Marc Chengliang Zhang, PhD

PRESENT Data Scientist, Associate Phone: +852–6216–2287

POSITION Seleya Technologies Email: marc.zhang@seleyatech.com

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

Department of Computer Science and Engineering

♦ Ph.D. Computer Science and Engineering

September 2016 - March 2021

♦ Supervisor: Wei Wang

Hong Kong PhD Fellowship recipient: prestigious and highly selective

Harbin Institute of Technology, Harbin, China School of Computer Science and Technology

♦ **B.Eng.** Software Engineering

September 2012 - June 2016

♦ Honors: National Scholarship (Top 2%), Fuji Xerox Scholarship

PUBLICATIONS

Conference

Chengliang Zhang, Junzhe Xia, Baichen Yang, Huancheng Puyang, Wei Wang, Ruichuan Chen, Istemi Ekin Akkus, Paarijaat Aditya, Feng Yan "Citadel: Protecting Data Privacy and Model Confidentiality for Collaborative Learning," in the *Proceedings of ACM Symposium on Cloud Computing (SoCC '21)*, Seattle, WA, Nov 2021.

Chengliang Zhang, Suyi Li, Junzhe Xia, Wei Wang, Feng Yan, Yang Liu "BatchCrypt: Efficient Homomorphic Encryption for Cross-Silo Federated Learning," in the *Proceedings of USENIX Annual Technical Conference (ATC'20)*, Boston, MA, July 2020.

Jun Yi, <u>Chengliang Zhang</u>, Wei Wang, Cheng Li, Feng Yan, "Not All Explorations Are Equal: Harnessing Heterogeneous Profiling Cost for Efficient MLaaS Training," in the *Proceedings of IEEE International Parallel and Distributed Processing Symposium (IPDPS'20)*, New Orleans, LA, May 2020.

<u>Chengliang Zhang</u>, Minchen Yu, Wei Wang, Feng Yan, "MArk: Exploiting Cloud Services for Cost-Effective, SLO-Aware Machine Learning Inference Serving," in the *Proceedings of USENIX Annual Technical Conference (ATC'19)*, Renton, WA, July 2019.

<u>Chengliang Zhang</u>, Huangshi Tian, Wei Wang, Feng Yan, "Stay Fresh: Speculative Synchronization for Fast Distributed Machine Learning," in the *Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS'18)*, 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 *IEEE Transactions on Cloud Computing*.

Yinghao Yu, <u>Chengliang Zhang</u>, Wei Wang, Jun Zhang, Khaled Letaief, "Towards Dependency-Aware Cache Management for Data Analytics Applications," in the *IEEE Transactions on Cloud Computing*.

PhD Thesis

Chengliang Zhang, "Towards Efficient and Secure Large-Scale Systems for Distributed Machine Learning Training.

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 laaS 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

 Cantonese: limited working proficiency ♦ English: proficient, IELTS 8

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/