Curriculum Vita

Yi Zhou

September 19, 2022

Address:

Center for Commerce & Technology, Suite 435 TSYS School of Computer Science 4225 University Ave, Columbus, GA 31907

Phone: (706) 507-8048

E-Mail: zhou yi@columbusstate.edu or yzhou056@gmail.com

1. Personal Information

1.1 Research Interests

Parallel/Distributed Systems, Green computing, Storage Systems, Cloud Computing, Database Systems, Data science, and Dynamic Resource Management, Undergraduate STEM Education

1.2 Education GPA:3.91/4.00

• Ph.D. in Computer Science & Software Engineering, Auburn University, AL, 2018

• M.S. in Computer Software Engineering,

Auburn University, AL, 2018

- M.E. in Microelectronics and Solid-State Electronics, Beijing University of Technology, China, 2010
- **B.E. in Electronic Science and Technology,** Beijing University of Technology, China, 2006

1.3 Employment

2018 - Present: Columbus State University

Assistant professor, TSYS school of Computer Science.

04/2011 - 10/2013: Alcatel-Lucent R&D

Senior software engineer, Department of Network Performance Optimizer System.

07/2010 - 04/2011: SOHU R&D

Android software engineer, Department of SOHU Video Mobile.

1.4 Honors and Awards

- Outstanding Doctoral Student Award at Auburn University 2018. One of the ten award winners at Auburn University Graduate School.
- Merriwether Fellowship (Nomination Award) 2017. One CSSE graduate student was nominated for the Merriwether fellowship.
- Alley Family Graduate Student Leadership Fellowship (AFGS Nomination Award) 2017. One of the seven nominated engineering graduate students.
- Alabama EPSCoR Fellowship (Nomination Award) 2016. One CSSE graduate student was nominated for the EPSCoR fellowship.

2. Research, Scholarly, and Creative Activities

2.1 Research Experience

08/2018 – present, Columbus State University

Assistant Professor, TSYS School of Computer Science.

<u>Improve the Performance and Energy Efficiency of Big-data Computing and Cloud Computing</u> Systems

- Novel Container Placement Strategies to Improve Performance and Energy-efficiency of Containerized Data Centers:
 - We proposed a container placement approach *-Blender* -by taking into account the Zipf-like distribution to reduce network traffic and improve the performance of containerized data centers.
 - We proposed a heat-recirculation-aware Virtual Machine placement (VMP) strategy for reducing the energy consumption of data centers. This novel VMP strategy considers heat recirculation coupled with multiple physical resource allocation to reduce the energy consumption of data centers.
 - We developed a structure-based Storm scheduler called TOSS, which is capable of accelerating the performance of Storm clusters by shortening process latency.
- Deduplication and Caching Techniques to Boost Performance of Large-scale Data Storage Systems:
 - We designed a container utilization based hot fingerprint entry distilling strategy to improve the performance of deduplication-based backup systems.
 - We proposed a Classified-Metadata based restoring method (*CMR*) that classifies backup metadata into file metadata and chunk metadata, aiming to boost the restoring performance of deduplication system.
 - We developed a caching method for identifying and managing hot data by leveraging resource utilization efficiency for dynamic large-scale data streams.
- Towards Green Hadoop Clusters in Data Centers:
 - We proposed an energy-efficient database system called *GreenDB* running on clusters, which applies a workload-skewness strategy by managing hot nodes coupled with a set of cold nodes in a database cluster.
 - We proposed a new resource- and thermal-aware scheduler in Hadoop clusters, aiming at minimizing peak inlet temperature across all nodes to reduce power consumption and cooling cost in data centers.
 - We investigated the thermal behavior of a MapReduce application called Pi running on Hadoop clusters by varying two input parameters number of maps and number of sampling points per map.
 - We proposed a self-driven energy-efficient-aware data placement strategy on clusters and developed a prediction model to place data in an energy-efficient manner.
 - We proposed *ThermoBench* to evaluate the thermal efficiency of computing and storage clusters deployed in data centers.

