Jiyue **Huang**

TU Delft · Netherlands

(+31) 625502430 | j.huang-4@tudelft.nl | gillhuang-xtler.github.io/

Education

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

Ph.D in distributed Machine Learning — *Sep, 2020 - present*

- First 3 Years **7** academic papers **published** while **4 as first author** in the area of privacy/security of Federated learning.
- Supervised **2 master thesis** (with publication) and **9 bachelor thesis** (with submission). **Guest lecturer** and **TA** for CS4290 Seminar (2 years) at TU Delft and 62122 distribute deep learning systems at University of Neuchatel.
- Dutch language ongoing, now at level: A1–A2.

Peking University (School of Electronics and Computer Engineering) — *China*

M.Phil. in Computer Application Technology — *Sep, 2017 - June, 2020*

- **Academic degree** of 3-year independent research oriented projects.
- **6** academic papers **published**, **4** named as the first student author.
- **Leading member** of organizing IEEE HotICN 2018 conference.
- Applied **4 funding** and **2 grants approved** summed up to 2,200,000 CNY.

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) — *China*

B.E. in Electronic Information Engineering — *Sep, 2013 - July, 2017*

- GPA ranking **Top 10%**, **5** awards and **1** scholarship.

Research Experience

Privacy of diffusion models (ongoing) — *Project: DML*

Distributed Intelligent Systems Lab (TU Delft) — *Dec. 2023 - present*

- Diffusion models are state-of-the-art generative models to generate high-quality data. We explore the privacy risk of diffusion models in terms of **model stealing** and **data reconstruction**.
- Research results on single-fold model stealing of diffusion models was submitted to (**ICML 2024**)[–].

Privacy of Multi-server Federated Learning — *Project: DML*

Distributed Intelligent Systems Lab (TU Delft) — *Dec. 2022 - Dec. 2023*

- 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 the increased risk of joining multiple tasks and how multi-server is able to **collude** for **gradient inversion** attacks with Nash Bargaining Game.
- Research results on quantifying privacy risk on data reuse was submitted to (**PETS 2024**)[–] and **presented** by (ICT.OPEN 2024).

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 classifiers, 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].
- Research results 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 fails 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 is from the model training prospective based on **game theory**, 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 **incentives** and **attacks** 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*

- Researching on similarity measure of Chinese medical semantic entities. 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**)[10] 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*

- Proposing a relation extraction method 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**)[9] 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`.

Teaching Experience

Guest Lecturer

University of Neuchatel

62122 Distributed Deep Learning Systems — *2024*

- Providing **guest lecture** on the topic of advance attacks in distributed learning systems. Designing group task on novel idea for privacy attack to inspire students' research interests.

Guest Lecturer and Teaching Assistant

TU Delft

CS4290 Seminar on Distributed Machine Learning — *2021-2023*

- Providing **guest lecture** on the topic of attacks and defenses on Federated Learning. Selecting papers for students reviewing and **grading** with **detailed** feedback to cultivate students' research skills.

Supervisor

TU Delft

Master thesis, Bachelor thesis — *2021-2023*

- Supervising 2 master students on designing gradient inversion attack. One research results are **published** in **SRDS 2022**.
- Supervising **9** bachelor students on attacks and defenses of learning models. One bachelor thesis turns into a **submitted** workshop paper.

Teaching Assistant

Peking University

Master course 04711990-Internet FinTech — *2018-2020*

- Preparing course materials and design assignments for 2 consecutive academic years.

Supervisor

Tianjin University

Bachelor 4-year project — *2015-2017*

- **Solely** supervising 30 bachelor students on answering lessons' questions and helping with career plans.

Funding Application

Funding Ref.No. JCYJ20170412151008290

Peking University

Research focus funding of Shenzhen — *Jul. 2017 - Dec. 2020*

- Research on the efficient and controllable security architecture of the sensor based Internet of Things integrating blockchain and content network. **Granted for 2,000,000 CNY**.
- For all funding applications in Peking University, Our group of 3 master students are responsible for choosing research questions, proposing solutions and writing application proposal. My supervising team is responsible for presentation and interview.

