

Song Huang

📞 +1 940-273-0976 • ✉ mailrichardhuang@gmail.com
🌐 <https://shuang-site.herokuapp.com/>

Song Huang currently works as a big data developer in Allstate Insurance Company. He has industrial experience on ETL pipeline, which includes data extracting, data processing and data loading. Also, he has experience on distributed computer systems, data analytics, machine learning and deep learning.

Education

- **University of North Texas** **Denton, TX**
Ph.D. in Computer Science and Engineering, GPA:4.0 2013–2019
- **Texas A&M University-Commerce** **Commerce, TX**
Master in Computer Science, GPA:4.0 2011–2013
- **Guangdong University of Technology** **Guangzhou, China**
Bachelor of Engineering in Network Engineering, 2002–2006

Teaching, Fellowship and Assistantship

- **Adjunct Professor**, Texas A&M University-Commerce
 - CSCI-352 Introduction to Digital Forensics (Fall 2020)
 - CSCI-525 Networking I (Fall 2020)
- **Teaching Fellow**, University of North Texas
 - CSCE 2610: Assembly Language and Computer Organization. (Fall 2018)
- **Teaching Assistant / Research Assistant**, University of North Texas (Fall 2013 – Spring 2019)
 - CSCE 1020: Program Development
 - CSCE 1030: Computer Science I
 - CSCE 2100: Foundations of Computing
 - CSCE 2610: Assembly Language and Computer Organization.
- **Graduate Assistant**, Texas A&M University-Commerce (Fall 2011 - Spring 2013)
 - CSCI 515: Fund Of Programming C/C++ (Mentor: Dr. Daniel Creider)
 - CSCI 520: Data Structures (Mentor: Dr. Daniel Creider)

Research Interest

- Applying Machine Learning on building reliable distributed systems.
- Workload characterization and resource management on large scale computer systems.
- Deep Learning, Reinforcement Learning and Feedback Control

Research Projects and Experience

- **Enhancing Dependability of Storage Systems by Failure Prediction on Cloud Platform**
@University of North Texas
 - Characterize workload and resource utilization on large-scale cloud computing platform.
 - Apply machine learning techniques to distinguish different types of jobs on batch mode.
 - Schedule resources on the distributed system to optimize the application runtime and resource consumption.
- **Enhancing Dependability of Storage Systems by Failure Prediction on Cloud Platform**
@University of North Texas

- Collected S.M.A.R.T. data of Hard Disk Drives using Linux system tools, stored the data into files.
- Applied statistical methods to analyze data and characterize health status of hard disk drives.
- Built machine learning models and predicted disk failures to proactively protect the storage systems.

Failure Analysis and Dependability Enhancement on Software-Defined Networks

- *Internship at Cisco Systems, Inc (Summer 2016);*
 - Designed the project for the failure analysis and prediction on Software-Defined Network in OpenStack.
 - Extracted the data from MongoDB, visualized the network topology of Software-Defined Network.
 - Developed machine learning models to quantify criticality of entities in virtualized network.
 - The predicting results outperform the generic PageRank method developed by Google.

Power and energy efficiency on large-scale computer systems

- *Internship at Los Alamos National Laboratory (Summer, Fall 2015);*
 - Ran benchmarks and collected power and energy data using Linux System tools on HPC servers.
 - Characterized and visualized the power and energy consumption, and predicted the behaviors.
 - Used the prediction results to control runtime settings to optimize power and energy consumption.

