

Marc Chengliang Zhang

| | | |
|--------------------|--|--|
| PRESENT POSITION | Ph.D Candidate Dept. of Computer Science & Engineering Hong Kong University of Science and Technology | Phone: +852-6216-2287 Email: czhangbn@cse.ust.hk Web: https://marcoszh.github.io/ |
| RESEARCH INTERESTS | My interests cover big data analytics systems and cloud computing , with a special focus on machine learning systems . I enjoy identifying system design, performance, and privacy issues in large-scale ML systems for both training and inference, and searching for general and efficient solutions. | |
| EDUCATION | Hong Kong University of Science and Technology , Hong Kong SAR <i>Department of Computer Science and Engineering</i> ◇ Ph.D. Computer Science and Engineering September 2016 - March 2021 ◇ <i>Supervisor: Wei Wang</i> ◇ Hong Kong PhD Fellowship recipient: prestigious and highly selective Harbin Institute of Technology , Harbin, China <i>School of Computer Science and Technology</i> ◇ B.Eng. Software Engineering September 2012 - June 2016 ◇ <i>Honors:</i> National Scholarship (Top 2%), Fuji Xerox Scholarship | |
| PUBLICATIONS | Conference †: top venue acc. CSRankings Chengliang Zhang , Suyi Li, Junzhe Xia, Wei Wang, Feng Yan, Yang Liu “ BatchCrypt: Efficient Homomorphic Encryption for Cross-Silo Federated Learning ,” in the <i>Proceedings of USENIX Annual Technical Conference (ATC’20)</i> †, Boston, MA, July 2020. Jun Yi , Chengliang Zhang , Wei Wang, Cheng Li, Feng Yan, “ Not All Explorations Are Equal: Harnessing Heterogeneous Profiling Cost for Efficient MLaaS Training ,” in the <i>Proceedings of IEEE International Parallel and Distributed Processing Symposium (IPDPS’20)</i> , New Orleans, LA, May 2020. Chengliang Zhang , Minchen Yu, Wei Wang, Feng Yan, “ MARk: Exploiting Cloud Services for Cost-Effective, SLO-Aware Machine Learning Inference Serving ,” in the <i>Proceedings of USENIX Annual Technical Conference (ATC’19)</i> †, Renton, WA, July 2019. Chengliang Zhang , Huangshi Tian, Wei Wang, Feng Yan, “ Stay Fresh: Speculative Synchronization for Fast Distributed Machine Learning ,” in the <i>Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS’18)</i> , Vienna, Austria. Journal Chengliang Zhang , Minchen Yu, Wei Wang, Feng Yan, “ Enabling Cost-Effective, SLO-Aware Machine Learning Inference Serving on Public Cloud ,” in the <i>IEEE Transactions on Cloud Computing</i> . Yinghao Yu , Chengliang Zhang , Wei Wang, Jun Zhang, Khaled Letaief, “ Towards Dependency-Aware Cache Management for Data Analytics Applications ,” in the <i>IEEE Transactions on Cloud Computing</i> . Under Review Chengliang Zhang , Junzhe Xia, Baichen Yang, Huancheng Puyang, Wei Wang, Ruichuan Chen, Istemi Ekin Akkus, Paarijaat Aditya, Feng Yan, “ Citadel: Protect Data Privacy and Model | |

Confidentiality for collaborative Learning with SGX.

RESEARCH EXPERIENCE

Research Scientist Intern @ Bell Labs Germany

- Project owner and implementation lead
- Designed and realized a distributed system for confidential collaborative machine learning
- Aspired to protect both data privacy and model privacy
- Utilized Trusted Execution Environment

Fast Secure Federated Learning System

- Project owner and implementation lead
- Proposed and implemented a method to speedup inter-enterprise federated learning
- Accelerated Homomorphic Encryption for federated learning training
- Mitigated encryption and network overhead by quantizing and batching intermediate results

MArk: ML Serving on Public Cloud

- Project owner and implementation lead
- Designed a large-scale machine learning inference serving system on public cloud
- Aspired to be both cost-effective and SLO-aware
- Characterized ML serving and its performance on cloud services and hardware accelerators
- Combined FaaS and IaaS to reduce over-provisioning

Speculative Synchronization

- Project owner and implementation lead
- Aimed at distributed data parallel training with relaxed consistency
- Actively improved consistency by re-synchronization if deemed beneficial

SKILLS

- | | |
|---------------------------------------|--|
| ◇ Python: proficient | ◇ Java, Scala, C++: familiar with |
| ◇ Keras, TensorFlow: proficient | ◇ Mandarin Chinese: native |
| ◇ English: proficient, <i>IELTS 8</i> | ◇ Cantonese: limited working proficiency |

REFERENCES

Wei Wang, Assistant Professor

Room 3524, CSE Department, Hong Kong University of Science and Technology

Phone: +852-2358-6972 Email: weiwa@cse.ust.hk

Web: <https://www.cse.ust.hk/~weiwa/>

Feng Yan, Assistant Professor

SEM 233, CSE Department, University of Nevada, Reno

Phone: +1-775-784-6448 Email: fyan@unr.edu

Web: <https://www.cse.unr.edu/~fyan/>

Ruichuan Chen, Distinguished Member of Technical Staff

Nokia Bell Labs

Lorenzstrasse 10, 70435 Stuttgart, Germany

Email: ruichuan.chen@nokia-bell-labs.com

Web: <https://www.ruichuan.org/>