14.9%).
2) R. Ning, J. Li, C. Xin, C. Wang, X. Li, R. Gazda, J.-H. Cho, and H. Wu, "ScanFed: Scalable Behavior-based Backdoor Detection in Federated Learning", in *IEEE International Conference on Distributed Computing Systems (**ICDCS**)*, 2023. (Acceptance ratio: 18.9%).
3) Z. Li, M. Yang, Y. Liu, J. Wang, H. Hu, W. Yi, and X. Xu, “GAN You See Me? Enhanced Data Reconstruction Attacks against Split Inference”, in the 37th *Conference on Neural Information Processing Systems (**NeurIPS**)*, 2023.
4) N. Vishwamitra, K. Guo, H. Hu, Z. Zhao, L. Cheng, and F. Luo, “Understanding and Measuring Robustness of Vision and Language Multimodal Models”, in the *International Conference on Secure Knowledge Management (**SKM**),* 2023.
5) F. Wei, H. Li, Z. Zhao, and H. Hu. "xNIDS: Explaining Deep Learning-based Network Intrusion Detection Systems for Active Intrusion Responses", in the 32nd *USENIX Security Symposium (**USENIX Security**)*, 2023.
6) L. Zhu, R. Ning, J. Li, C. Xin, and H. Wu, "Most and Least Retrievable Images in Visual-Language Query Systems", in *European Conference on Computer Vision (**ECCV**),* 2022. (Acceptance ratio: 28%).
7) Y. Cai, Q. Zhang, R. Ning, C. Xin, and H. Wu, "HE-Friendly Structured Pruning for Efficient Privacy-Preserving Deep Learning", in *ACM ASIA Conf. on Computer and Communications Security (**AsiaCCS**),* 2022. (Acceptance ratio: 18.4%).
8) R. Ning, J. Li, C. Xin, H. Wu, and C. Wang, "Hibernated Backdoor: A Mutual Information Empowered Backdoor Attack to Deep Neural Networks", in *AAAI Conference on Artificial Intelligence (**AAAI**)*, 2022. Oral Presentation. (Acceptance ratio: 15% for conference and 4.7% for oral presentation).
9) R. Ning, C. Xin, and H. Wu, "TrojanFlow: A Neural Backdoor Attack to Deep Learning-based Network Traffic Classifiers", in *IEEE Int’l Conference on Computer Communications (**INFOCOM**)*, 2022. (Acceptance ratio: 20%).
10) Y. Zhang, Y. Zhu (equal contribution), Z. Liu, C. Miao, F. Hajiaghajani, L. Su, and C. Qiao. “Towards Backdoor Attacks against LiDAR Object Detection in Autonomous Driving.” In *ACM Conference on Embedded Networked Sensor Systems (**SenSys**)*, 2022.
11) L. Zhu, R. Ning, C. Xin, C. Wang, and H. Wu, "CLEAR: Clean-up Sample-Targeted Backdoor in Neural Network", in IEEE International Conference on Computer Vision (**ICCV**), 2021. (Acceptance ratio: 25:9%).
12) Q. Zhang, C. Xin, and H. Wu, "GALA: Greedy ComputAtion for Linear Algebra in Privacy-Preserved Neural Networks", in *Network and Distributed System Security Symposium (**NDSS**),* 2021. (Acceptance ratio: 15:3%).
13) R. Ning, J. Li, C. Xin, and H. Wu, "Invisible Poison: A Blackbox Clean Label Backdoor Attack to Deep Neural Networks", in *IEEE Int’l Conference on Computer Communications (**INFOCOM**)*, 2021. (Acceptance ratio: 252/1266=19.9%).
14) Yi Zhu, Chenglin Miao, Foad Hajiaghajani, Mengdi Huai, Lu Su, and Chunming Qiao. “Adversarial Attacks against LiDAR Semantic Segmentation in Autonomous Driving.” In *ACM Conference on Embedded Networked Sensor Systems (**SenSys**)*, 2021.
15) Yi Zhu, Chenglin Miao, Tianhang Zheng, Foad Hajiaghajani, Lu Su, and Chunming Qiao. “Can We Use Arbitrary Objects to Attack LiDAR Perception in Autonomous Driving?” In *ACM Conference on Computer and Communications Security (**CCS**),* 2021.

National Science Foundation
WHERE DISCOVERIES BEGIN