

Jin Huang

PH.D. · RESEARCH FELLOW · SOFTWARE ENGINEER

Room 714 Doug McDonnell Building The University of Melbourne Victoria Australia

☎ (+61) 0414 980 977 | ✉ huang.j@unimelb.edu.au | 🏠 iojin.com | 📱 jinhuang | 📺 natejinhuang

Education

The University of Melbourne

Melbourne, VIC, Australia

PH.D. IN COMPUTER SCIENCE

February 2012 - December 2015

- Thesis Title: Similarity Analysis with Advanced Relationships on Big Data
- Recipient of Full Fee Remission and Living Allowance Scholarships

South China University of Technology

Guangzhou, Guangdong, P.R. China

B.E. IN SOFTWARE ENGINEERING

September 2007 - July 2011

- GPA 3.62 Ranked 4th among 276 students

Skills

Speciality Apache Hadoop, Apache Spark, Apache Hama

Knowledgeable Java, Scala, Javascript, Meteor, Machine Learning Models

Exposure C/C++, Python, React, MongoDB, Elasticsearch, AWS

Communication English (Fluent), Chinese Mandarin (Native Speaker)

Experience

The University of Melbourne

Melbourne, VIC, Australia

POSTDOCTORAL RESEARCH FELLOW (DATABASE AND KNOWLEDGE MINING)

February 2016 - Present

- Implemented a data integrate pipeline for cleaning, mining, and querying unstructured/semi-structured data
- Implemented a Meteor based web application for users to make use of the integrated data
- Supervised Ph.D. students on topics in large scale data processing and machine learning

The University of Melbourne

Melbourne, VIC, Australia

PH.D. STUDENT (DISTRIBUTED DATA PROCESSING)

February 2012 - December 2015

- Studied topics in distributed data processing, machine learning, and spatial data management and mining
- Developed original ideas that improve the efficiency of distributed data processing techniques by an order of magnitude
- Published **7** papers in prestigious conference and **4** papers in top-tier peer review journals
- Served as a reviewer for leading academic venues including ACM TODS, ACM SIGMOD, VLDBJ, VLDB, IEEE TKDE, and IEEE ICDE

The University of Melbourne

Melbourne, VIC, Australia

TUTOR

February 2013 - December 2015

- COMP20005 Engineering Computation 2015: taught basics of C to undergraduates
- COMP90058 Advanced Spatial-temporal Data Analytics 2015: taught spatial database techniques to postgraduates
- COMP90050 Advanced Database Systems 2015 and 2014: taught literature review and critical analysis in database research to postgraduates
- COMP90041 Programming and Software Development 2015, 2014, and 2013: taught basics of Java and software engineering to postgraduates

GSegement Space Technologies

Beijing, P.R. China

SOFTWARE ENGINEER INTERN

July 2011 - October, 2011

- Assisted on migrating development infrastructure from an in-house platform to the de-facto industry standard
- Developed an astronomical data processing prototype for digesting telescope data in a European Space Agency project

Projects

Large Scale Hypergraph Learning and Processing

GitHub

SOLO


August 2014 - PRESENT

- Designed and developed a novel distributed framework for efficient and effective hypergraph learning and processing
- Tackled challenges of inflated data size, enormous replication cost, and difficulty in balancing distributed workloads
- Implemented in Scala on Spark, which is **orders of magnitude faster** than Spark's GraphX
- 📄 <https://github.com/jinhuang/hyperx/>

Earth Mover's Distance Based Similarity Joins on Hadoop

SOLO


August 2012 - July 2014

- Designed and developed a distributed framework for efficient earth mover's distance joins in a distributed manner
- Tackled challenges of prohibitive computation overhead and highly skewed data distribution
- Implemented in Java on Hadoop with MapReduce and BSP, which is **an order of magnitude faster** than the state-of-the-art join algorithm
-  <https://github.com/jinhuang/heads-join/>

