Jiyue **Huang**

TU Delft · Peking University

(+31) 625502430 | j.huang-4@tudleft.nl | GillHuang-Xtler

*"I am seeking for a (roughly) 3-month **research visit** opportunity in the area of **security and privacy** challenges with emphasis on **distributed learning** systems. My research background is mainly on attacks and defenses of heterogeneous Federated Learning."*

Education

Delft University of Technology (The Distributed Systems Group, EWI)

Delft, the Netherlands

Ph.D in distributed Machine Learning

Nov, 2020 - present

- First 3 Year **4 academic papers** accepted in federated learning.
- **3 cooperated paper** accepted.
- 1 journal paper under review.
- **Fully-funded** Ph.D position.

Peking University (School of Electronics and Computer Engineering)

Shenzhen, China

M.S. in Computer Application Technology

Sep, 2017 - June, 2020

- **GPA 88.8%,** ranking Top 10%. **IELTS 7.0.**
- **6 Academic papers**, 4 named as the first student author.
- **1 First Prize** of Shenzhen Graduate School on the 26th Challenge Cup Competition.
- 1 of 39 **Merit Student Award** recipients in **564** students for promising students in research and study.
- Advanced Individual Award for **Scientific Research** in Peking University.

City University of Hong Kong (Electronic Engineering Department)

Hong Kong

Exchange Student in Electronic Information Engineering

Aug, 2016 - Jan, 2017

- **Sole recipient** of annual exchange opportunity in Tianjin University to City University of Hong Kong.

Tianjin University (School of Electronic Information Engineering)

Tianjin, China

B.E. in Electronic Information Engineering

Sep, 2013 - July, 2017

- **GPA 87.9%**, **Mathematical Contest in Modeling**, Electronic Design Contest.
- Won the **2/55** place in National Undergraduate Innovation and Entrepreneurship Training Programs.
- **Merit Student** for **3** consecutive years,. **Zhonghuan Electronic Scholarship**.
- **Outstanding Graduates Award** for postgraduate recommendation without exam.

Publications

[1] **Huang J**, Zhao Z, et al. Fabricated Flips: Poisoning Federated Learning without Data. DSN 2023 (CORE rank: A).

[2] **Huang J**, Hong C, et al. Maverick Matters: Client Contribution and Selection in Federated Learning. PAKDD 2023 (CORE rank: A).

[3] **Huang J**, Hong C, et al. Tackling Mavericks in Federated Learning via Adaptive Client Selection Strategy. IEEE Transactions on Parallel and Distributed Systems (Under review).

[4] Zhao Z[1], **Huang J**[1], et al. Defending Against Free-Riders Attacks in Distributed Generative Adversarial Networks. FC 2023 (CORE rank: A).

[5] Hong C, **J Huang**, et al. Exploring and Exploiting Data-Free Model Stealing. ECML-PKDD 2023 (CORE rank: A).

[6] Xu J, Hong C, **Huang J**, et al. AGIC: Approximate Gradient Inversion Attack on Federated Learning. SRDS 2022 (CORE rank: B).

[7] **Huang J**, Talbi, R, et al. An Exploratory Analysis on Users' Contributions in Federated Learning, TPS 2020.

[8] **Huang J**, Du M, et al. Attention-based Updates Aggregation in Federated Learning, IJCAI 2019 (CORE rank: A*) workshop on Federated Machine Learning.

[9] Lei K[1], **Huang J**[1], et al. Semantic Similarity Measures to Disambiguate Terms in Medical Text. ICONIP 2018 (CORE rank: B).

[10] Shen Y[1], **Huang J**[1], et al. Discovering Medical Entity Relations from Texts using Dependency Information. IJCAI 2019 (CORE rank: A*) workshop.

[7] Lei K, Du M, **Huang J**, et al. Groupchain: Towards a Scalable Public Blockchain in Fog Computing of IoT Services Computing, IEEE Transactions on Services Computing (IF: 5.823).

