YUZHE MA

University of Wisconsin–Madison
ma234@wisc.edu · 6082133287

Education

University of Wisconsin–Madison *05/2021 (expected)*
Ph.D. in Computer Sciences, Minor in Statistics
Advisor: Professor Xiaojin (Jerry) Zhu

University of Wisconsin-Madison *05/2018*
M.S. in Computer Sciences

Huazhong University of Science and Technology *06/2016*
B.S. in Computer Science and Technology

Research Interests

My research interest lies in Adversarial Machine Learning (AML), with a focus on attacks and defenses in sequential decision making such as Multi-Armed Bandit (MAB) and Reinforcement Learning (RL). From attack perspective, I study how to adversarially manipulate the training samples such that the bandit player or RL agent learns a compromised policy. From defense perspective, I study how to design robust algorithms that are resistant to various types of attacks. Other topics I am interested in include machine teaching, personalized education, unsupervised learning (e.g., dimensionality reduction and clustering), and differential privacy.

Publication

Superscript * for alphabetic author order.

Yuzhe Ma, Jon A. Sharp, Ruizhe Wang, Earlence Fernandes, Xiaojin Zhu. Adversarial Attacks on Kalman Filter-based Forward Collision Warning Systems. *In Submission* to **AAAI**, 2021.

Xuezhou Zhang, Shubham Bharti, **Yuzhe Ma**, Adish Singla, Xiaojin Zhu. The Teaching Dimension of Q-learning. Preprint in arXiv:2006.09324. *In Submission* to **AAAI**, 2021.

Xuezhou Zhang, **Yuzhe Ma**, Adish Singla. Task-agnostic Exploration in Reinforcement Learning. In The 34th Conference on Neural Information Processing Systems (**NeurIPS**), 2020.

Xuezhou Zhang, **Yuzhe Ma**, Adish Singla, Xiaojin Zhu. Adaptive Reward-Poisoning Attacks against Reinforcement Learning. In The 37th International Conference on Machine Learning (**ICML**), 2020.

Yuzhe Ma, Xuezhou Zhang, Wen Sun, Xiaojin Zhu. Policy Poisoning in Batch Reinforcement Learning and Control. In The 33rd Conference on Neural Information Processing Systems (**NeurIPS**), 2019.

Yuzhe Ma, Xiaojin Zhu, Justin Hsu. Data Poisoning against Differentially-Private Learners: Attacks and Defenses. In The 28th International Joint Conference on Artificial Intelligence **(IJCAI**), 2019.

Yufei Han, **Yuzhe Ma**, Chris Gates, Kevin Roundy, Yun Shen. Collaborative and Privacy-Preserving Machine Teaching via Consensus Optimization. In International Joint Conference on Neural Networks (**IJCNN**), 2019.

Kwang-Sung Jun*, Lihong Li*, **Yuzhe Ma***, Xiaojin Zhu*. Adversarial Attacks on Stochastic Bandits. In The 32nd Conference on Neural Information Processing Systems (**NeurIPS**), 2018.

Yuzhe Ma, Kwang-Sung Jun, Lihong Li, Xiaojin Zhu. Data Poisoning Attacks in Contextual Bandits. In The 9th Conference on Decision and Game Theory for Security (**GameSec**), 2018.

Ayon Sen, Scott Alfeld, Xuezhou Zhang, Ara Vartanian, **Yuzhe Ma**, Xiaojin Zhu. Training Set Camouflage. In The 9th Conference on Decision and Game Theory for Security (**GameSec**), 2018.

Yuzhe Ma, Robert Nowak, Philippe Rigollet, Xuezhou Zhang, Xiaojin Zhu. Teacher Improves Learning by Selecting a Training Subset. In The 21st International Conference on Artificial Intelligence and Statistics (**AISTATS**), 2018.

Yuzhe Ma, Kun He, John Hopcroft, Pan Shi. Neighbourhood-Preserving Dimension Reduction via Localised Multidimensional Scaling. In Theoretical Computer Science (**TCS**), 2017.

Yuzhe Ma, Kun He, John Hopcroft, Pan Shi. Nonlinear Dimension Reduction by Local Multidimensional Scaling. In The 10th International Frontiers of Algorithmics Workshop (**FAW**), 2016.

Yuzhe Ma, Kun He, Leihua Qin, Yan Wang. A Primary Research on Overlapping Community Detection. In The 32nd National Conference on Theoretical Computer Science (**NCTCS**), 2014.

Yun-Shiuan Chuang, Xuezhou Zhang, **Yuzhe Ma**, Mark K. Ho, Joseph L. Austerweil, Xiaojin Zhu. Using Machine Teaching to Investigate Human Assumptions when Teaching Reinforcement Learners. Preprint in arXiv:2009.02476.

Work Experience

Research Intern, IBM Research *06-09/2020*
- Built a two-stage machine learning model to construct rasterized maps of American cities based on satellite imagery data. The model is able to identify buildings and roads on images with high accuracy.
- Applied cycle Generative Adversarial Network (cycle-GAN) to accomplish image-to-map translation.
- Applied YOLO to augment training data, and increased the density of detected buildings by 16 times.
- Evaluated the model performance on the satellite data of Las Vegas and detected 80% buildings of the city.

Applied Scientist Intern, Amazon *06-09/2019*
- Applied Gradient Boosted Tree (GBT) to identify students among Amazon customers.
- Carried out the whole pipeline of machine learning, including customer data acquisition with SQL scripts, data preprocessing, training with GBT, and post-processing of model prediction.
- Evaluated the model performance on Amazon Prime data, and achieved 85% accuracy.

Research Intern, Symantec Research Labs (NortonLifeLock) *05-08/2018*
- Developed a federated machine teaching framework that coordinates the training process of local nodes to jointly learn a desired model, while respecting the local data privacy during the communication between nodes.
- Applied block coordinate descent on the duality of a convex optimization, which requires each node to share only aggregated statistics of local data, thus can preserve privacy.
- Published a paper in International Joint Conference on Neural Networks (**IJCNN**).

Research Scholar, Cornell University *07-09/2015*
- Proposed a method to determine the intrinsic dimensionality of high-dimensional manifolds.
- Developed a nonlinear dimensionality reduction algorithm based on Multi-dimensional Scaling (MDS) that can maintain both local and global properties of the manifolds after reduction.

Honors and Awards

Student Travel Award, NeurIPS	*2019*
Top 50% Reviewer, NeurIPS	*2019*
Student Travel Award, GameSec	*2018*
Honorarium Award, GameSec Special Track	*2018*
Student Travel Award, AISTATS	*2018*
UW CS Summer Research Award	*2017*
Outstanding Bachelor Thesis Award	*2016*
CCF Outstanding Undergraduate Award	*2015*
Outstanding Student Leader Award	*2014*
China National Scholarship Award	*2013*

Academic Service

Program Committee: AAAI21, ACML20, AAAI20, ACML19, AAAI19
Conference Reviewer: AISTATS21, ICLR21, NeurIPS20, ICML20, AISTATS20, NeurIPS19, ICML19, AISTATS19
Journal Reviewer: TON, TPAMI, IEEE Access, Machine Learning
Student Volunteer: AISTATS18

Skills & Expertise

Programming Skills:	Python, Pytorch, SQL, Matlab, C, C++, R, AMPL, Verilog
Machine Learning:	Multi-Armed Bandit, Reinforcement Learning, Deep Learning, Differential Privacy