Ariful Azad
Curriculum Vitae

September 2021

 Department of Intelligent Systems Engineering, Indiana University, Bloomington, IN, USA.
 arifulazad.com
 azad@iu.edu
 hipgraph

Education

2006	B.Sc. in Computer Science & Engineering	Bangladesh University of Engineering and Technology
2014	Ph.D. in Computer Science	Purdue University

Positions held

2018–	**Assistant Professor**, Department of Intelligent Systems Engineering, Indiana University, Bloomington.
2020–	**Director of Graduate Studies**, Department of Intelligent Systems Engineering, Indiana University, Bloomington.
2018–	**Affiliate Research Scientist**, Lawrence Berkeley National Laboratory, Berkeley, CA
2016–2018	**Research Scientist**, Lawrence Berkeley National Laboratory, Berkeley, CA

Grants

2021–2026	**Intelligent Primitives for Scalable Graph Analytics and Learning**, DOE Early Career Award	$752,000
2018–2022	**Research and development of high-performance Markov clustering**, subcontract from Lawrence Berkeley National Laboratory under the DOE ExaBiome project.	$400,000

Journal Papers

1. Azad, A, O Selvitopi, MT Hussain, J Gilbert, and A Buluç (2021). Combinatorial BLAS 2.0: Scaling combinatorial algorithms on distributed-memory systems. *IEEE Transactions on Parallel and Distributed Systems*.

2. Omodior, O, MR Saeedpour-Parizi, MK Rahman, A Azad, and K Clay (2021). Using convolutional neural networks for tick image recognition–a preliminary exploration. *Experimental and Applied Acarology*, 1–16.

3. Saeedpour-Parizi, MR, SE Hassan, A Azad, KJ Baute, T Baniasadi, and JB Shea (2021). Target position and avoidance margin effects on path planning in obstacle avoidance. *Scientific Reports* **11**(1), 1–18.

4. Azad, A, A Buluç, XS Li, X Wang, and J Langguth (2020). A distributed-memory algorithm for computing a heavy-weight perfect matching on bipartite graphs. *SIAM Journal on Scientific Computing* **42**(4), C143–C168.

5. Wanyan, T, H Honarvar, A Azad, Y Ding, and BS Glicksberg (2020). Deep Learning with Heterogeneous Graph Embeddings for Mortality Prediction from Electronic Health Records. *Data Intelligence*, 1–13.

6. Wanyan, T, A Vaid, JK De Freitas, S Somani, R Miotto, GN Nadkarni, A Azad, Y Ding, and BS Glicksberg (2020). Relational Learning Improves Prediction of Mortality in COVID-19 in the Intensive Care Unit. *IEEE Transactions on Big Data*, 1–1.

7. Yelick, K, A Buluç, M Awan, A Azad, B Brock, R Egan, S Ekanayake, M Ellis, E Georganas, G Guidi, et al. (2020). The parallelism motifs of genomic data analysis. *Philosophical Transactions of the Royal Society A* **378**(2166), 20190394.

8. Zhang, Y, A Azad, and A Buluç (2020). Parallel algorithms for finding connected components using linear algebra. *Journal of Parallel and Distributed Computing* **144**, 14–27.

9. Nagasaka, Y, S Matsuoka, A Azad, and A Buluç (2019). Performance optimization, modeling and analysis of sparse matrix-matrix products on multi-core and many-core processors. *Parallel Computing* **90**, 102545.

10. Azad, A, GA Pavlopoulos, CA Ouzounis, NC Kyrpides, and A Buluç (2018). HipMCL: a high-performance parallel implementation of the Markov clustering algorithm for large-scale networks. *Nucleic Acids Research* **46**(6), e33–e33.

11. Azad, A, A Buluç, and A Pothen (2017). Computing Maximum Cardinality Matchings in Parallel on Bipartite Graphs via Tree-Grafting. *IEEE Transactions on Parallel and Distributed Systems (TPDS)* **28**(1), 44–59.

12. Azad, A, G Ballard, A Buluc, J Demmel, L Grigori, O Schwartz, S Toledo, and S Williams (2016). Exploiting multiple levels of parallelism in sparse matrix-matrix multiplication. *SIAM Journal on Scientific Computing* **38**(6), C624–C651.

13. Azad, A and A Buluç (2016). A matrix-algebraic formulation of distributed-memory maximal cardinality matching algorithms in bipartite graphs. *Parallel Computing* **58**, 117–130.

14. Azad, A, A Pothen, and B Rajwa (2016). Immunophenotype Discovery, Hierarchical Organization, and Template-based Classification of Flow Cytometry Samples. *Frontiers in Oncology* **6**, 188.

