Jinzhen Wang

+1-551-580-4404 jw447@njit.edu jinzhen-wang-54b6b0a8 jw447.github.io Jersey City, NJ

I'm a senior Ph.D. candidate in Computer Engineering, focusing on data compression and HPC data management. I'm a motivated and result-driven individual that can work effectively under limited supervision. I have experience working in a collaborative environment and am responsible for leading and managing research projects. I'm also proficient in writing and presentation at all levels.

Skills

Programming languages	Python, C, C++
Software & tools	ADIOS, HDF5, Foresight, Git, LaTeX, MarkDown
Communication	English (fluent), Mandarin (native)

Experience

Graduate research intern **Sep 2022 — now**
Los Alamos National Laboratory *Los Alamos, New Mexico, US*
Design a user study on the feedback on image quality of volume rendering with different compression techniques, and correlate user feedback with image quality assessment.

Summer research intern **May 2022 — Aug 2022**
Los Alamos National Laboratory *Los Alamos, New Mexico, US*
Perform research on the effect of error-bounded lossy data compression on data visualization for large-scale scientific datasets.

Education

Ph.D. in Computer Engineering, *New Jersey Institute of Technology, NJ, US* Jan 2018 - Present

M.S. in Electrical Engineering, *New Jersey Institute of Technology, NJ, US* Sep 2015 - May 2017

B.S. in Internet of Things (Electrical Engineering), *Shandong University, China* Sep 2011 - May 2015

Research

Analyzing the Impact of Lossy Data Reduction on Volume Rendering of Cosmology Data
Motivation — As HPC scientific users seek to reduce the amount of storage needed for application data while preserving enough quality for analysis and visualization. The key question is how much we can reduce the data before noticeable artifacts are introduced.

- Perform a thorough analysis of the effects of 3 cutting-edge error-bounded lossy compression techniques on large-scale scientific data;
- Investigate more than 20 image quality assessment metrics on quantifying visualization artifacts for volume rendering.

High-Ratio Lossy Compression: Exploring the Auto-encoder to Compress Scientific Data
Motivation — Auto-encoder has been widely used in image compression and dimension reduction and achieving superior performance, motivating the adoption of scientific data compression.

- Propose and design the first error-bounded auto-encoder compression framework that achieves superior reduction performance over traditional scientific lossy compression techniques.
- Optimize the auto-encoder compression framework via data pre-processing techniques and back-end lossless encoding techniques.
- Evaluate the compression framework over 20 large-scale scientific datasets (100+ GB) from more than 10 disciplines to show the generality.

Jinzhen Wang

+1-551-580-4404 jw447@njit.edu jinzhen-wang-54b6b0a8 jw447.github.io Jersey City, NJ

Estimating Lossy Compressibility of Scientific Data Using Deep Neural Networks
Motivation — The expected lossy compression performance is hard to obtain, making it hard for users to make decisions on the trade-off between data accuracy and compression performance.

- Characterize the compression performance under various configurations using deep learning.
- Propose a deep-learning-based framework to predict the compressibility of scientific data.
- Significantly improve the prediction accuracy by 30% leveraging compressor-dependent features.
- Evaluate the proposed work against the sampling-based and analytical model that was used to estimate the compression performance.

Publications

- **Wang, J.**, Liu, T., Liu, Q., He, X., Luo, H., & He, W. (2019). Compression ratio modeling and estimation across error bounds for lossy compression. IEEE Transactions on Parallel and Distributed Systems, 31(7), 1621-1635.

- Liu, T., **Wang, J.**, Liu, Q., Alibhai, S., Lu, T., & He, X. (2021). High-ratio lossy compression: Exploring the autoencoder to compress scientific data. IEEE Transactions on Big Data.

- Qin, Z., **Wang, J.**, Liu, Q., Chen, J., Pugmire, D., Podhorszki, N., & Klasky, S. (2020). Estimating Lossy Compressibility of Scientific Data Using Deep Neural Networks. IEEE Letters of the Computer Society, 3(1), 5-8.

- Liu, T., Alibhai, S., **Wang, J.**, Liu, Q., He, X., & Wu, C. (2019, August). Exploring transfer learning to reduce training overhead of hpc data in machine learning. In 2019 IEEE International Conference on Networking, Architecture and Storage (NAS) (pp. 1-7). IEEE.

- Qiao, Z., Lu, T., Luo, H., Liu, Q., Klasky, S., Podhorszki, N., & **Wang, J.** (2018). Sirius: Enabling progressive data exploration for extreme-scale scientific data. IEEE Transactions on Multi-Scale Computing Systems, 4(4), 900-913.

Activities

Oral presentation: IPDPS2019, DRBSD-8 (SC22)
Reviewer: DRBSD-5, DRBSD-6, ICDE2022