MINH H. PHAM

https://minhhpham.github.io

QUALIFICATION SUMMARY

- Ph.D. Candidate in Computer Science (2024), Master's Degree in Statistics (2018)
- Winner of the 2021 IEEE Big Data Cup Challenge in Reinforcement Learning
- Runner up of the eBay 2022 University Machine Learning Competition
- 3 years of working experience as system administrator for a high-performance computing cluster
- Ph.D. concentration in parallel processing, GPGPU, and database

EDUCATION

University of South Florida, Tampa, FL 5/2024
- Ph.D. in Computer Science (GPA 4.0)

University of South Florida, Tampa, FL 5/2018
- Master of Arts in Statistics (GPA 3.97)
- Thesis: Signal Detection of Adverse Drug Reaction using the Adverse Event Reporting System: Literature Review and Novel Methods

University of South Florida, Tampa, FL 5/2016
- Bachelor of Arts in Statistics (GPA 3.97, USF Dean's List, Honor College, and Summa Cum Laude)

WORKING EXPERIENCE

PhD Intern, Meta – Menlo Park, CA 5/2023 – 8/2023
- Build simulation for the Presto distributed query engine
- Optimize resource usage on >200 query clusters
- Increased query throughput by 10% and average wall time by 9%, ≈ $2.2 million (0.2 MW) saving

PhD Intern, Meta – Menlo Park, CA 5/2022 – 8/2022
- Query optimization for Presto distributed query engine with machine learning
- Testing showed average saving of 7% in CPU time and 4% wall time across all queries

Graduate Assistant, University of South Florida– Tampa, FL 8/2019 – Now
- System administrator for a high-performance computing cluster
- Develop web servers (back-end and front-end)

SOFTWARE

Dynamic Memory Management on CUDA	https://github.com/minhhpham/parallel-GPU-memory-management
Credit Card Approval Chance	https://minhhpham.github.io/credit-cards
Wastewater Management Web Application	https://wastewater.csee.usf.edu/
Pharmacovigilance Algorithms	https://github.com/minhhpham/MultiPharma
NGS short-read aligner implementation	https://github.com/minhhpham/bwa

PROGRAMMING SKILLS

- C, C++, CUDA: used in all research projects
- Python: used in all machine learning/data mining projects

PUBLICATIONS

Under Review

Minh Pham, Chengcheng Mou, Benjamin Hsu, and Yicheng Tu. Computing Group-By and Aggregate on Massively Parallel Systems.

Minh Pham, Yongke Yuan, Hao Li , Chengcheng Mou, Zichen Xu, and Yicheng Tu. Dynamic Buffer Management in Massively Parallel Systems: The Power of Randomness.

Conference

Minh Pham, Yicheng Tu , Weijia Xu. Accelerating BWA-MEM Read Mapping on GPUs. In *Proceedings of the 37th International Conference on Supercomputing* (pp. 155-166).

Minh Pham, Hao Li, Yongke Yuan, Chengcheng Mou, Kandethody Ramachandran, Zichen Xu, and Yicheng Tu. 2022. Dynamic memory management in massively parallel systems: a case on GPUs. In *Proceedings of the 36th ACM International Conference on Supercomputing (ICS '22)*.

M. Pham, H. Nguyen, L. Dang, J.A. Nieves (2021). Compressive Features in Offline Reinforcement Learning for Recommender Systems. *IEEE International Conference on Big Data 2021*.

J. Adorno Nieves, **M. Pham,** S. Barbeau, A. Labrador (2019). Scalable Real-Time Transit Data Archiving: A Framework for Performance Assessment and Machine Learning Prediction. In *Transportation Research Board Conference Proceedings* (No. 55)

M. Pham, J. Lin, and Y. Zhang, "Diagnose Voice Disorder with Machine Learning," In *IEEE International Conference on Big Data*, 2018.

Journal

M. Pham, F. Cheng, and K. Ramachandran, "A Comparison Study of Algorithms to Detect Drug–Adverse Event Associations: Frequentist, Bayesian, and Machine-Learning Approaches," *Drug Safety*, vol. 42, no. 6, pp. 743–750, 2019.

Y. Lu, A. Ramachandra, **M. Pham**, Y.-C. Tu, and F. Cheng, "CuDDI: A CUDA-Based Application for Extracting Drug-Drug Interaction Related Substance Terms from PubMed Literature," *Molecules*, vol. 24, no. 6, p. 1081, 2019.

Z. Tang, **M. Pham**, Y. Hao, F. Wang, D. Patel, L. Jean-Baptiste, L. Fan, W. Wang, Y. Wang, and F. Cheng, "Sex, Age, and BMI Modulate the Association of Physical Examinations and Blood Biochemistry Parameters and NAFLD: A Retrospective Study on 1994 Cases Observed at Shuguang Hospital, China," *BioMed Research International*, 2019.

Y. Hao, F. Cheng, **M. Pham**, H. Rein, D. Patel, Y. Fang, Y. Feng, J. Yan, X. Song, H. Yan, and Y. Wang, "A Noninvasive, Economical, and Instant-Result Method to Diagnose and Monitor Type 2 Diabetes Using Pulse Wave: Case-Control Study," *JMIR mHealth and uHealth*, vol. 7, no. 4, 2019.

M. H. Pham, C. Tsokos, and B.-J. Choi, "Maximum Likelihood Estimation for the Generalized Pareto Distribution and Goodness-of-Fit Test with Censored Data," *Journal of Modern Applied Statistical Methods*, vol. 17, no. 2, 2019.

Thesis

M. H. Pham. (2018). *Signal Detection of Adverse Drug Reaction using the Adverse Event Reporting System: Literature Review and Novel Methods*.

INVITED REVIEWER FOR:

- IEEE Transactions on Intelligent Transportation Systems
- Scientific Reports by Nature
- Clinical Drug Investigation by Springer

AWARDS & HONORS

- 2021 IEEE Big Data Cup Challenge: 1st place out of 22 teams 9/2021
- 2018 IEEE Big Data Cup Challenge: 8th place out of 109 teams from 27 countries 11/2018

LEADERSHIPS & ACTIVITIES

- **Web Chair**, IEEE International Conference on Data Mining 2021
- **Web Chair**, International Conference on Scientific and Statistical Database Management 2021