Yuetian Chen

yuetian.me
Stry233

(518)-244-0845
yuetian@purdue.edu
yuetian-purdue

Education

Purdue University — West Lafayette, Indiana
- *Ph.D. in Computer Science* — *July 2024 – Present*
  Advisor: Prof. Ninghui Li
  Research Focus: *AI Privacy, Large Language Model Membership Inference Attack*

Rensselaer Polytechnic Institute (RPI) — Troy, New York
- *Bachelor of Computer Science* — *July 2020 – December 2023*
  GPA: 3.85 / 4.0; Junior & Senior GPA: 3.93 / 4.0
  Relevant Coursework: *Computational Creativity, Machine Learning & Optimization, Rensselaer Center for Open Source*

Technical Skills

- **AI/ML Research:** PyTorch, Transformers, LLMs (GPT/LLaMA/Mistral), LoRA/QLoRA, Diffusion Models, RLHF, Multi-GPU Training, CUDA
- **Privacy & Security:** Membership Inference Attacks, Differential Privacy, Backdoor Attacks, Model Extraction
- **Programming:** Python, C++, CUDA, SQL, Bash, Git, LaTeX, Docker, Linux, SLURM/PBS
- **Research Tools:** Weights & Biases, TensorBoard, Jupyter, NumPy, Pandas, Opacus, Statistical Analysis, A/B Testing
- **Infrastructure:** AWS (EC2, S3, SageMaker), HPC Clusters, Distributed Training (DeepSpeed, FSDP), Ray, Kubernetes

Selected Publications

- **Membership Inference Attacks on Finetuned Diffusion Language Models**
  Y. Chen, K. Zhang, Y. Du, E. Stoppa, C. Fleming, A. Kundu, B. Ribeiro, N. Li
  International Conference on Learning Representations (ICLR) [🔗] **2026**
- **Window-based Membership Inference Attacks Against Fine-tuned Large Language Models**
  Y. Chen, Y. Du, K. Zhang, C. Fleming, A. Kundu, B. Ribeiro, N. Li
  USENIX Security Symposium [🔗] **2026**
- **Imitative Membership Inference Attack**
  Y. Du, **Y. Chen**, H. Xiao, B. Ribeiro, N. Li
  USENIX Security Symposium [🔗] **2026**
- **Cascading and Proxy Membership Inference Attack**
  Y. Du, J. Li, **Y. Chen**, K. Zhang, Z. Yuan, H. Xiao, B. Ribeiro, N. Li
  Network and Distributed System Security Symposium (NDSS) [🔗] **2026**
- **SOFT: Selective Data Obfuscation for Protecting LLM Fine-tuning against MIA**
  K. Zhang, S. Cheng, H. Guo, **Y. Chen**, Z. Su, S. An, Y. Du, C. Fleming, A. Kundu, X. Zhang, N. Li
  USENIX Security Symposium [🔗] **2025**
- **Membership Inference Attacks as Privacy Tools: Reliability, Disparity and Ensemble**
  Z. Wang, C. Zhang, **Y. Chen**, N. Baracaldo, S. Kadhe, L. Yu
  ACM Conference on Computer and Communications Security (CCS) [🔗] **2025**
- **Evaluating the Dynamics of Membership Privacy in Deep Learning**
  Y. Chen, Z. Wang, N. Baracaldo, S. R. Kadhe, L. Yu
  arXiv preprint arXiv:2507.23291 [🔗] **2025**
- **Reflections & Resonance: Two-Agent Partnership for Advancing LLM-based Story Annotation**
  Y. Chen, M. Si
  Joint International Conference on Computational Linguistics (LREC-COLING) [🔗] **2024**
- **Enhancing Sentiment Analysis Results through Outlier Detection Optimization**
  Y. Chen, M. Si
  arXiv preprint arXiv:2311.16185 [🔗] **2023**
- **Prompt to GPT-3: Step-by-Step Thinking Instructions for Humor Generation**
  Y. Chen, B. Shi, M. Si
  International Conference on Computational Creativity (ICCC) [🔗] **2023**
- **Automated Visual Story Synthesis with Character Trait Control**
  Y. Chen, B. Shi, P. Liu, R. Li, M. Si
  Applied Human Factors and Ergonomics (AHFE) [🔗] **2023**
- **Visual Story Generation Based on Emotion and Keywords**
  Y. Chen, R. Li, B. Shi, P. Liu, M. Si
  AAAI Conf. on AI and Interactive Digital Entertainment (AIIDE) [🔗] **2023**
- **Automated Cell Recognition using Single-cell RNA Sequencing with Machine Learning**
  C. Xu, **Y. Chen**, Y. Cao
  International Conference on Computational Biology and Bioinformatics (ICCBB) [🔗] **2021**

