Joshua C. Zhao

Purdue University zhao1207@purdue.edu joshuaczhao.github.io joshua-c-zhao

Research summary

My research focuses on machine learning with an emphasis on building scalable, reliable, and robust systems. As a machine learning security researcher, I am particularly passionate when learning about new systems and improving the safety (reliability of the models and the ethical uses) of real world deep learning applications.

I have worked in a variety of research areas. Some of my previous works explore data privacy (CVPR 23 & 24, S&P 24) or model training reliability under data heterogeneity (AsiaCCS 23, DSN-S 23), key challenges for real-world deployments of large-scale settings in federated learning. I have also worked in model robustness (reliability), studying the trade-off between robustness and generalization accuracy, along with developing more efficient approaches for robust transfer learning. Some of my recent work has also been using foundation models to improve weakly supervised point cloud semantic segmentation.

Computer skills: Deep Learning (PyTorch, Tensorflow, Keras), Machine Learning (Scikit-learn, XGBoost, etc.), Python (NumPy, Pandas, etc.), MATLAB, C/C++, LaTeX, Microsoft Office (Excel, Powerpoint)

Education

Purdue University, West Lafayette, IN — *Aug 2021 - May 2026 (Anticipated)*
Elmore Family School of Electrical and Computer Engineering
Ph.D. in Electrical and Computer Engineering

Northwestern University, Evanston, IL — *Sept 2020 – March 2021*
Robert R. McCormick School of Engineering and Applied Science
MS in Computer Science

Northwestern University, Evanston, IL — *Sept 2017 – March 2021*
Robert R. McCormick School of Engineering and Applied Science
BS in Computer Engineering

Publications

C = Conference, W = Workshop, J = Journal, S = Under Submission

[C.1] **Joshua C. Zhao**, Atul Sharma, Ahmed Roushdy Elkordy, Yahya H. Ezzeldin, Salman Avestimehr, Saurabh Bagchi, "*LOKI: Large-scale Data Reconstruction Attack against Federated Learning through Model Manipulation.*" IEEE Symposium on Security & Privacy, 2024 (pp. 1287-1305). (Acceptance rate: 261/1463 = 17.8%)

[C.2] **Joshua C. Zhao**, Ahaan Dabholkar, Atul Sharma, Saurabh Bagchi. "*Leak and Learn: An Attacker's Cookbook to Train Using Leaked Data from Federated Learning.*" IEEE/CVF CVPR, 2024 (pp. 12247-12256). (Acceptance rate: 2719/11532 = 23.6%)

[C.3] **Joshua C. Zhao**, Ahmed Roushdy Elkordy, Atul Sharma, Yahya H. Ezzeldin, Salman Avestimehr, Saurabh Bagchi, "*The Resource Problem of Using Linear Layer Leakage Attack in Federated Learning.*" IEEE/CVF CVPR, 2023 (pp. 3974-3983). (Acceptance rate: 2360/9155 = 25.8%)

[C.4] Atul Sharma, Wei Chen, **Joshua C. Zhao**, Qiang Qiu, Saurabh Bagchi, and Somali Chaterji. "*FLAIR: Defense against Model Poisoning Attack in Federated Learning.*" ACM AsiaCCS, 2023 (pp. 553-566). (Acceptance rate: 32/200 = 16.0%)

[W.1] Atul Sharma, **Joshua C. Zhao**, Wei Chen, Qiang Qiu, Saurabh Bagchi, and Somali Chaterji. "*How to Learn Collaboratively – Federated Learning to Peer-to-Peer Learning and What's at Stake.*" DSN Disrupt (DSN-S), 2023 (pp. 122-126). (Acceptance rate: 17/36 = 47.2%)

[J.1] **Joshua C. Zhao**, S. Bagchi, S. Avestimehr, K. Chan, S. Chaterji, D. Dimitriadis, J. Li, N. Li, A. Nourian, H. Roth, "*The Federation Strikes Back: A Survey of Federated Learning Privacy Attacks, Defenses, Applications, and Policy Landscape.*" ACM Computing Surveys (CSUR), 2025 (pp. 1-37).

[J.2] Yingyi Luo, **Joshua C. Zhao**, Arnav Aggarwal, Seda Ogrenci-Memik, and Kazutomo Yoshii. "*Thermal Management for FPGA Nodes in HPC Systems.*" ACM TODAES, 2020 (26, pp.1-17).