Funding Ref.No. 2020B0101090003

Peking University

Key-Area Research and Development Program of Guangdong — *Aug. 2020 - Aug. 2023*

- Research on key technologies of independent and controllable consortium blockchain. **Granted for 200,000 CNY**.

Applied

- HealthDEAL Federated Learning Business Framework with Health Data Equity and Access via Ledger. *Trustchain, European Union:* under review
- Research on Privacy-Preserving Time Series Forecasting. *AI Convergence FinTech Fund, the Netherlands:* under review
- Research on secure transmission and privacy protection technology of network data based on blockchain authentication and authorization. *National Natural Science Foundation of China:* **Proposal round passed**, project funding cancelled for that year.
- Demonstration project of fruit and vegetable product traceability system based on big data and blockchain. *Guangdong Department of Science and Technology:* **Proposal round passed**, interview round failed.

Internships

Tencent Holdings Limited *Shenzhen, China*

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

- Independently completed the **Concept Graph** prototype 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%**.
- Working for 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*

- Joining as a selected member of **Fintech Elite Boot Camp**, applying Deep learning algorithms in the financial field, get the **first place** in the **Financial Event Extraction** project, and win the title of "diamond elite team" with 10,000 CNY.

Honors & Awards

2023	**Travel Grants**, Invited talk for ISSRE 2023 Workshop on Dependability Modeling and Design by Huawei	*Florence, Italy*
2023	**Selected Mentee**, 2nd F+Cube workshop on famale researchers in STEM fields	*TU Delft*
2021	**Selected student chair**, 22nd ACM/IFIP International Middleware Conference	*Online*
2019	**"Diamond Elite Team"**, Award for National Fintech Elite Boot Camp, Offline Competition	*Merchants Bank*
2019	**Top 100 Certification**, Award for National Fintech Online Competition over 8000+ competitors	*Merchants Bank*
2019	**Advanced Individual for Scientific Research**, Annual Student Achievement Assessment	*Peking University*
2018	**Merit Student Award**, Annual Student Achievement Assessment of Peking University	*Peking University*
2018	**First Prize**, 26th Challenge Cup Competition for Innovation in Shenzhen Graduate School	*Peking University*
2017	**Selected Outstanding Undergraduate**, National summer camp in Information Engineering	*Peking University*
2017	**Selected Member**, Awarded camp member of Institution of Microsystem and Information Technology	*CAS*
2017	**Outstanding Graduate Award**, Annual Student Achievement Assessment for Graduations (2/88)	*Tianjin University*
2016	**Zhonghuan Electronic Scholarship**, Awards for Promising Researchers	*Tianjin University*
2016	**Advanced Individual for Volunteering.**, Annual Student Achievement Assessment of Tianjin University	*Tianjin University*
2016	**Advanced Individual for Social Services.**, Annual Student Achievement Assessment of Tianjin University	*Tianjin University*
2014-16	**Merit Student Award**, Annual Student Achievement Assessment of Tianjin University	*Tianjin University*

Community

2024	**Volunteer**, 41st International Conference on Machine Learning (ICML 2024, under selection).	*Vienna, Austria*
2021	**Volunteer**, 22nd ACM/IFIP International Middleware Conference (Middleware 2021).	*Online*
2019	**Volunteer**, IEEE International Conference on Hot Information-Centric Networking (HotICN 2019).	*Chongqing, China*
2018	**Organizers**, IEEE International Conference on Hot Information-Centric Networking (HotICN 2018).	*Shenzhen, China*
2020-23	**Reviewer**, IEEE TPDS, IEEE TDSC.	*Online*
2018-23	**Reviewer**, USENIX ATC /USENIX Security/NDSS/S&P/DSN/Infocom/Middleware/WWW/AAAI/SigMetrics.	*Online*

Publications

[-] Hong C, **Huang J**, et al, SFDDM: Single-fold Distillation for Diffusion model. ICML 2024 (CORE rank: A*, under review).

[-] **Huang J**, Roos S, On Quantifying the Gradient Inversion Risk of Data Reuse in Federated Learning Systems. PETS 2024 (CORE rank: A, under review).

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

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

[11] 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 (CORE rank: A*).