Dong Hu

Rensselaer Polytechnic Institute
hud3@rpi.edu
https://jurohd.github.io

INTERESTS Matrix Completion, Sketching, Machine Learning, High-dimensional Statistics, Applied Linear Algebra, Randomized Algorithms.

EDUCATION

Rensselaer Polytechnic Institute Summer 2019 – Spring 2024 (expected)
Ph.D. in Computer Science Advisor: *Prof.* Alex Gittens
GPA: 3.9/4.0

Rensselaer Polytechnic Institute Spring 2016 – Spring 2019
Bachelor of Science in Mathematics Advisor: *Prof.* Jeffery Banks
Bachelor of Science in Computer Science Advisor: *Prof.* Heng Ji
GPA: 3.86/4.0, *Dean's Honor List*

RESEARCH EXPERIENCE

Rensselaer Polytechnic Institute Summer 2019-present
Graduate Research Assistant, Computer Science Department

- Proposed a proximal regularized sketched alternating least squares (Tucker-ALS) algorithm for low rank Tucker decomposition of large tensors, and prove that a sublinear rate of convergence of proximally regularized sketched CPD algorithms also holds for the proposed algorithm. Showed that the iterative nature of the Tucker-ALS approach can be exploited algorithmically to choose more performant sketching rates at different iterations to improve the convergence of the overall algorithm.

- Proposed and investigated in a regression-based matrix completion algorithm (noisyCUR) for low budget matrix completion setting and experimentally verify the performance of our algorithm on both synthetic and real data, compared our algorithm with state-of-the-art Matrix completion algorithms. Paper accepted by ECML 2020.

IBM, Yorktown Heights Summer 2020-present
Artificial Intelligence Research Collaboration(AIRC) scholar

- Investigated sketching matrices that are obtained from bipartite graphs that are sparse, and explored two popular classes of sparse graphs, expander graphs and magical graphs. Proved that for a subspace with dimention k, the magical graph with left degree $s = 2$ yields a $(1 \pm \varepsilon)$ ℓ_2-subspace embedding if the number of right vertices (the sketch size) $m = \mathcal{O}(k^2/\varepsilon^2)$, and the expander graph with $s = \mathcal{O}(\log k/\varepsilon)$ yields a subspace embedding for $m = \mathcal{O}(k \log k/\varepsilon^2)$.

PUBLICATIONS

D. Hu, S. Ubaru, A. Gittens, K. Clarkson, L. Horesh, and V. Kalantzis. "Sparse graph based sketching for fast numerical linear algebra." in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP),* 2021.

D. Hu, A. Gittens, and M. Magdon-Ismail, "NoisyCUR: An algorithm for two-cost budgeted matrix completion," in *Machine Learning and Knowledge Discovery in Databases - European Conference(ECML-PKDD),* 2020