Deduplication and Caching Techniques for Storage Systems

• Deduplication and Caching Techniques to Boost Performance of Large-scale Data Storage Systems:

- We designed a container utilization based hot fingerprint entry distilling strategy to improve the performance of deduplication-based backup systems.
- We proposed a Classified-Metadata based restoring method (CMR) that classifies backup metadata into file metadata and chunk metadata, aiming to boost the restoring performance of deduplication system.
- We developed a caching method for identifying and managing hot data by leveraging resource utilization efficiency for dynamic large-scale data streams.

Authentication Detection Mechanism for Edge Computing Platforms

• Energy-based abnormal IT authorization and authentication detection mechanism: we proposed a model of energy consumption of IT authorization and authentication operations and designed a detection algorithm to identify abnormal IT authorization and authentication.

08/2014 – 05/2018, Auburn University

Graduate Teaching Assistant, Department of Computer Science and Software Engineering

Energy-efficiency Modeling and Management of Database System

- An Asynchronized Prefetching Mechanism: we designed an asynchronous replication method and a pre-fetching strategy to reduce energy cost of DBMS.
- A Thermal-aware Scheduler: we proposed a novel resource-aware and thermal-aware scheduler to reduce peak inlet temperatures across all nodes.
- Modeling CPU/Disk Utilization and Heat Re-Circulation: we developed a thermal model to project thermal behaviors and evaluate heat re-circulation effect in data centers.

09/2017 – 07/2010, Beijing University of Technology, China

Reaching Assistant, Department of Electronic Science and technology

FPGA-based wireless communication systems

- RFID wireless communication system based on FPGA: we implemented a RFID wireless communication system.
- A Serial communication system interface module: we designed and tested an interface module for FPGA-based serial communication systems.

2.2 Publications

Book and Book Chapters

[1] Ahmed Mostafa, **Yi Zhou**, and Rania Hodhod, "Improving Energy-Efficiency through Smart Data Placement in Hadoop Clusters" (18676-67233), *Book chapter in "Advances in Computer Science Research"*. Publisher: Nova Science Publishers, October 2020.

Journal Articles and Conference Papers

- [2] Yuan, Jingling, Yao Xiang, Yuhui Deng, Yi Zhou, and Geyong Min. "UPOA: A User Preference Based Latency and Energy Aware Intelligent Offloading Approach for Cloud-edge Systems." *IEEE Transactions on Cloud Computing* (2022).
- [3] Si, Lei, Shujie Pang, Yuhui Deng, Weiheng Zhu, Yi Zhou, and Yifeng Zhu. "DM-Pages: Improving energy efficiency of disk storage systems and cache performance using deduplication-based mixed pages." *Journal of Circuits, Systems and Computers* (2022).
- [4] Li, Jie, Yuhui Deng, Yi Zhou, Zhen Zhang, Geyong Min, and Xiao Qin. "Towards Thermal-Aware Workload Distribution in Cloud Data Centers Based on Failure Models." *IEEE Transactions on Computers* (2022).
- [5] Lin, Lifang, Yuhui Deng, and Yi Zhou. "Improving Restore Performance of Deduplication Systems via a Greedy Rewriting Scheme." *In 2021 IEEE 27th International Conference on Parallel and Distributed Systems (ICPADS)*, pp. 291-298. IEEE, 2021.