15. Azad, A, B Rajwa, and A Pothen (2016). flowVS: channel-specific variance stabilization in flow cytometry. *BMC bioinformatics* **17**(1), 291.

16. Halappanavar, M, A Pothen, A Azad, F Manne, J Langguth, and A Khan (2015). Codesign lessons learned from implementing graph matching on multithreaded architectures. *Computer* **48**(8), 46–55.

17. Langguth, J, A Azad, M Halappanavar, and F Manne (2014). On parallel push–relabel based algorithms for bipartite maximum matching. *Parallel Computing* **40**(7), 289–308.

18. Aghaeepour, N, G Finak, F Consortium, D Consortium, H Hoos, TR Mosmann, R Brinkman, R Gottardo, and RH Scheuermann (2013). Critical assessment of automated flow cytometry data analysis techniques. *Nature Methods*.

19. Azad, A, S Pyne, and A Pothen (2012). Matching phosphorylation response patterns of antigen-receptor-stimulated T cells via flow cytometry. *BMC Bioinformatics* **13**(Suppl 2), S10.

Conference Proceedings

1. Hussain, MT, O Selvitopi, A Buluç, and A Azad (2021). Communication-Avoiding and Memory-Constrained Sparse Matrix-Matrix Multiplication at Extreme Scale. In: *2021 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*.

2. Rahman, M, MH Sujon, and A Azad (2021). FusedMM: A Unified SDDMM-SpMM Kernel for Graph Embedding and Graph Neural Networks. In: *2021 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*.

3. Azad, A, MM Aznaveh, S Beamer, M Blanco, J Chen, L D'Alessandro, R Dathathri, T Davis, K Deweese, J Firoz, et al. (2020). In: *2020 IEEE International Symposium on Workload Characterization (IISWC)*. IEEE, pp.216–227.

4. Gu, Z, J Moreira, D Edelsohn, and A Azad (2020). Bandwidth-Optimized Parallel Algorithms for Sparse Matrix-Matrix Multiplication using Propagation Blocking. In: *Symposium on Parallelism in Algorithms and Architectures (SPAA) 2020*.

5. Majeske, N, B Abesh, C Zhu, and A Azad (2020). Inductive Predictions of Extreme Hydrologic Events in The Wabash River Watershed. In: *AI for Earth Sciences Workshop at NeurIPS 2020*.

6. Rahman, M, MH Sujon, and A Azad (2020). BatchLayout: A Batch-Parallel Force-Directed Graph Layout Algorithm in Shared Memory. In: *IEEE Pacific Visualization Symposium (PacificVis 2020)*.

7. Rahman, M, MH Sujon, and A Azad (2020). Force2Vec: Parallel force-directed graph embedding. In: *20th IEEE International Conference on Data Mining (IEEE ICDM 2020), best paper nomination*.

8. Selvitopi, O, S Ekanayake, G Guidi, G Pavlopoulos, A Azad, and A Buluc (2020). Distributed many-to-many protein sequence alignment using sparse matrices. In: *SC20, International Conference for High Performance Computing, Networking, Storage and Analysis, 2020*.

9. Selvitopi, O, MT Hussain, A Azad, and A Buluç (2020). Optimizing high performance markov clustering for pre-exascale architectures. In: *2020 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*.

10. Wanyan, T, Y Ding, A Azad, and BS Glicksberg (2020). Coupling Heterogeneous Graph Embeddings with Convolution Neural Networks Improves Mortality Prediction. In: *Workshop on Knowledge Graph: Mining Knowledge Graph for Deep Insights (with KDD 20)*.

11. Zhang, Y, A Azad, and Z Hu (2020). FastSV: A distributed-memory connected component algorithm with fast convergence. In: *Proceedings of the 2020 SIAM Conference on Parallel Processing for Scientific Computing (SIAM PP)*. Society for Industrial and Applied Mathematics, pp.46–57.

12. Azad, A and A Buluc (2019). LACC: a linear-algebraic algorithm for finding connected components in distributed memory. In: *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. IEEE, pp.2–12.

13. Nagasaka, Y, S Matsuoka, A Azad, and A Buluç (2019). Performance optimization, modeling and analysis of sparse matrix-matrix products on multi-core and many-core processors. *Parallel Computing* **90**, 102545.

14. Rahman, MK and A Azad (2019). Evaluating the Community Structures from Network Images Using Neural Networks. In: *International Conference on Complex Networks and Their Applications*. Springer, Cham, pp.866–878.

15. Chen, Y, RD Calvert, A Azad, B Rajwa, J Fleet, T Ratliff, and A Pothen (2018). Phenotyping Immune Cells in Tumor and Healthy Tissue Using Flow Cytometry Data. In: *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM BCB)*, pp.73–78.

