Saeed Mahloujifar

Curriculum Vitae

Electrical and Computer Engineering
Princeton University

Homepage: smahloujifar.github.io
Email: sfar@princeton.edu

Education

Postdoctoral Research Associate (2020 - present)

- Princeton University, Princeton, NJ, USA
- Advisor: Prateek Mittal

Ph.D. (2015 - 2020)

- University of Virginia, Charlottesville, VA, USA
- Department of Computer Science
- Advisor: Mohammad Mahmoody

B.Sc. (2010-2015)

- Sharif University of Technology, Tehran, Iran
- Department of Computer Engineering
- Major: Software Engineering, Minor: Mathematics

Research Interests

- Foundations of Adversarial Machine Learning
- Foundations of Cryptography

◁ *My research statement is available here.*

Honors and Awards

- **John A Stankovic Research Award**, University of Virginia (2020).
- Top reviewer for **NeurIPS2021**, **ICLR2021**, **ICML 2020** and **NeurIPS 2019**
- Travel award to present at **ICML 2019** and **SODA 2020**.
- **Outstanding Research Graduate Student Award**, University of Virginia (2018).
- **Silver Medalist** in **Iranian National Olympiad in Mathematics** (2009).
- Member of **Iranian National Foundation of Elites** (2009-Present).

Publications

In the following * indicates equal contribution and [$\alpha\beta$] indicates alphabetical order.

☐ **Conference Publications**

- Saeed Mahloujifar, Esha Ghosh, Melissa Chase *Property Inference fro m Poisoning* IEEE Symposium on Security and Privacy (**S&P**), 2022.
- Chong Xiang, Saeed Mahloujifar, Prateek Mittal *PatchCleanser: Certifiably Robust Defense against Adversarial Patches for Any Image Classifier* **USENIX** Security Symposium 2022
- Xinyu Tang, Saeed Mahloujifar, Liwei Song, Virat Shejwalkar, Milad Nasr, Amir Houmansadr, Prateek Mittal *Mitigating Membership Inference Attacks by Self-Distillation Through a Novel Ensemble Architecture* **USENIX** Security Symposium 2022
  ◁ Preliminary version appeared in NeurIPS 2021 Workshop Privacy in Machine Learning

- Ashwinee Panda, Saeed Mahloujifar, Arjun N. Bhagoji, Supriyo Chakraborty, Prateek Mittal *SparseFed: Mitigating Model Poisoning Attacks in Federated Learning with Sparsification* International Conference on Artificial Intelligence and Statistics (**AISTATS**) 2022
- Vikash Sehwag, Saeed Mahloujifar, Tinashe Handina, Sihui Dai, Chong Xiang, Mung Chiang, Prateek Mittal *Improving Adversarial Robustness Using Proxy Distributions.* International Conference on Learning Representations (**ICLR**) 2022.

  ◁ Preliminary version appeared in ICLR 2021 Security and Safety in Machine Learning Systems Workshop
- [$\alpha\beta$] Samuel Deng, Sanjam Garg, Somesh Jha, Saeed Mahloujifar, Mohammad Mahmoody, and Abhradeep Thakurta. *A Separation result between data-oblivious and data-aware poisoning attacks* Conference on Neural Information Processing Systems (**NeurIPS**), 2021.

  ◁ A preliminary version presented at the Uncertainty and Robustness in Deep Learning workshop at ICML 2020.
- [$\alpha\beta$] Omid Etesami, Ji Gao, Saeed Mahloujifar, Mohammad Mahmoody
  Polynomial-time targeted attacks on coin tossing for any number of corruptions
  Theory of Cryptography Conference (**TCC**) 2021, 718-750
- Fnu Suya, Saeed Mahloujifar, Anshuman Suri, David Evans, and Yuan Tian. *Model-Targeted Poisoning Attacks with Provable Convergence.* International Conference on Machine Learning (**ICML**) 2021.
- [$\alpha\beta$] Nicholas Carlini, Samuel Deng, Sanjam Garg, Somesh Jha, Saeed Mahloujifar, Shuang ,Mohammad Mahmoody, Abhradeep Thakurta, Florian Tramer. *An Attack on Instahide: Is Private Learning Possible with Instance Encoding?.* IEEE Symposium on Security and Privacy (**S&P**), 2021.

  ◁ Also presented at NeurIPS Privacy Preserving Machine Learning Workshop, 2020. (Oral Presentation).
- Dimitrios I. Diochnos*, Saeed Mahloujifar*, Mohammad Mahmoody *Lower Bounds on Adversarially Robust PAC Learning.* International Conference on Machine Learning and Applications (**ICMLA**) 2020.

  ◁ *Also presented at* Security and Privacy of Machine Learning *workshop at ICML 2019 and* Robustness in Decision Making *workshop at NeurIPS 2019.*
- [$\alpha\beta$] Sanjam Garg, Somesh Jha, Saeed Mahloujifar, Mohammad Mahmoody *Adversarially Robust Learning Could Leverage Computational Hardness.* Algorithmic Learning Theory (**ALT**), 2020.