- [6] Wu, Zhaorui, Yuhui Deng, Hao Feng, Yi Zhou, Geyong Min, and Zhen Zhang. "Blender: A Container Placement Strategy by Leveraging Zipf-like Distribution within Containerized Data Centers." IEEE Transactions on Network and Service Management (2021).
- [7] Ai, Liang, Yuhui Deng, Yi Zhou, and Hao Feng. "RUE: A caching method for identifying and managing hot data by leveraging resource utilization efficiency." Software: Practice and Experience 51, no. 11 (2021): 2252-2273.
- [8] Zhang, Datong, Yuhui Deng, Yi Zhou, Yifeng Zhu, and Xiao Qin. "Improving the Performance of Deduplication-Based Backup Systems via Container Utilization Based Hot Fingerprint Entry Distilling." ACM Transactions on Storage (TOS) 17, no. 4 (2021): 1-23.
- [9] Feng, Hao, Yuhui Deng, Yi Zhou, and Geyong Min. "Towards Heat-Recirculation-Aware Virtual Machine Placement in Data Centers." *IEEE Transactions on Network and Service Management* (2021).
- [10] Yang, Ru, Yuhui Deng, Yi Zhou, and Ping Huang. "Boosting the Restoring Performance of Deduplication Data by Classifying Backup Metadata." ACM/IMS Transactions on Data Science 2, no. 2 (2021): 1-16.
- [11] Wu, Zhaorui, Yuhui Deng, Hao Feng, Yi Zhou, and Geyong Min. "Blender: A Traffic-Aware Container Placement for Containerized Data Centers." In 2021 *Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 986-989. IEEE, 2021.
- [12] Li, Yaojie, **Yi Zhou**, Thomas Stafford, and Xuan Wang. "Significant Stakeholders: Toward an Agile Knowledge Management System in the Time of Coronavirus Crisis." *IEEE Engineering Management Review* 49, no. 1 (2020): 38-49.
- [13] **Yi Zhou**, Yuanqi Chen, Shubbhi Taneja, Ajit Chavan, Xiao Qin and Jifu Zhang, "ThermoBench: A Thermal Efficiency Benchmark for Clusters in Data Centers," *Parallel Computing*, August 2020.
- [14] **Yi Zhou**, Yangyang Liu, Chaowei Zhang, Xiaopu Peng, Xiao Qin, "TOSS: A Topology-based Scheduler for Storm Clusters," *The 34th Proceedings of IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, May 2020.
- [15] H. Feng, Y. Deng, and Y. Zhou, "A Heat-Recirculation-Aware VM Placement Strategy for Data Centers," *The 23rd Design, Automation and Test in Europe Conference (DATE 2020)*, 2020.
- [16] **Yi Zhou**, Yuanqi Chen, Chaowei Zhang, Xiao Qin and Jifu Zhang, "Thermal-Efficiency Benchmark on High-Performance Clusters," *Proceedings of IEEE the 10th International Green and sustainable computing Conference*, 2019.
- [17] S. Taneja, Y. Zhou*, A. Chavan, X. Qin, "Improving Energy Efficiency of Hadoop Clusters using Approximate Computing," *Proceedings of IEEE Intl Conference on High Performance and Smart Computing*, 2019.
- [18] Y. Zhou, S. Taneja, C.-W. Zhang, X. Qin, "GreenDB: Energy-Efficient Prefetching and Caching in Database Clusters," *IEEE Transactions on Parallel and Distributed Systems*, Oct 2018.
- [19] Y. Zhou, S. Taneja, G. Dudeja, X. Qin, J.-F. Zhang, M.-H. Jiang and M.-I. Alghamdi, "Towards Thermal-Aware Hadoop Clusters," *Future Generation Computer Systems*. Vol. 88, 2018.
- [20] Y. Zhou, M. Alghamdiy, S. Taneja, A. Chavan and X. Qin, "Improving Energy Efficiency of Database Clusters through Prefetching and Caching," *IEEE International Symposium on Cluster Computing and the Grid*. May 2018.
- [21] Y.-Q. Chen, Y. Zhou*, S. Taneja, X. Qin, J.-Z Huang, "aHDFS: An Erasure-Coded Data Archival System for Hadoop Clusters," *IEEE Transactions on Parallel and Distributed Systems*, Vol. 28, 2017.
- [22] Y. Zhou, M. Alghamdiy, G. Dudeja, S. Taneja, and X. Qin, "Towards Thermal-Efficient Hadoop Clusters through Scheduling," *Proceedings of the 13th IEEE International Conference on Green Computing and Communications*, 2017.
- [23] Y. Zhou, S. Taneja, X. Qin, W.-S. Ku and J.F. Zhang, "EDOM: Improving Energy Efficiency of Database Operations on Multicore Servers," Future Generation Computer Systems, SI: Energy for CPC Systems, 2017.