Research Experience

- **Graduate Research Assistant – TruSe Lab** Purdue University
  Advisor: Prof. Ninghui Li, Department of Computer Science *Jul 2024 – Present*
  - Developed window-based membership inference attack using novel sliding-window analysis to achieve 30% higher AUC than existing baselines for detecting training data exposure in fine-tuned LLMs.
  - Co-architected SOFT framework implementing selective data obfuscation techniques to reduce privacy leakage by 60% while maintaining 95% model utility across benchmark tasks.
  - Authored and submitted 5 research papers to top-tier venues (ICLR'26, USENIX Security'26, NDSS'26, USENIX Security'25) to advance state-of-the-art in LLM privacy and security.

- **Undergraduate Research Assistant – DSP Lab** RPI
  Advisor: Prof. Lei Yu, Department of Computer Science *May 2023 – Aug 2024*
  - Pioneered evaluation framework for membership inference attacks through systematic analysis to identify critical flaw in data uniqueness assumption affecting 70% of existing defenses.
  - Co-developed MIAE toolkit with IBM Research implementing 8 attack algorithms and 3 evaluation metrics to enable production-level privacy auditing at IBM Watson.
  - Published 2 papers at CCS'25 documenting novel attack methodologies to establish new benchmarks for privacy evaluation in machine learning systems.

- **Undergraduate Research Assistant – ISL** RPI
  Advisor: Prof. Qiang Ji, Department of ECSE *May 2023 – Dec 2023*
  - Optimized object detection pipeline through novel backbone pruning techniques to reduce MS-COCO inference latency by 18% while maintaining 95% mAP accuracy.
  - Integrated real-time emotion and pose recognition models on Pepper robot platform to enable naturalistic human-robot interaction for 250-participant behavioral study.

- **Undergraduate Research Assistant – CISL** RPI
  Advisor: Prof. Mei Si, Department of Cognitive Science *Mar 2022 – Dec 2023*
  - Published 5 papers (LREC-COLING'24, ICCC'23, AHFE'23, AIIDE'23) investigating LLM applications in computational creativity to advance automated content generation research.
  - Engineered prompt-chaining methodology using iterative refinement techniques to reduce GPT-3.5 hallucination rates by 40% in narrative generation tasks.
  - Built multimodal story-generation system combining LLMs with Stable Diffusion to serve 500+ users with 4.2/5 satisfaction rating for interactive storytelling applications.

Teaching Experience

- **Head Teaching Assistant** *RPI*
  Department of Computer Science, Supervisor: Lecturer Konstantin Kuzmin
  - **CSCI 2500: Computer Organization (400+ students)** *Fall 2023*
    Led team of 29 TAs; reduced grading turnaround from 14 to 3 days via optimized workflow automation. Created 6 new review labs, improving average scores by 12%. Received a perfect 5/5 faculty evaluation.
  - **CSCI 2600: Principles of Software (350+ students)** *Spring/Summer 2023*
    Managed 17 TAs and maintained 98% help-desk response rate. Implemented a Git-based peer review system, resulting in 15% improvement in student satisfaction scores.

- **Teaching Assistant / Undergraduate Mentor** *RPI*
  Department of Computer Science & Cognitive Science
  - **CSCI 2500: Computer Organization** *Fall 2022*
    Mentored 30 students in MIPS assembly and digital logic, improving exam averages by 10%.
  - **CSCI 2600: Principles of Software** *Summer/Fall 2022*
    Taught Java design patterns and JUnit testing to 25 students; guided 5+ capstone teams to A grades.
  - **COGS 2140: Introduction to Logic** *Fall 2022*
    Facilitated weekly recitation sessions for 60 students; 90% achieved B or higher on final exam.

Honors and Awards

- Rensselaer Polytechnic Institute Dean's Honor List (6 semesters): Fall 2020 – Fall 2023
- Academic Recognition Letters: Charles V. Stewart (Spring 2022); Mohammed J. Zaki (Summer 2022)

Presentations & Talks

- **RHC Academic Showcase – Poster Presentation** *October 2023*
  "Understanding the Dynamics of Membership Privacy in Deep Learning" – Presented MIA framework and privacy evaluation methods to 200+ attendees at RPI research symposium.

- **Canada-China International Film Festival (CCIFF) – Invited Talk** *July 2023*
  "AI in Creative Arts" – Demonstrated LLM-based story generation pipeline for film applications at an international festival in Montreal.

- **RPI Undergraduate Research Fair – Best Poster Award Nominee** *April 2023*
  "Visual Story Generation with LLMs and Diffusion Models" – Showcased multimodal AI system combining GPT-3.5 and Stable Diffusion for interactive storytelling.