Publications

- **Song Huang**, Shuwen Liang, Song Fu, Weisong Shi, Devesh Tiwari, and Hsing-bung Chen, "Characterizing Disk Health Degradation and Proactively Protecting Against Disk Failures for Reliable Storage Systems", submitted to The 16th IEEE International Conference on Automatic Computing (ICAC), June 2019
- **Song Huang**, Song Fu, Scott. Pakin and Michael. Lang, "Characterizing Power and Energy Efficiency of A Data-Centric HPC Runtime and Applications". Published in book: *High Performance Parallel Computing*, November 2018.
- Shuwen Liang, Zhi Qiao, Jacob Hochstetler, **Song Huang**, Song Fu, Weisong Shi, Devesh Tiwari, Hsing-bung Chen, Bradley Settlemyer, and David Montoya, "Reliability Characterization of Solid State Drives in a Scalable Production Datacenter", in Proceedings of IEEE Big Data Conference (BigData) 2018.
- **Song Huang**, "Research on Power Saving and Energy Efficiency for Data-Centric Computing on Production HPC Systems", in Proceedings of *IEEE International Green and Sustainable Computing Conference (IGSC)*, October, 2017
- **Song Huang**, S. Fu, W. Shi and D. Tiwari, "Proactive Disk Failure Management and Data Protection for Highly Available Storage Systems", *ACM Symposium on High-Performance Parallel and Distributed Computing (HPDC), extended abstract*, July 2017
- **Song Huang**, Song Fu, Scott Pakin and Michael Lang, "Characterizing Power and Energy Efficiency of Legion Runtime and Applications: An Early Experience", *IEEE International Green and Sustainable Computing Conference (IGSC)*, November 2016.
- **Song Huang**, Zhiang Deng, Song Fu, "Quantifying Entity Criticality for Fault Impact Analysis and Dependability Enhancement in Software-Defined Networks", *35th IEEE International Performance Computing and Communications Conference(IPCCC). Las Vegas, December 2016.*
- **Song Huang**, Song Fu, Quan Zhang, Weisong Shi, "Characterizing Disk Failures with Quantified Disk Degradation Signatures: An Early Experience", in *Proceedings of IEEE International Symposium on Workload Characterization (IISWC)*, October 2015.
- **Song Huang**, Song Fu, Nathan DeBardleben, Qiang Guan, and Chengzhong Xu, "Differentiated Failure Remediation with Action Selection for Resilient Computing", in *Proceedings of the 21st IEEE/IFIP International Symposium on Dependable Computing (PRDC)*, November 2015.
- **Song Huang**, Michael Lang, Scott Pakin, and Song Fu, "Measurement and Characterization of Haswell Power and Energy Consumption", in *Proceedings of the 3rd International Workshop on Energy Efficient Supercomputing (E2SC '15), in conjunction of IEEE/ACM International Conference for High Performance Computing, Networking, Storage and Analysis (SC)*, November,

2015.

- Xiajun Wang, **Song Huang**, Song Fu and Krishna Kavi, "Characterizing Workload of Web Applications on Virtualized Servers", *Big Data Benchmarks, Performance Optimization, and Emerging Hardware*, pp 98-108, Springer, November 2014.

Services

Technical Program Committee.....

International Conference on Green Communications, Computing and Technologies (GREEN) 2018

International Conference on Green Communications, Computing and Technologies (GREEN) 2019

Peer Reviews.....

TDSC; ICCD'19; IGSC'18; NAS2018; IGSC17; ICCCN 2017; ICPADS 2016; IGSC16; ICPP-2016; SELSE 2016; BodyNets 2015; Bodynets2014

Industrial Work Experience

- **Big Data Developer** **Allstate Insurance Company**
August 2019 – Present *Irving, TX*
 - ETL. Extract data from data warehouses to Hadoop Platform, and transform data into desired format. Load data to different environments for user accesses.
 - Matching. Match data from different sources using IBM Bulk Cross Matching, and Big Match Environment.
 - Data Analysis. Analyze complex data to provide deep knowledge and business insight on the data.
 - Production support. Setup CICD pipeline, load data into different environments, and trouble shootings.
- **IT Team Lead** **Guangdong Century Jiahua Trading Co., Ltd**
July 2008–December 2010 *Guangzhou, China*

Led a team to develop and maintain the network and website publishing system.

 - Coordinated software development team to develop a website publishing system.
 - Specified the requirements, designed the software, conducted integrated testing and deployment.
- **Software Developer** **Guangzhou TWO Information Technology Co., Ltd**
July 2006–July 2008 *Guangzhou, China*

Developed ERP software for construction management.

 - Developed Enterprise Resource Planning (ERP) software for release (iTWO).
 - Requirement analysis (communicated with clients), software design (High level and low level).
 - Software implementation (coding and documentation), and testing (unit and integrated testing).

Activities and Awards

- IEEE Travel Grant for the 16th IEEE International Conference on Autonomic Computing(ICAC'2019)
- ACM Travel Grant for High-Performance Parallel and Distributed Computing (HPDC'2017)
- Student Volunteer at IEEE/ACM Super Computing conference (SC'2015)
- College of Engineering Travel Grant, UNT, November 2015 / December 2016 / June 2017
- College of Engineering Graduate Student Scholarship, UNT, 2014 – 2015
- Toulouse Graduate School: Graduate Assistantship Teaching Scholarship, UNT, 2013-2014
- CyberQ Consulting Company: CMM3 Completion Certificate, March 2008
- Guangdong University of Technology: Outstanding Study Performance / Outstanding Student, 2002-2005