
Barna Saha

180 Park Avenue, Room C-257
AT&T Shannon Research Laboratory
Florham Park, NJ 07932. Phone: 617 331 1896
Email: barna@research.att.com
Homepage: <http://www.cs.umd.edu/~barna>

Research Interests

- Design and Analysis of Algorithms
 - Algorithms for Large Scale Data Management, Approximation Algorithms, Randomized Algorithms, Graph Algorithms, Optimization on Cloud.
- Databases, Data Quality and Data Integration, Distributed Data Management, Data Streaming Algorithms.
- Probabilistic Methods, Combinatorial Optimization.

Current Position

Senior Member of Technical Staff Research (2011–),
AT&T Shannon Research Laboratory, Florham Park, NJ.

Education

Ph.D. (2007-2011) in Computer Science,
University of Maryland, College Park.
GPA: 4.0/4.0.
Advisor: Prof. Samir Khuller.
Dissertation: Approximation Algorithms for Resource Allocation.

Master of Technology in Computer Science & Engineering,
Indian Institute of Technology (IIT), Kanpur, India.
GPA: 10.0/10.0.
Thesis: Graph Streaming Algorithms for Counting Patterns in Massive Graphs.

Bachelor of Engineering in Computer Science & Engineering,
Jadavpur University, Kolkata, India.
GPA: 3.86/4.00.

Publications

1. Yury Polyanskiy, Arya Mazumdar, Barna Saha, “The Closest String Problem and the Chebyshev Radius”, Accepted, *IEEE International Symposium on Information Theory (ISIT)*, 2013, Istanbul, Turkey, 2013.
2. Xin Luna Dong, Barna Saha, Divesh Srivastava, “Less is More, Selecting Sources Wisely for Integration”, Accepted, *39th International Conference on Very Large Data Bases (VLDB)*, 2013, Riva del Garda, Trento, 2013 .
3. Lukasz Golab, Howard Karloff, Flip Korn, Barna Saha, Divesh Srivastava, “Discovering Conservation Rules”, Proc. *28th IEEE International Conference on Data Engineering (ICDE)*, Washington DC, 2012.
Journal version, to appear in *IEEE Transactions on Knowledge and Data Engineering (TKDE)* **Special issue on the Best Papers of ICDE 2012**.
4. Barna Saha, Samir Khuller, “Set Cover Revisited: Hypergraph Cover with Hard Capacities,” Proc. *International Colloquium on Automata, Languages, and Programming (ICALP)*, Track A, Warwick, UK, 2012.