[S.1] **Joshua C. Zhao**, Ahaan Dabholkar, Saurabh Bagchi, "*Separate Classes, Separate Features? Separate Class Adversarial Training Reconciles Robust and Non-robust Features.*" (pp. 1-18).

[**S.2**] **Joshua C. Zhao**, Saurabh Bagchi, "*Are Fast Methods Stable in Adversarially Robust Transfer Learning?*" (pp. 1-13).

[**S.3**] Kasra Derakhshandeh, **Joshua C. Zhao**, Somali Chaterji, "*SEEMSeg: Calibration of SEEM for Weakly Supervised Point Cloud Semantic Segmentation.*" (pp. 1-16).

Research experience

Graduate Research *Purdue University*
Prof. Saurabh Bagchi's Lab, Dependable Computing Systems Laboratory (DCSL) *April 2021 - Present*

- Developed an attack against secure aggregation **federated learning** that arbitrarily scales towards FedAVG aggregation. The attack **leaks 85% of total user data** while prior work leaks less than 1% of user data. Furthermore, identified that FedAVG (a typically more challenging scenario) is more susceptible to attacks than FedSGD. Work accepted to **IEEE S&P 2024** as first author.
- Explored the reconstruction quality of various methods of data reconstruction attacks for downstream tasks. Models trained centrally on the leaked data **perform up to 20.4% better than federated learning** even with an extremely limited amount of data or low quality images. Accepted for publication at **CVPR 2024** as first author.
- Discovered the resource problems of privacy attacks on large scale federated learning with linear layer leakage. Using a sparse attack method we **decrease model size overhead by over 327x and computation time by 3.34x compared to SOTA** while maintaining an equivalent leakage rate. Work is accepted at **CVPR 2023** as first author.
- Evaluated the robustness of aggregation techniques in federated learning against malicious clients and developed a new aggregation design that provides **robustness with up to 45% malicious clients**. Work published in **AsiaCCS 2023**.

Undergraduate Research *Northwestern University*
Prof. Seda Memik's Lab, Dept. of Electrical Computer Engineering *June 2019 - March 2021*

- Investigated deep learning on chips with hardware-induced weight loss (coming from loss of power).
- Training DS-CNN networks for keyword spotting tasks and testing accuracy with different dropout rates being applied during the testing phase. Also testing model accuracy when random weight decay is applied on model parameters.
- Developed machine learning models using python to predict the temperature of FPGAs and applied them in task placement, decreasing the peak temperature in a system of FPGAs by up to 26.4°C.
- Performed feature selection, model selection, and parameter tuning to minimize the prediction error of a machine learning model to below 1.3 °C on average when predicting FPGA peak temperature.

Undergraduate Research *Northwestern University*
Prof. Oliver Cossairt's Lab, Dept. of Computer Science *Oct 2019 - March 2021*

- Developed a convolutional neural network in PyTorch to generate 3D face structure from single in- the-wild images. Training data generated through a 3D morphable head model.
- Designed and trained an encoder deep learning network to generate 3D face information using inputs of rendered images through PyTorch3D. In addition, applied the rendering process into the training as a custom regularization term in order to improve the actual rendering quality.
- Implemented a rendering process in PyTorch3D that takes in 2.5D (color and depth) information and renders multiple new images using light sources at different locations.
- Created an optimization framework to iteratively estimate surface normals of a scene using images under different lighting.
- Investigated image registration for 3D surface measurements of specular objects using feature-based alignment algorithms.

Other (awards, services, teaching)

- ***Guest lecturer at Purdue University*** for ECE 60872 (Fault-Tolerant Computer System Design). I gave the two following lectures on ML security:
  - How reliable is your model? (adversarial machine learning)
  - Distributed machine learning, a secure and private alternative?

- ***Reviewer:*** ICCV, CVPR FedVision, NeurIPS

- ***Head of NSF Center CHORUS Student Committee.*** This involves organizing virtual seminars/panels across the 4 universities and industry partners. Once a semester, there is also a newsletter summarizing the happenings in

the CHORUS center.

Awards:

- Bilsland Dissertation Fellowship
- Purdue Andrews Fellowship
- DCSL Best Fresher Award, Group Champion Award
- Eta Kappa Nu, Beta Tau Chapter (Electrical Engineering Honor Society)
- Tau Beta Pi (Engineering Honor Society)
- Northwestern University Summer Undergraduate Research Fellowship