

# Chenghao Lyu

Email: [chenghao@cs.umass.edu](mailto:chenghao@cs.umass.edu)

Homepage: <https://people.cs.umass.edu/~chenghao>

---

## EDUCATION

- University of Massachusetts Amherst, MA, USA** 2018.9 - Present
- **Ph.D.**, Computer Science, advisors: Prof. Yanlei Diao, Prof. Prashant J. Shoney
  - **MSc**, Computer Science (2021.6)
- Fudan University, Shanghai, China** 2015.9 - 2018.7
- **MSc**, Computer Science, advisor: Prof. X. Sean Wang
- Fudan University, Shanghai, China** 2011.9 - 2015.7
- **BS**, Electronic Engineering (Math-related GPA: 3.79/4.0, Ranking 10/100)
- 

## RESEARCH EXPERIENCE

**Scientific Collaborator** 2021.10 - Present

*Inria and LIX, Ecole Polytechnique, Paris, France*

Working on the [BigFastData](#) project in the CEDAR project-team of Inria and LIX.

**Research Intern (Mentor: Kai Zeng)** 2021.3 – 2021.12

*Alibaba DAMO Academy, Remote & Hangzhou, China*

Proposed a new architecture for big data processing that breaks the resource optimization (RO) into a series of simpler problems, new fine-grained predictive models, and novel optimization methods that exploit these models to make effective instance-level recommendations in a hierarchical MOO framework. Our evaluation using 0.6M production workloads and a simulator of the extended Alibaba MaxCompute environment showed that the new RO system saved 36-37% latency and 37-75% cost compared to the current scheduler.

**Research Assistant (advisor: Yanlei Diao, Prashant Shenoy)** 2019.1 – Present

*College of Information and Computer Sciences, UMass Amherst, MA, USA*

Work on a Unified Data Analytics Optimizer (UDAO) that can recommend parameters of a running system (e.g., Apache Spark), so that user performance goals and budgetary constraints, collectively referred to as “objectives”, can be achieved automatically. UDAO collects the runtime traces of submitted, learns objective models, and does multi-objective optimization (MOO) to recommend configurations according to the user preferences. In this project, I contributed a triplet-loss based modeling approach to do the job modeling, and a multi-objective gradient descent (MOGD) solver to boost our MOO approach. With the MOGD solver, our MOO approach beats SOTAs.

**Research Assistant (advisor: X. Sean Wang)** 2015.9 – 2018.7

*School of Computer Science, Fudan University, Shanghai, China.*

Propose an Adaptive Partition Strategy (APS) to distribute data from multi-stream in the mini-batch based distributed stream processing system. Apply the APS to stream data with changeable characteristics and asynchronous arrivals. A maximum improvement of 26.7% in latency can be achieved in real data, compared with the state-of-the-art static partition methods when implemented atop Apache Spark.

**Research Assistant (advisor: Patrick Chiang)** 2015.1 - 2016.9

*OSU VLSI Research Group, School of Electrical Engineering and Computer Science, OSU, OR, USA.*

(1) Applied the Unscented Kalman Filtering method over a fusion of measured trajectories from a bundle of inaccurate Inertial measurement unit (IMU) sensors, and improved the position estimation by A 26% compared to using a single IMU. (2) Proposed a new fitting algorithm capable of estimating the distance among multiple IMU-based devices by (only) using the nine-axis kinematic data. (3) Applied ML algorithms to recognize the gestures by the trajectory of wearable IMU sensors.

---

## PUBLICATION

Lyu C, Fan Q, Song F et al. Fine-Grained Modeling and Optimization for Intelligent Resource Management in Big Data Processing (in submission)

Song F\*, Zaouk K, Lyu C\*, et al. Spark-based Cloud Data Analytics using Multi-Objective Optimization[C]//2021 IEEE 37th International Conference on Data Engineering (ICDE). IEEE, 2021 (\*equal contribution)

Zaouk K, Song F, Lyu C, et al. Neural-based Modeling for Performance Tuning of Spark Data Analytics[J]. arXiv preprint arXiv:2101.08167, 2021.

Zaouk L, Song F, Lyu C, et al. UDAO: A Next-Generation Unified Data Analytics Optimizer[J]. Proceedings of the VLDB Endowment, 2019, 12(12): 1934-1937, 2019.

Zhang Y, Lyu C, Xu H, et al. Improved position estimation by fusing multiple inaccurate inertial measurement unit sensors[C]//Wireless Symposium (IWS), 2016 IEEE MTT-S International. IEEE, 2016: 1-4.

Xu H, Lyu C, Zhang Y, et al. Method for estimating the distance between multiple IMU-based wearable devices[C]//Wireless Symposium (IWS), 2016 IEEE MTT-S International. IEEE, 2016: 1-4.

---

## TEACHING EXPERIENCE

### Teaching Assistant

*UMass Amherst, MA, USA*

- COMPSCI 645 Databases (2020 spring)
- COMPSCI 105 Computer Literacy (2018 fall)

*Fudan University, Shanghai, China*

- SOFT130039.01 Discrete Mathematics (2016 fall)
- 

## PROFESSIONAL SKILLS

Familiar Big Data Framework: Apache Spark, Apache Flink, Apache Kafka, etc.

Familiar ML Frameworks: Scikit-learn, PyTorch, Tensorflow, Keras.

Programming language: Python, Scala, Java, etc.

---

## SELECTED HONORS AND AWARDS

### Academic:

- Tung OOCL Scholarship (First Prize Scholarship) at Fudan University
- Scholarship of Academic Excellence at Fudan University
- Excellent Student Pacesetter at Fudan University
- Outstanding student at Fudan University
- 1st runner-up (top 5) in 2014 Fudan iShamrock Software Competition
- Honorable Mention in 2014 Mathematical Contest in Modeling
- Entrance examination waived for graduate education

### Sports:

- Advanced Individual in Sports at Fudan University
- 1-time Championship in Fudan Basketball Cup.
- 3-time Championships in Fudan Basketball League
- 1-time First Scoring Player in Sub-Campus Basketball Cup.
- First Prize in 2015 Dragon Boat Competition compared with C9 Colleges in China.
- Former Member of Fudan Dragon Boat Team