- [24] Y. Zhou, M. Alghamdiy, S. Taneja, W.-S. Ku and X. Qin, "Towards Energy-Efficient Multicore Database Systems," *Proceedings of the Seventh International Green and Sustainable Computing Conference*, 2016.
- [25] Y. Zhou, C. Zhu, W. Wu, L. Dong, "Implementation of RFID wireless communication system based on FPGA," *Modern Electronic Technology*, Vol.17, February 2010.
- [26] Y. Zhou, L.-G. Hou, W. Wu, "Design and optimization of Serial communication system interface module," 2nd International Conference on Computer Engineering and Technology, April 2010.
- [27] S. Taneja, Y. Zhou*, X. Qin: "Thermal Benchmarking and Modeling for HPC using Big Data Applications," Future Generation Computing Systems (FGCS), Vol. 87 October 2018.
- [28] S. Taneja, Y. Zhou, X. Qin, M. Alghamdi, "Thermal-Aware Job Scheduling of MapReduce Applications on High Performance Clusters," *Proceedings of the 46th International Conference on Parallel Processing Workshops* (ICPPW), 2017.
- [29] S. Taneja, A. Chavan, Y. Zhou*, M. Alghamdiy, X. Qin, "Thermal Profiling and Modeling of Hadoop Clusters using BigData Applications," *IEEE BigDF Workshop*, August 2017
- [30] S. Taneja, S. Kulkarni, Y. Zhou*, X. Qin, "Thermal-aware task assignments in high performance computing clusters," *Proceedings of the Concurrency and Computation: Practice and Experience*, Vol. 29, No.18, 2017.
- [31] Y.-Y. Liu, M. Alghamdi, W.-S. Ku, **Yi Zhou**, S. Taneja, and X. Qin. "Profiling Energy Usage of Web-Service Applications on Clusters," *Proceedings of the IEEE International Conference on* Networking, Architecture and Storage (NAS), 2016.
- [32] X. Kong, W. Wu, L.-G. Hou, Y. Zhou, "Verilog design and optimization of FSMs," *Microeletronic and Computer*, Vol. 27, No. 2, February 2010.

Google Scholar: https://scholar.google.com/citations?user=YihIJiUAAAAJ&hl=en

2.3 Grant Experiences

1. (PI) CSU, Interdisciplinary Initiative Grant Application

PI: Yi Zhou; Co-PIs: Yi Zhou, Kristin Lilly, and Rania Hodhod

Smart Course Advising Tool: A Smart Class Planning Tool for Academic Advising,

Duration:8/2022 – 8/2023 Award Amount: \$3,000.00

2. (Co-PI) University System of Georgia, Affordable Learning Georgia Round 20

PI: Linqiang Ge; Co-PIs: Yi Zhou, Anastasia Angelopoulou, Suk Lee, and Rania Hodhod

"Design computer science courses that offer affordable open educational resources (OER) material and provide cost savings of \$137.11 per student in five computer science courses",

Duration: 12/2021 – 12/2022 Award Amount: \$25,000.00

3. (Co-PI) NCWIT Extension Services (ES):

PI: Radhouane Chouchane Co-PIs: Shamim Khan, Hillary Fleenor, Alfredo Perez, Lydia Ray, and Hyrum Carroll.

"To implementing systemic change in our undergraduate programs in ways that increase enrollment, retention, and graduation of women"

Award Amount: \$9,540

4. STEM Initiative Mini Grants Proposal, Columbus State University

"Development of Hands-on Activities in Big-data Computing"

5. NSF proposal participant: Training computer science and mathematics undergraduates from Alabama and nearby regions in data science 2019

"To build courses into which Data Science related curricular content and assignments/projects can be incorporated."

Not Funded

3. Teaching, Mentoring, and Advising Activities

3.1 Teaching Interest

Operating Systems, Big-data Computing, Machine Learning & AI, Neural Networks, Database Systems, Distributed Computing, Computer Architecture, Software Engineering, Software Design, Data Structures