16. Gholami, A, A Azad, P Jin, K Keutzer, and A Buluc (2018). Integrated model, batch, and domain parallelism in training neural networks. In: *Proceedings of the 30th on Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pp.77–86.

17. Koanantakool, P, A Ali, A Azad, A Buluc, D Morozov, L Oliker, K Yelick, and SY Oh (2018). Communication-avoiding optimization methods for distributed massive-scale sparse inverse covariance estimation. In: *International Conference on Artificial Intelligence and Statistics (AISTATS)*. PMLR, pp.1376–1386.

18. Nagasaka, Y, S Matsuoka, A Azad, and A Buluç (2018). High-performance sparse matrix-matrix products on Intel KNL and multicore architectures. In: *Proceedings of the 47th International Conference on Parallel Processing Companion*, pp.1–10.

19. Azad, A and A Buluc (2017). A work-efficient parallel sparse matrix-sparse vector multiplication algorithm. In: *Parallel & Distributed Processing Symposium (IPDPS)*.

20. Azad, A and A Buluç (2017). Towards a GraphBLAS Library in Chapel. In: *IPDPS Workshops*.

21. Azad, A, M Jacquelin, A Buluc, and EG Ng (2017). The Reverse Cuthill-McKee Algorithm in Distributed-Memory. In: *Parallel & Distributed Processing Symposium (IPDPS)*.

22. Azad, A and A Buluç (2016). Distributed-Memory Algorithms for Maximum Cardinality Matching in Bipartite Graphs. In: *Parallel & Distributed Processing Symposium (IPDPS)*.

23. Koanantakool, P, A Azad, A Buluç, D Morozov, SY Oh, L Oliker, and K Yelick (2016). Communication-Avoiding Parallel Sparse-Dense Matrix-Matrix Multiplication. In: *Parallel & Distributed Processing Symposium (IPDPS)*.

24. Azad, A and A Buluç (2015). Distributed-memory algorithms for maximal cardinality matching using matrix algebra. In: *2015 IEEE International Conference on Cluster Computing*. IEEE, pp.398–407.

25. Azad, A, A Buluç, and J Gilbert (2015). Parallel Triangle Counting and Enumeration using Matrix Algebra. In: *Workshop on Graph Algorithms Building Blocks (GABB), in conjunction with IPDPS*. IEEE.

26. Azad, A, A Buluç, and A Pothen (2015). A Parallel Tree Grafting Algorithm for Maximum Cardinality Matching in Bipartite Graphs. In: *Parallel & Distributed Processing Symposium (IPDPS)*. IEEE.

27. Azad, A, A Khan, B Rajwa, S Pyne, and A Pothen (2013). Classifying immunophenotypes with templates from flow cytometry. In: *Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics (ACM BCB)*, pp.256–265.

28. Azad, A, M Halappanavar, S Rajamanickam, EG Boman, A Khan, and A Pothen (2012). Multithreaded Algorithms for Maximum Matching in Bipartite Graphs. In: *Parallel & Distributed Processing Symposium (IPDPS)*. IEEE, pp.860–872.

29. Azad, A and A Pothen (2012). Multithreaded algorithms for matching in graphs with application to data analysis in flow cytometry. In: *2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum*. IEEE, pp.2494–2497.

30. Azad, A, M Halappanavar, F Dobrian, and A Pothen (2011). Computing maximum matching in parallel on bipartite graphs: worth the effort? In: *Proceedings of the 1st Workshop on Irregular Applications: Architectures and Algorithms*, pp.11–14.

31. Azad, A, J Langguth, Y Fang, A Qi, and A Pothen (2010). Identifying rare cell populations in comparative flow cytometry. In: *International Workshop on Algorithms in Bioinformatics (WABI)*. Springer, Berlin, Heidelberg, pp.162–175.

32. Al Islam, AA, A Azad, K Alam, and S Alam (2007). Security attack detection using genetic algorithm (ga) in policy based network. In: *2007 International Conference on Information and Communication Technology*. IEEE, pp.341–347.

33. Azad, MA, AA Al Islam, MK Alam, and MS Alam (2007). Router Oriented Traffic Flow Analysis for IP Backbone Network. In: *2007 International Conference on Information and Communication Technology*. IEEE, pp.348–351.