Research Experience

Privacy and Security of Multi-server Federated Learning (Ongoing)

Project: DML

Distributed Intelligent Systems Lab (TU Delft)

Dec. 2022 - present

- Gradient transmission in Federated Learning system leaks information of the training data. It is even possible for the server in Federated Learning systems to reconstruct clients' training data. We study how multi-server is able to **collude** for better **gradient inversion** attacks.

Security in Federated Learning and Distributed GAN

Project: DML

Distributed Intelligent Systems Lab (TU Delft) *Aug. 2021 - Dec. 2022*

- Federated Learning systems appear to be vulnerable towards attacks due to multiple anonymous parties. We propose the **data-free** untargeted attack and defense of classifier, also attack and defense of free-riders in **multi-discriminator GAN**.
- Research results on data-free untargeted attack were **published** on (**DSN 2023**)[1] and free-riding MD-GAN was **published** to (**FC 2023**)[4].
- Cooperated research work as the 2rd author on knowledge extraction and model stealing was **published** on (**ECML 2023**)[5]

Data Heterogeneity in Federated Learning

Project: DML

Distributed Intelligent Systems Lab (TU Delft) *Dec. 2020 - Aug. 2021*

- **Maverick** is an important but overlooked heterogeneous data distribution, where specific clients own exclusive data in the system. We theoretically analyzed why **contribution-based** client selection failed and propose the distance-based way with **convergence guarantee** provided.
- Research results were published on (**PAKDD 2023**)[2] and the extended version was **submitted** to (**IEEE TPDS**)[3].

User's Contribution in Federated Learning

Project: IEN

TU Delft & Peking University *June. 2018 - Dec. 2020*

- Measuring client's contribution in Federated Learning system is challenging as the real data from each client is not reachable. This work measures the contribution from the model training prospective, and also observes into the existing malicious behaviors.
- Partial study results on **Attention-based Updates Aggregation** were **published** on (**IJCAI2019**)[8] workshop of Federated Learning.
- An exploratory vision and survey on **incentive** and **attack** were **published** on (**TPS2020**)[7].

Medical Semantic Similarity Measures

Project: IASO

Institute of Big Data Technologies Shenzhen Key Lab for Cloud Computing Technology & Applications *Aug. 2017 - June. 2018*

- Chinese Medical Semantic Similarity Measure. Improved recognition efficacy and accuracy by introducing **linguistic** features including pinyin, radical and edit distance, and **global context** extracted from search engines.
- Research results were **published** on (**ICONIP2018**)[9] as the first student author.

Medical Entity Relation Extraction

Project: IASO

Institute of Big Data Technologies Shenzhen Key Lab for Cloud Computing Technology & Applications *Aug. 2017 - Apr. 2018*

- Proposed RED, a neural network for Relation Extraction that fully explores Dependency information and incorporates that information into deep neural networks. The **long-range relation** between entities can be captured by organizing a sentence as a **dependency tree**, while the requirement for a large amount of training data can be reduced with the abstract-level features generated by dependency information.
- Research results were **published** on (**IJCAI2019**)[10] workshop, as the first student author.
- System framework: Java/SpringBoot/Nignx/FreeMarker/MySQL/Mybatis/Redis. The project displays a **drug knowledge map** with **45W** entities to researchers in the form of RESTful API, with thousands of page views monthly. See in `www.iasokg.com/treeShow`.

Internships

Tencent

Shenzhen, China

Intern of researching on Artifical Intelligence in Platform and Content Group *04-09. 2019*

- Independently completed the **Concept Graph** mining task of QQ browser, gradually learned Spark applications, extracted millions of general concepts and trained a **hypernym and hyponym** matching model, reaching accuracy of **98.5%**.
- Work in Knowledge Graph Guided Semantic Distance for Object Detection was **submitted** to **ISWC2021**.

China Mechants Bank

Shenzhen, China

as one of the Selected 300 members out of 8000+ candidates for Fintech Elite Boot Camp *06. 2019*

- **Fintech Elite Boot Camp**, Learn the application of artificial intelligence algorithm in the financial field, get the **first place** in the **Financial Event Extraction** project, and win the title of "diamond elite team" with 10000 RMB.