3.2 Teaching Experience

Columbus State University

Academic Terr	m, Course, Enroll	ment		
2022 Fall,	CPSC 6985 Research and Thesis: Towards energy-efficient edge computing for	or tiny AI		
applications		1		
2022 Fall,	CPSC 1105 Introduction to Computing Principles and Technology	28		
2022 Fall,	CPSC 1301K Computer Science I	30		
2022 Fall,	CPSC 4175 Software Engineering	25		
2022 Fall,	CPSC 6179 Software Project Planning and Management, sections 01 and V01	9,7		
2022 Summer,	CPSC 6985 Research and Thesis	1		
2022 Summer,	CPSC 1302 Computer Science II	10		
2022 Spring,	CPSC 6127 Contemporary Issues in Database Management Systems, sections 01	and V01		
		8,7		
2022 Spring,	CPSC 4176 Senior Software Engineering Project	26		
2022 Spring,	CPSC 3125 Operating Systems	13		
2021 Fall,	CPSC 6125 Operating System Design and Implementation, sections 01 and V01	2,4		
2021 Fall,	CPSC 3125 Operating Systems	20		
2021 Fall,	CPSC 4175 Software Engineering	30		
2021 Summer,	CPSC 6106 Fundamentals of Computer Programming and Data Structures	10		
2021 Spring,	CPSC 3125 Operating Systems	26		
2021 Spring,	CPSC 4176 Senior Software Engineering Project	27		
2021 Spring,	CPSC 6127 Contemporary Issues in Database Management Systems, sections 01	and V01		
		2,14		
2020 Fall,	CPSC 3125 Operating Systems	20		
2020 Fall,	CPSC 4175 Software Engineering	30		
2020 Fall,	CPSC 6119 Object-Oriented Development, sections 01 and V01	3,7		
2020 Fall,	CPSC 4505 Undergraduate Research: Investigating the impact of virtual	machine		
placement on energy consumption of a data center 1				
2020 Summer,	CPSC 6106 Fundamentals of Computer Programming and Data Structures	9		

2020 Spring,	CPSC 4176 Senior Software Engineering Project	25
2020 Spring,	CPSC 6127 Contemporary Issues in Database Management Systems, sections	$01 \ and \ V01$
		5,11
2020 Spring,	CPSC 6175 Web Engineering and Technologies, sections 01 and V01	2,10
2020 Spring,	CPSC 6985 Research and Thesis	1
2020 Spring,	CPSC 6986 Thesis Defense	1
2019 Fall,	CPSC 3125 Operating Systems	21
2019 Fall,	CPSC 4175 Software Engineering	30
2019 Fall,	CPSC 6177 Advanced Software Design, sections 01 and V01	3,11
2019 Fall,	CPSC 6985 Research and Thesis	1
2019 Summer,	CPSC 1301K Computer Science I	17
2019 Spring,	CPSC 4176 Senior Software Engineering Project	26
2019 Spring,	CPSC 6127 Contemporary Issues in Database Management Systems, sections	$01 \ and \ V01$
		14,8
2019 Spring,	CPSC 6179 Software Project Planning and Management, sections 01 and V01	1,9
2018 Fall,	CPSC 1301 Computer Science I, sections 4,5, and P13.	28,10,17
2018 Fall,	CPSC 1301L Computer Science I Lab, sections 4,5 and P13	27,10,17
2018 Fall,	CPSC 4175 Software Engineering	31

Auburn University

2014 - 2017, Instructor of COMP1000 - Personal Computer Application

- Teaching evaluation: 5/6

2015 - 2017, Instructor of K-12 Robotics & Game Development Camp

2016 Spring, Course Coordinator of COMP1000 - Personal Computer Application

- Served as the course coordinator, managed a group of 18 teaching sessions.

2018 Spring, Graduate Teaching Assistance of COMP4710-Senior Design

2016 Spring, Graduate Teaching Assistance of COMP5000/6000 - Web Application Development

3.3 Course or Curriculum Development

Thanks to my software engineering experience in Fortune Global 500 and my research background in big data computing and cloud computing, energy-efficient computing, and data science, I have taught a wide variety of classes of topics, ranging from low-level undergraduate classes to high-level graduate classes since I joined CSU in Fall2018. As of now, I have taught **sixteen (16)** different courses. Moreover, to improve comprehension of students' learning. I have fully re-designed **six (6)** courses by incorporating both cutting-edge techniques and novel pedagogies. More specifically, cutting-edge techniques such as distributed computing, data science have been introduced into these courses. Besides, novel pedagogies like scaffolding strategy, small-team-based teaching, problem-solving strategy have been applied. In addition, to sharpen students' programming skills and software engineering skills, I have brought advanced software engineering topics and real-world skills into computer science courses, students who graduated have provided positive feedback on such topics. Furthermore, I have conducted a wide range of research and teaching collaborations with both internal and external faculties and researchers.