Trajectory Mining and User Mobility Prediction

COLLABORATIVE

February 2012 - August 2015

- Designed mining algorithm to predict user mobility based on historical GPS trajectory
- Tackled the fundamental data sparsity problem with novel hidden markov model based synthesis method
- Developed an interactive web demo to showcase the capability of the algorithm
-  <http://spatialanalytics.cis.unimelb.edu.au/subsyndemo/>

Publications

Large Scale Hypergraph Learning and Processing

PRINCIPAL RESEARCHER

August 2014 - PRESENT

- **Jin Huang**, Rui Zhang, and Jeffrey Xu Yu, "Scalable Hypergraph Processing", in *Proceedings of the IEEE International Conference on Data Mining (ICDM)*, 2015, Atalantic City, NJ, USA

Earth Mover's Distance Based Similarity Joins on Hadoop

PRINCIPAL RESEARCHER

August 2012 - July 2014

- **Jin Huang**, Rui Zhang, Jian Chen, Rajkumar Buyya, and Yongwei Wu, "HEADS-JOIN: Efficient Earth Mover's Distance Join on Hadoop", accepted by *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, July 2015
- **Jin Huang**, Rui Zhang, Rajkumar Buyya, Jian Chen, "MELODY-JOIN: Efficient Earth Mover's Distance Similarity Join Using MapReduce", in *Proceedings of the 30th IEEE International Conference on Data Engineering (ICDE)*, Chicago, IL, USA, 2014

Trajectory Mining and User Mobility Prediction

RESEARCHER

February 2012 - August 2015

- Andy Yuan Xue, Jianzhong Qi, Xing Xie, Rui Zhang, **Jin Huang**, and Yuan Li, "Solving the Data Sparsity Problem in Destination Prediction", in *The Very Large Data Base Journal (VLDBJ)*, August 2014
- Andy Yuan Xue, Rui Zhang, Yu Zheng, Xing Xie, **Jin Huang**, and Zhenghua Xu, "Destination Prediction by Sub-Trajectory Synthesis and Privacy Protection Against Such Prediction", in *Proceedings of the 29th IEEE International Conference on Data Engineering (ICDE)*, Brisbane, QLD, Australia, 2013

Index Structures and Querying Algorithms for Spatial Data

RESEARCHER

February 2011 - August 2015

- Rui Zhang, Jianzhong Qi, Martin Stradling, **Jin Huang**, "Towards a Painless Index for Spatial Objects", in *ACM Transactions on Database Systems (TODS)*, September 2014
- Jian Chen, **Jin Huang**, Zeyi Wen, Zhen He, Kerry Taylor, and Rui Zhang, "Analysis and Evaluation of the Top-k Most Influential Location Selection Query", in *Knowledge and Information Systems (KAIS)*, January 2014
- Yu Sun, **Jin Huang**, Yueguo Chen, Rui Zhang, and Xiaoyong Du, "Location Selection for Utility Maximization with Capacity Constraints", in *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM)*, 2012
- Yu Sun, **Jin Huang**, Yueguo Chen, Xiaoyong Du, and Rui Zhang, "Top-*k* Most Influential Location Selection with Capacity Constraint", in *Proceedings of the 13th International Conference on Web-Age Information Management (WAIM)*, 2012
- **Jin Huang**, Zeyi Wen, Jianzhong Qi, Rui Zhang, Jian Chen, and Zhen He, "Top-*k* Most Influential Location Selection", in *Proceedings of the 20th ACM International Conference on Information and Knowledge Management (CIKM)*, 2011
- **Jin Huang**, Zeyi Wen, Mukaddim Pathan, Kerry Taylor, Yuan Xue, and Rui Zhang, "Ranking Locations for Facility Selection", in *Proceedings of the 27th IEEE Annual Conference of the IEEE Industrial Electronics Society (IECON)*, 2011

Personal

- Permanent resident of Australia with full right to work in Australia