

Zheguang Zhao

Brown University
Box 1910, Department of Computer Science
115 Waterman St
Providence, RI 02912
United States of America

Phone: +1-608-630-1677
Email: zheguang.zhao@gmail.com
Homepage: zheguang.github.io
LinkedIn: www.linkedin.com/in/zheguang
Google Scholar: goo.gl/DR8pSa

Education

Ph.D. Candidate in Computer Science, Brown University, Rhode Island, expected 2020.

Advisor: Prof. Seny Kamara and Stan Zdonik.

Thesis: *Relational Database Encryption*

M.S. in Computer Science, Brown University, Rhode Island, 2016.

Advisor: Prof. Stan Zdonik.

Thesis: *Approximate Data Structures for Visualization*

B.S. in Computer Science, University of Wisconsin at Madison, Wisconsin, 2012.

Advisor: Prof. Jignesh Patel

Certification

Deep Learning Specialization, *Coursera / deeplearning.ai*

Neural Networks and Deep Learning

Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization

Structuring Machine Learning Projects

Convolutional Neural Networks

Sequence Models

Technology

ML: Tensorflow, Keras; JuMP

Systems: Apache Spark, Hadoop; PostgreSQL

PL: Scala; Java, C#; C; Python, Julia

Infra: Amazon EC2; Jenkins; Git; Jira; Docker

Honors

Eta Kappa Nu

Upsilon Pi Epsilon

Golden Key International Honour Society

Experiences

Los Alamos National Laboratory, 2019 – present.

Machine learning model reconstruction of mixed dynamics in cyber-physical systems, with application to network verification and security.

Aroki Systems, 2018 – present.

Provably-secure end-to-end database encryption for PostgreSQL and Apache Spark.

Microsoft AI & Research, 2017.

Constraint learning for puzzle-solving AI.

Intel Labs, 2015.

Efficiency of machine learning algorithms in Apache Spark.

In-memory transactional processing using non-volatile memory.

Hadapt (Acquired by Teradata), 2013.

Enterprise SQL-on-Hadoop system including query execution, storage engine, high availability and analytics.

Kosmix (Acquired by @WalmartLabs), 2012.

In-memory distributed queue system for the in-house stream processing.

Great Lakes Bioenergy Research Center, 2010.

Scientific database for biological enzyme research.

Open-Source Software¹

ML framework for Cyber-physical Systems

Encrypted Spark

Encrypted Searchable Signal

Macau: statistical hypothesis testing based on resampling

Machine learning algorithms in Spark

Consistency control for machine learning algorithms

R-tree in Rust

Spark performance analysis tool

VoltDB on non-volatile memory

¹All hosted on Github: <https://github.com/zheguang>

Publication

An Optimal Relational Database Encryption Scheme

Cryptology ePrint Archive: Report 2020/274.

Dynamic Query Refinement for Interactive Data Exploration

EDBT/ICDT Joint Conference, March 2020.

Behavior of Large Random Graph.

Randomized Algorithms for Counting, Integration and Optimization, Brown University, April 2017.

Investigating the Effect of the Multiple Comparisons Problem in Visual Analysis.

CHI Conference, April 2018.

Signal Search.

Press release: <http://esl.cs.brown.edu/blog/signal>, April 2017.

Controlling False Discoveries During Interactive Data Exploration.

SIGMOD Conference, May 2017.

Safe Visual Data Exploration.

SIGMOD Conference, Demo, May 2017.

Bridging the Gap between HPC and Big Data frameworks.

VLDB Journal, 2017.

Towards Sustainable Insights.

CIDR Conference, January 2017.

Towards a Benchmark for Interactive Data Exploration.

IEEE Data Engineering Bulletin, 2016.

Larger-than-memory Data Management on Modern Storage Hardware for In-memory OLTP Database Systems.

SIGMOD DaMoN Workshop, June 2016.

VisTrees: Fast Indexes for Interactive Data Exploration.

SIGMOD HILDA Workshop, June 2016.

Data Tiering in Heterogeneous Memory Systems.

EuroSys Conference, April 2016.