Fully Redesigned Courses:

1. CPSC3125 Operating System

2.	CPSC4175 Software Engineering	3 cr, 3 cl hrs
3.	CPSC4176 Senior Software Engineering	3 cr, 3 cl hrs
4.	CPSC6125 Operating System Design and Implementation	3 cr, 3 cl hrs
5.	CPSC6127 Contemporary Issues in Database Management Systems	3 cr, 3 cl hrs
6.	CPSC6179 Software Project Planning and Management	3 cr, 3 cl hrs

3.4 Advising

3.4.1 Academic Advising for Undergraduate Students

Academic Year,	Number of Undergraduate Students Advised
2022-	50
2021-2022	41
2020-2021	59
2019-2020	66
2018-2019	41

3.4.2 Advising for Graduate Students

I have graduated 1 master's student (thesis option). Currently, I am supervising a master student. And I am co-supervising 2 Ph.D. students, in collaboration with Dr.Xiao Qin from Auburn University.

3.4.3 M.S. Theses and Projects

Vamsi Krishna [Fall 2022; Thesis Option; co-advised with Drs. Shamim Khan and Suk Lee]

Project: Towards energy-efficient edge computing for tiny AI applications

In this project, we energy-aware scheduling module that predicts the energy consumption of AI algorithms running on a Raspberry Pi and thus schedules AI algorithms based on their energy behaviors.

Ahmed Mostafa [Spring 2020; Thesis Option; co-advised with Drs. Shamim Khan and Rania Hodhod] **Project: Improving Energy-Efficient through Smart Data Placement in Hadoop Clusters**

In this project, we introduce an energy-aware Hadoop model that estimates the energy efficiency of MapReduce jobs from job types and Hadoop clusters' characteristics (e.g., CPU and memory utilization). A node manager is seamlessly integrated to the energy-aware Hadoop aiming to dynamically manage nodes within the cluster.

4. Services

4.1 University and Departmental Services

Columbus State University

- 2021, 2022: Faculty Development Committee
- 2019, 2020: Judge, Tower Day
- 2019: Moderator, Day of Service
- 2019: Moderator, NSF Workshop on Interdisciplinary and Collaborative Research in Cybersecurity at CSU
- 2018: Moderator, Faculty Research Conference

TSYS School of Computer Science

• 2022: Student Scholarship/Honors Committee

- 2021: Discovery Day Volunteer
- 2019, 2020 The National Center for Women & Information Technology (NCWIT Learning Circle)
- 2018-Present: Graduate Programs Committee
- 2019: USG Chancellor's faculty cohort workshop
- 2019, 2020: Moderator, Faculty Research Conference
- 2019: Chancellor's Learning Scholars Faculty Learning Communities (FLC)
- 2019: LSAMP summer camp in the robot workshop
- 2018-2019: Coordinator, CPSC1301 Group
- 2018: Peer Review of Teaching Form Revising Committee
- 2018: CS Faculty Search Committee

D. Abbott Turner College of Business

- 2022: Assurance of Learning (AOL) Committee
- 2018-2021: Faculty Resource and Development Committee (FRDC)
- 2018: Assurance of Learning (AOL) Committee

Auburn University

• 2016, 2017 E-Day Volunteer, Samuel Ginn College of Engineering, Auburn University

4.2 Professional Service

- 2022: Reviewer, Journal of Parallel and Distributed Computing
- 2021: Reviewer, ICDCS'21: 41st IEEE International Conference on Distributed Computing Systems
- 2021: Program Committee, 23rd Workshop on Advances in Parallel and Distributed Computational Models (APDCM 2021)
- 2020: Session chair of Advanced Technologies in Big Data Systems, The 13th IEEE International Conference on Cyber, Physical and Social Computing (CPSCom 2020)
- 2020: Reviewer, IEEE Access
- 2019: Reviewer, IEEE Transactions on Industrial Informatics
- 2017: Reviewer, IEEE Transactions on Systems, Man and Cybernetics: Systems
- 2017: Reviewer, Journal of Information Science and Engineering
- 2016: Reviewer, Journal of Future Generation Computing Systems.
- 2017: Reviewer, IEEE International Symposium on Cluster Computing and the Grid
- 2016: Reviewer, Int'l Conf. Parallel Computing.

4.3 Other Services

- 2020-Present: Committee Member, Auburn Chinese Bible Study Group.
- 2016-2018: Leader, Auburn Chinese Student Christian Fellowship.
- 2015: Mission of Carpenter for Christ, Moberly, Missouri