Teaching

➤ **ENGR-E599 Graph Analytics**: Spring 19, 20, 21 (a newly-developed course at IU)
➤ **ENGR-E503 Introduction to Intelligent Systems**: Fall 19, 20, 21 (a newly-developed course at IU)
➤ **Contributed lectures**: (a) Fall 20: T100 What is Data Science? (b) Spring 19: ENGR-E599 Big Data Systems

Conference Organization

➤ **Program Co-Chair**: HiCOMB workshop, IPDPS 2020
➤ **Minimymposium organizer**: SIAM Annual Meeting 2018
➤ **Program committee member**:

2022	IPDPS
2021	SC, IPDPS, IEEE CLUSTER, ACM BCB, PacificVis, HiCOMB
2020	IPDPS, SPAA, SIAM CSC, ACM BCB, HiPC, NSysS
2019	SC, IPDPS, HiPC, HiCOMB, NSysS
2018	SC, IPDPS, IEEE CLUSTER, HiCOMB, Supercomputing-Asia, NSysS
2017	SC, IPDPS, IEEE CLUSTER, PASC, BigDF (HiPC), HiCOMB, NSysS
2016	IPDPS, BigGraphs, HPGDMP (SC)

Journal Reviewer

IEEE Transactions on Parallel and Distributed Systems (TPDS), IEEE Transactions on Multi-Scale Computing Systems (TMSCS), Parallel Computing, Journal of Parallel and Distributed Computing (JPDC), SIAM Journal on Scientific Computing (SISC), ACM Transactions on Knowledge Discovery from Data (TKDD), Cytometry A

Advising and Mentoring

➤ **Current Students**

PhD Students	Md Khaledur Rahman (CS), Md Taufique Hussain (ISE), Nicholas Majeske (ISE), Selahattin Akkas (ISE) Tingyi Wanyan (ISE; co-advised by Ying Ding from UT Austin)

➤ **Past Students**

MS Students	Anuj Godase (now at Amazon), Abhigya Agrawal, Zhixiang Gu (now at Facebook), Piyush Vyas
Undergraduates	Kyle Combs, James Dumas
Interns	Guttu Sai Abhishek (from IIT Bombay, India)
Visiting Scholars	Yongzhe Zhang (from SOKENDAI, Japan)

Significant Software

1. Force2Vec. Parallel algorithms for graph embedding and visualization. Written using Python, C++, and OpenMP.

2. The Combinatorial BLAS. A distributed-memory parallel graph library offering a small but powerful set of linear algebra primitives specifically targeting graph analytics. Written using C++, MPI and OpenMP. Scales up to 200,000 cores on NERSC supercomputers.

3. HipMCL. A high-performance software for large-scale network clustering, parallelizing the popular Markov Clustering (MCL) algorithm. Written using C++, MPI and OpenMP. *Scales up to more than 1 million hardware threads on Cori supercomputer at NERSC.*

4. MS-BFS-Graft. A shared-memory parallel implementation of bipartite cardinality matching algorithms. Written using C++ and OpenMP.

5. flowMatch. A Biconductor package for matching cell clusters and building templates from flow cytometry data. Written in R and C++. *More than 15,000 downloads*.

6. FlowVS : A Biconductor package to stabilize variance in flow cytometry and microarray data. Written in R and C++. *More than 12,000 downloads*.

Invited Talks and Seminars

1. Mar 2021. Performance Portable Algorithms for Machine Learning on Graphs. In DATA lab speaker series, Northeastern University, Boston, MA. Invited Seminar, Host: Wolfgang Gatterbauer.

2. Jan 2021. Computational Building Blocks for Machine Learning on Graphs. In EECS, MIT, Cambridge, MA. Invited Seminar, Host: Julian Shun.

3. Oct 2020. Computational Building Blocks for Machine Learning on Graphs. In CS, Rensselaer Polytechnic Institute, Troy, NY. Invited Seminar, Host: George Slota.

4. Feb 2020. Big Graph Analytics and Learning. In Data Science Immersion Day, Indiana University.

5. May 2019. HipMCL: A High-Performance Parallel Algorithm for Clustering Large-scale Networks. In HiCOMB workshop at IPDPS 19, Invited Speaker.

Awards and Honors

1. DOE Early Career Award, 2021

2. IBM PhD Fellowship, 2013

3. Best poster award (CYTO 2013)

4. Travel awards: SIAM PP 2016, Cyto 2015, ACM BCB 2013, NIH 2012, Purdue 2011

5. Dean's merit award 2003-05, Bangladesh University of Engineering and Technology

Synergistic Activities

➤ **Technical forum participation:** The GraphBLAS Forum, the flowCAP Forum at NIH

➤ **Active member:** Berkeley Institute of Data Science (BIDS), Berkeley Benchmarking and OPtimization Group (BeBOP)

➤ **Training:**

- Argonne Extreme Scale Computing Program, 2014
- Early Career Program, Supercomputing, 2017

➤ **Professional membership:** SIAM, ACM, IEEE

➤ **Leadership Activities:** Organizer, LBNL postdoc coordination program (2014-16), President, Purdue Bangladesh Student Association (2012-13), Webmaster, Purdue Graduate Student Government (2009-11), Volunteer, Bay Area Scientists in Schools (2014-2018).