

2022

HKUST CSE AWARD PAPERS SEMINAR

NOVEMBER 12
SATURDAY

09:30am - 11:30am | Zoom 

ZOOM LINK

<https://cutt.ly/USTCSE>



Meeting ID: 928 9690 5625
Passcode: 532599

SIGCOMM Best Paper

Software-Defined Network Assimilation: Bridging the Last Mile Towards Centralized Network Configuration Management with NAssim



DR. HUANGXUN CHEN

CSE PhD Graduate

Researcher,
Huawei Hong Kong
Research Center



DR. LI CHEN

CSE PhD Graduate

Assistant Researcher,
Zhongguancun
Laboratory

Co-authors:

Wei Wang (HKUST GZ)

Yukai Miao, Haifeng Sun, Hong Xu, Libin Liu, Gong Zhang

SIGMOD Best Paper

R2T: Instance-optimal Truncation for Differentially Private Query Evaluation with Foreign Keys



MR. WEI DONG

CSE PhD Candidate

Co-authors:

Juanru Fang (HKUST), Ke Yi (HKUST),

Yuchao Tao, Ashwin Machanavajjhala

VLDB Best Paper

SANCUS: Staleness-Aware Communication-Avoiding Full-Graph Decentralized Training in Large-Scale Graph Neural Networks



DR. JINGSHU PENG

CSE PhD Graduate

Co-authors:

Zhao Chen (HKUST), Lei Chen (HKUST),

Yingxia Shao, Yanyan Shen, Jiannong Cao



DR. HUANGXUN CHEN

SIGCOMM
Best Paper Awardee



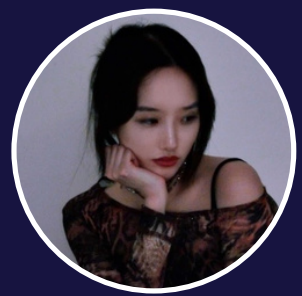
DR. LI CHEN*

SIGCOMM
Best Paper Awardee



MR. WEI DONG

SIGMOD
Best Paper Awardee



DR. JINGSHU PENG

VLDB
Best Paper Awardee

SIGCOMM BEST PAPER

Title: Software-Defined Network Assimilation: Bridging the Last Mile Towards Centralized Network Configuration Management with NAssim

Authors: Huangxun Chen(HKUST), Yukai Miao, Li Chen(HKUST), Haifeng Sun, Hong Xu, Libin Liu, Gong Zhang, Wei Wang(HKUST GZ)

Abstract

On-boarding new devices into an existing SDN network is a pain for network operations (NetOps) teams, because much expert effort is required to bridge the gap between the configuration models of the new devices and the unified data model in the SDN controller. In this work, we present an assistant framework NAssim, to help NetOps accelerate the process of assimilating a new device into a SDN network. Our solution features a unified parser framework to parse diverse device user manuals into preliminary configuration models, a rigorous validator that confirm the correctness of the models via formal syntax analysis, model hierarchy validation and empirical data validation, and a deep-learning-based mapping algorithm that uses state-of-the-art neural language processing techniques to produce human-comprehensible recommended mapping between the validated configuration model and the one in the SDN controller. In all, NAssim liberates the NetOps from most tedious tasks by learning directly from devices' manuals to produce data models which are comprehensible by both the SDN controller and human experts. Our evaluation shows, NAssim can accelerate the assimilation process by 9.1x. In this process, we also identify and correct 243 errors in four mainstream vendors' device manuals, and release a validated and expert-curated dataset of parsed manual corpus for future research.

Speaker Bio

Dr. Li CHEN, the corresponding author of the award-winning paper, was a 2018 Ph.D. graduate of the CSE department supervised by Prof. Kai CHEN and is currently an assistant researcher at Zhongguancun Laboratory. Dr. Huangxun CHEN, the first author of the award-winning paper, was a 2020 Ph.D. graduate of the CSE department supervised by Prof. ZHANG Qian and is currently a researcher at Huawei Hong Kong Research Center.

*Dr. Li CHEN will be in Q&A session

SIGMOD BEST PAPER

Title: R2T: Instance-optimal Truncation for Differentially Private Query Evaluation with Foreign Keys

Authors: Wei Dong(HKUST), Juanru Fang(HKUST), Ke Yi(HKUST), Yuchao Tao, Ashwin Machanavajjhala

Abstract

Answering SPJA queries under differential privacy (DP), including graph pattern counting under node-DP as an important special case, has received considerable attention in recent years. The dual challenge of foreign-key constraints and self-joins is particularly tricky to deal with, and no existing DP mechanisms can correctly handle both. For the special case of graph pattern counting under node-DP, the existing mechanisms are correct (i.e., satisfy DP), but they do not offer nontrivial utility guarantees or are very complicated and costly. In this paper, we propose the first DP mechanism for answering arbitrary SPJA queries in a database with foreign-key constraints. Meanwhile, it achieves a fairly strong notion of optimality, which can be considered as a small and natural relaxation of instance optimality. Finally, our mechanism is simple enough that it can be easily implemented on top of any RDBMS and an LP solver. Experimental results show that it offers order-of-magnitude improvements in terms of utility over existing techniques, even those specifically designed for graph pattern counting.

Speaker Bio

Mr. Wei Dong is now a Ph.D. candidate in the Department of Computer Science and Engineering at HKUST supervised by Prof. Ke Yi. His general areas of interest include database theory and algorithms, data security and privacy, and statistics. His research has been recognized by the academic community and appeared in top conferences, such as SIGMOD, PODS, NeuIPS, CCS, and KDD.

VLDB BEST PAPER

Title: SANCUS: Staleness-Aware Communication-Avoiding Full-Graph Decentralized Training in Large-Scale Graph Neural Networks

Authors: Jingshu Peng(HKUST), Zhao Chen(HKUST), Yingxia Shao, Yanyan Shen, Lei Chen(HKUST), Jiannong Cao

Abstract

Graph neural networks (GNNs) have emerged due to their success at modeling graph data. Yet, it is challenging for GNNs to efficiently scale to large graphs. Thus, distributed GNNs come into play. To avoid communication caused by expensive data movement between workers, we propose SANCUS, a staleness-aware communication-avoiding decentralized GNN system. By introducing a set of novel bounded embedding staleness metrics and adaptively skipping broadcasts, SANCUS abstracts decentralized GNN processing as sequential matrix multiplication and uses historical embeddings via cache. Theoretically, we show bounded approximation errors of embeddings and gradients with convergence guarantee. Empirically, we evaluate SANCUS with common GNN models via different system setups on large-scale benchmark datasets. Compared to SOTA works, SANCUS can avoid up to 74% communication with at least 1.86X faster throughput on average without accuracy loss.

Speaker Bio

Jingshu Peng holds a PhD in Computer Science and Engineering from HKUST, supervised by Prof. Lei Chen and Prof. Lionel Ni. Her research focus is to scale graph learning for efficiency and effectiveness. But anything towards this goal - including new GNN models, algorithms, systems, and algorithm-system co-designs, can be of her interest. Her work has just received the Best Regular Research Paper Award at VLDB 2022.