Xinyi (Cindy) CHEN

Email: cxinyic@seas.upenn.edu | Mobile: (+1)215-578-1438

EDUCATION

University of Pennsylvania (UPenn) Sept. 2019 – present
Ph.D. student in Computer Science *Philadelphia, United States*
Advisor: Prof. Vincent Liu

Shanghai Jiao Tong University (SJTU) Sept. 2015 – Jun. 2019
B.S. in Computer Science *Shanghai, China*
GPA: 3.80/4.0

PUBLICATIONS

[1] Qizhen Zhang, **Xinyi Chen**, Sidharth Sankhe, Zhilei Zhang, Ke Zhong, Sebastian Angel, Ang Chen, Vincent Liu, Boon Thau Loo, "Optimizing Data-intensive Systems in Disaggregated Data Centers with TELEPORT", in *the ACM Internal Conference on Management of Data (SIGMOD)*, 2022, to appear.

[2] Qizhen Zhang, Yifan Cai, **Xinyi Chen**, Sebastian Angel, Ang Chen, Vincent Liu, Boon Thau Loo, "Understanding the Effect of Data Center Resource Disaggregation on Production DBMSs", in *the 46th International Conference on Very Large Data Bases (VLDB)*, 2020.

[3] Shaobo Wang, Hui Lyu, Jiachi Zhang, Chenyuan Wu, **Xinyi Chen**, Wenchao Zhou, Boon Thau Loo, Susan B. Davidson, Chen Chen, "Provenance for Probabilistic Logic Programs", in *the 23th International Conference on Extending Database Technology (EDBT)*, 2020 (*Best Paper Award*).

[4] Yin Lin, **Xinyi Chen**, Xiaofeng Gao, Bin Yao, Guihai Chen, "R2 -Tree: An Efficient Indexing Scheme for Server-Centric Data Center Networks", accepted by *International Conference on Database and Expert Systems Applications (DEXA)*, 2018

RESEARCH EXPERIENCE

Fault Tolerance in Spot VMs June. 2020 – Present

UPenn – Advisors: Prof. Vincent Liu

- Propose a learning-based fault tolerance system. It can decide when to make checkpoints based on the workload and spot vm configurations.
- Provide cheap checkpoint by combining flushing dirty pages with normal evictions.

DBMSs in Data Center Resource Disaggregation Sept. 2019 – Mar. 2020

UPenn – Advisors: Prof. Vincent Liu and Prof. Boon Thau Loo

- Run two popular open-source DBMSs (MonetDB and PostgreSQL) and test their performance with TPC-H benchmark in a recently released operating system for resource disaggregation.
- Evaluate the two DBMS with various configurations and compare their performance with that of single-machine Linux with the same hardware resources.
- The results show that there are significant performance degradation in disaggregation data center when running DBMS.

Data Provenance for Probabilistic Programming Jul. 2018 – Oct. 2018

Reseach Intern in distributed system group, UPenn – Advisor: Prof. Boon Thau Loo

- Designd a unified language called P3log (based on Datalog) which can capture different probabilistic programming language models in distributed system and machine learning.
- Supported provenance in system and proposed several novel provenance queries, for example, identifying the most influential tuples.
- Developed a prototype of P3log and the evaluation shows P3log can maintain provenance trees with low overhead.

SKILLS

C/C++, Python, Java, Latex