  ◁ *Additionally a preliminary version presented at* Security and Privacy of Machine Learning *workshop at ICML 2019 and Safety and* Robustness in Decision Making *workshop at NeurIPS 2019*
- [$\alpha\beta$] Omid Etesami, Saeed Mahloujifar, Mohammad Mahmoody *Computational Concentration of Measure: Optimal Bounds, Reductions, and More.* ACM-SIAM Symposium on Discrete Algorithms (**SODA**), 2020.
- Saeed Mahloujifar*, Xiao Zhang*, Mohammad Mahmoody, David Evans *Empirically Measuring Concentration: Fundamental Limits on Intrinsic Robustness.* Conference on Neural Information Processing Systems (**NeurIPS**), 2019 [Acceptance: 21%, (spotlight: 3%)].

  ◁ *Additionally, a preliminary version presented at* Safe Machine Learning *and* Debugging ML Models *workshops at ICLR 2019, as well as* Uncertainty and Robustness in Deep Learning *workshop at ICML 2019*
- Saeed Mahloujifar, Mohammad Mahmoody, Ameer Mohammad *Universal Multi-party Poisoning Attacks.* International Conference on Machine Learning (**ICML**) 2019. [Acceptance: 23%]

  ◁ *Additionally, selected for presentation at* ICLR 2019 Debugging Machine Learning Models *and* ICML 2019 Security and Privacy of Machine Learning *workshops.*

- Saeed Mahloujifar, Mohammad Mahmoody *Can Adversarially Robust Learning Leverage Computational Hardness?* Algorithmic Learning Theory (**ALT**), 2019.
- Saeed Mahloujifar, Dimitrios I. Diochnos, Mohammad Mahmoody *The Curse of Concentration in Robust Learning: Evasion and Poisoning Attacks from Concentration of Measure.* **AAAI** Conference on Artificial Intelligence , 2019 [Acceptance: 16%].

  ◁ *Additionally, presented at* NeurIPS 2018 Security in Machine Learning *workshop [Acceptance: 27%].*
- Dimitrios I. Diochnos*, Saeed Mahloujifar*, Mohammad Mahmoody *Adversarial Risk and Robustness: General Definitions and Implications for the Uniform Distribution.* Conference on Neural Information Processing Systems (**NeurIPS**), 2018 [Acceptance: 20%].
- Saeed Mahloujifar, Dimitrios I. Diochnos, Mohammad Mahmoody *Learning Under p-Tampering Attacks.* Algorithmic Learning Theory (**ALT**) pp. 572–596, 2018 [Acceptance: 34%].

  ◁ *Additionally, selected for presentation at* International Symposium on Artificial Intelligence and Mathematics (ISAIM) 2018.
- Saeed Mahloujifar, Mohammad Mahmoody *Blockwise p-tampering Attacks on Cryptographic Primitives, Extractors, and Learners.* Theory of Cryptography Conference (**TCC**) , Springer, Cham, pp. 245–279, 2017 [Acceptance: 34%].
- A. Rezaei, Saeed Mahloujifar, M. Soleymani *Near Linear-Time Community Detection in Networks with Hardly Detectable Community Structures.* ACM International Conference on Advances in Social Networks Analysis and Mining (**ASONAM**) 2015 [Acceptance: 18%].

□ Journal Publications

- Saeed Mahloujifar, Dimitrios I. Diochnos, Mohammad Mahmoody Learning under p-Tampering Poisoning Attacks. Annals of Mathematics and Artificial Intelligence.

□ Workshop papers and Preprints

- [$\alpha\beta$] Melissa Chase, Esha Ghosh, and Saeed Mahloujifar. *Property Inference from Poisoning.*
- Saeed Mahloujifar, Chong Xiang, Vikash Sehwag, Sihui Dai, Prateek Mittal *Robustness from Perception.*

  ◁ ICLR 2021 Security and Safety in Machine Learning Systems Workshop

Work Experience

- **Postdoctoral Research Associate at Princeton University** 2020-now
- **Research Intern at Microsoft Research Redmond** Summer 2020
- **Research Intern at Microsoft Research Redmond** Summer 2019
- **Research Assistant at University of Virginia** 2015-2020
- **Teaching Assistant at University of Virginia**
  - Program and Data Representation Fall 2015
  - Discrete Mathematics Fall 2015
  - Introduction to Cryptography Fall 2016
  - Algorithms Fall 2016
- **Teaching Assistant at Sharif University of Technology**
  - Compiler Design Fall 2014
  - Computer Networks Fall 2014
  - Introduction to Cryptography Fall 2014

Professional Service

- **Program Committee:** S&P 2023, CCS 2022, PETS 2022, ICML 2021, NeurIPS 2021, ICLR 2021, ICML 2020, NeurIPS 2020, ICLR 2020, AAAI 2021.
- **Journal Reviewer:** AMAI, JMLR, TBD, TDSCSI, Information and Computation
- **Conference Reviewer:** Crypto 2017, Eurocrypt 2018, Eurocrypt 2019, IJCAI 2019, Eurocrypt 2020, TCC 2020.