

李佳睿

(+1) 412-915-6531

[jiaruli@umich.com](mailto:jiaruli@umich.com)

[个人网站](#)

[英文简历](#)

## 教育经历

08 2024 - 05 2029

密歇根大学安娜堡分校 | 美国安娜堡 - 计算机科学与工程博士

课程: 高级操作系统, 高级编译器

09 2022 - 05 2024

卡内基梅隆大学 | 美国匹兹堡 - 人工智能工程硕士 - 信息安全方向

课程: 可信AI系统, 大规模互联网服务, 电信网络, 高级自然语言处理

论文: AI时代的网络安全挑战: 新的攻击与防御

08 2024 - 05 2029

香港中文大学(深圳) | 中国深圳 - 计算机科学与工程学士

课程: 分布式与并行计算, 云计算, 数据库系统, 数据结构与算法

## 研究兴趣

鲁棒且可信的人工智能。机器学习系统安全。

## 发表论文

J. Li, J. Brewington, Q. Zhang, Z.M. Mao. WIP: Hijacking Attacks on UAV Follow-Me Systems in Realistic Scenarios. The 3rd USENIX Symposium on Vehicle Security and Privacy (VehicleSec'25), 2025.

J. Li, Y. Yuan, Z. Zhang. Enhancing LLM Factual Accuracy with RAG to Counter Hallucinations: A Case Study on Domain-Specific Queries in Private Knowledge-Bases. arXiv, 2024.

P. Sharma, J. Li, and G. Joshi. On Improved Distributed Random Reshuffling Over Networks. In Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2023.

T. Kim, J. Li, N. Madaan, S. Singh, and C. Joe-Wong. Adversarial Robustness Unhardening via Backdoor Attacks in Federated Learning. In NeurIPS 2023 Workshop on Backdoors in Deep Learning, 2023.

## 研究经历

09 2024 - 至今

## 鲁棒AI驱动的无人机系统

导师: Morley Mao教授, Robustnet, 密歇根大学, 美国安娜堡

- 开发基于现代深度学习目标跟踪算法的无人机目标跟随系统(基于运动、外观和混合算法), 部署在DJI Matrice无人机和定制NUC计算单元上
- 在物理模拟器中演示系统在多种良性场景下的鲁棒性(风力、纹理环境、随机目标运动)以及在对抗性环境下的脆弱性(云台相机攻击)
- 发现复杂攻击者可以使跟随目标切换到选定对象, 攻击成功率超过80%, 并提出使用相机运动补偿算法提高攻击下的鲁棒性

09 2023 - 02 2024

## 检索增强生成大语言模型

导师: Graham Neubig教授, 卡内基梅隆大学, 美国匹兹堡

- 设计RAG大语言模型系统(在线数据集生成器、嵌入、向量数据库和大语言模型服务), 以减少在私有数据集上进行问答时的大语言模型幻觉
- 发现RAG大语言模型通过微调嵌入实现最佳事实准确性(0.437召回率, 0.304 F1分数), 以及在小规模数据集上微调核心大语言模型的局限性

09 2022 - 05 2024

## 安全联邦学习系统

导师: Carlee Joe-Wong教授, LIONS Group, 卡内基梅隆大学, 美国匹兹堡

- 通过中央服务器聚合算法改善联邦(分布式)学习系统在对抗性参与客户端场景下的安全性
- 设计并实现新型攻击, 将模型对规避攻击的鲁棒性降低89.73%
- 开发模块化攻击流水线, 代码复杂度降低57.47%

01 2023 - 08 2023

## 分布式优化

导师: Gauri Joshi教授, OPAL Lab, 卡内基梅隆大学, 美国匹兹堡

- 设计实验证明非凸收敛在理论上存在差距
- 从零开始实现网络图、优化器和梯度跟踪, 重现文献结果
- 改进凸和非凸情况下去中心化随机重洗算法的理论收敛结果

## 工作经历

---

08 2021 - 08 2022

### 多语言自然语言处理 - 工程实习生

AI实验室, 腾讯公司, 中国深圳

- 通过标记化和反向翻译预处理和增强大规模(1000万)非洲语言语料库
- 升级THUMT NLP工具包, 添加多语言神经网络、去噪自编码器功能, 同时保持单语言设置性能BLEU 26.85
- 通过获得最佳BLEU 14.09和ChrF++ 37.42在WMT22竞赛中获得第1名

11 2019 - 08 2021

### 农业物联网系统 - 主要工程师

AgriFuture (创业项目), 中国深圳

- 使用Arduino芯片、传感器、基站和树莓派服务器设计农场数据收集系统
- 构建管理网站, 使用Go后端、Vue.js前端和AWS QLDB数据库
- 将系统部署到深圳农场并在创业竞赛中获得铜牌

## 精选项目

---

02 2023

### Log结构键值存储性能评估

- 实现log结构内存管理、故障恢复、测试工作负载和基线naive-KV-store
- 得出log-KV-store在恢复吞吐量上优于48.28%, 在恢复错误率上优于77.48%的结论

02 2023

### Malloc和Free的系统内存管理

- 通过选择策略(包括first-two-fit和LIFO)进行块搜索和插入, 将吞吐量提高66.17%
- 通过减少外部和内部碎片, 在位级别将内存利用率优化24.11%

02 2021

### 大学体育协会应用

- 领导五人开发团队使用Vue和NoSQL数据库构建管理小程序
- 设计云API位置签到功能, 减少客户端计算负担并加速程序运行
- 实现超过100名用户在活动中同时进行手机访问、数据传输和GPS定位

## 教学经历

---

08 2023 - 05 2024

### 计算机系统导论 - 助教

卡内基梅隆大学, 美国匹兹堡

- 涵盖汇编、C编程和计算机体系结构等主题

06 2023 - 08 2023

### 人工智能数学基础 - 助教

卡内基梅隆大学, 美国匹兹堡

- 涵盖线性代数主题, 包括矩阵、线性变换、向量空间和特征值

01 2023 - 05 2025

### 对抗机器学习 - 助教

卡内基梅隆大学, 美国匹兹堡

- 涵盖监督垃圾邮件过滤、无监督网络异常检测 and 对抗深度学习等主题

## 技术栈

---

Python, C/C++, SQL, PyTorch, Git, Docker, Linux, Arduino, CUDA, Verilog

## 荣誉与奖项

---

WMT22大规模非洲语言机器翻译 - 约束赛道, 第1名 | 2022年9月

中国国际"互联网+"大学生创新创业大赛, 铜牌 (1%) | 2021年8月

香港中文大学(深圳)学勤书院奖学金 (5%) | 2020年9月