5. Samir Khuller, Barna Saha, Kanthi Sarpatwar, “New Approximation Results for Resource Replication Problems,” Accepted, *The 15th. International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*, Boston, 2012.
6. Ravishankar Krishnaswamy, Vishwanath Nagarajan, Barna Saha “The Matroid Median Problem,” Proc. *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, San Francisco, 2011. Merged with a paper by Amit Kumar and Yogish Sabharwal who obtained similar results for the case of partition matroids. Journal version, submitted to *ACM Transactions on Algorithms*, 2012.
7. Saeed Alaei, Mohammad Taghi Hajiaghayi, Vahid Liaghat, Dan Pei, Barna Saha, “AdCell-Ad Allocation in Cellular Networks,” Proc. *19th European Symposium on Algorithms (ESA)*, Saarbrücken, Germany, 2011. Journal version, submitted to *ACM Transactions on Algorithms*, 2012.
8. Nikhil Bansal, Ravishankar Krishnaswamy, Barna Saha, “On Capacitated Set Cover Problem,” Proc. *The 14th. International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*, Princeton, 2011.
9. Philip Anderson, Samir Khuller, Saket Navlakha, Louiqa Raschid, Barna Saha, Andreas Thor, Xiao-Ning Zhang, “Link Prediction for Annotation Graphs,” Proc. *International Semantic Web Conference (ISWC)*, Bonn, Germany, 2011.
10. Bernhard Haeupler, Barna Saha, Aravind Srinivasan, “New Constructive Aspects of the Lovász Local Lemma,” Proc. *IEEE Symposium on Foundations of Computer Science (FOCS)*, Las Vegas, 2010. Journal version, Published in *Journal of the ACM (JACM)*, 58(6), 28, 2011.
11. Samir Khuller, Jian Li, Barna Saha, “Energy Efficient Scheduling via Partial Shutdown,” Proc. *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, Austin, 2010.
12. Barna Saha, Aravind Srinivasan, “A New Approximation Technique for Resource-Allocation Problems,” Proc. *Innovations in Theoretical Computer Science (ITCS) 2010*, Beijing, China, 2010. Journal version, submitted to *Siam Journal on Computing*, 2012.
13. Barna Saha, Allison Hoch, Samir Khuller, Louiqa Raschid and Xiao-Ning Zhang, “Dense Subgraphs with Restrictions and Applications to Gene Annotation Graphs,” Proc. *International Conference on Research in Computational Molecular Biology (RECOMB)*, Lisbon, Portugal, 2010.
14. Barna Saha, Ioana Stanoi, Kenneth L. Clarkson, “Schema Covering: a Step Towards Enabling Reuse in Information Integration,” Proc. *26th IEEE International Conference on Data Engineering (ICDE)*, Long Island, 2010.
15. Jian Li, Barna Saha, Amol Deshpande, “A Unified Approach to Ranking in Probabilistic Databases,” Proc. *35th International Conference on Very Large Data Bases (VLDB), 2009*, (**BEST PAPER AWARD**), Leon, France, 2009. Journal version, published in THE VLDB JOURNAL **Special Issue on the Best Papers from VLDB 2009**.
16. Samir Khuller, Barna Saha “On Finding Dense Subgraphs,” Proc. *International Colloquium on Automata, Languages, and Programming (ICALP, TRACK A)*, Rhodes, Greece, 2009.

17. Barna Saha, Lise Getoor, “On Maximum Coverage in the Streaming Model & Application to Multi-topic Blog-Watch,” Proc. *9th SIAM International Conference on Data Mining (SDM)*, Reno, 2009.
18. Bogdan Alexe, Michael Gubanov, Mauricio A. Hernandez, Howard Ho, Jen-Wei Huang, Yannis Katsis, Lucian Popa, Barna Saha, Ioana Stanoi, “Simplifying Information Integration: Object-Based Flow-of-Mappings Framework for Integration,” *Lecture Notes in Business Information Processing, Revised Selected Papers from Second International VLDB Workshop, BIRTE*, 2009, Volume 27.
19. Barna Saha, Lise Getoor, “Group Proximity Measure for Recommending Groups in Online Social Networks,” Proc. *2nd ACM SIGKDD Workshop on Social Network Mining and Analysis (SNA-KDD)*, Las Vegas, 2008.
20. Sumit Ganguly, Barna Saha, “On Estimating Path-Aggregates over Streaming Graphs,” Proc. *17th International Symposium on Algorithms and Computations (ISAAC)*, Kolkata, India, 2006.

Submitted Papers

1. Howard Karloff, Barna Saha, “Renting a Cloud,” Submitted.
2. Lukasz Golab, Flip Korn, Feng Li, Barna Saha, Divesh Srivastava, “Summarizing Event Response Data,” Submitted.
3. Flip Korn, Barna Saha, Divesh Srivastava, Shanshan Yang, “On Repairing Structural Problems In Semi-structured Data”, Submitted.
4. Mohammadtaghi Hajiaghayi, Jian Li, Shi Li, Barna Saha, “Fault Tolerant K-Median Problem”, Submitted.

Patents

1. Lukasz Golab, Howard Karloff, Flip Korn, Barna Saha, Divesh Srivastava, “Conservation Dependency,” US patent US20120130935, AT&T Research Center, New Jersey, 2012.
2. Saeed Alaei, Mohammad Taghi Hajiaghayi, Vahid Liaghat, Dan Pei, Barna Saha, “AdCell-Ad Allocation in Cellular Networks,” US patent pending, AT&T Research Center, New Jersey, 2011.

Professional Activities **Panelist:** NSF Big Data.

Programm Committee Member: Grace Hopper Celebration of Women in Computing (GHC), 2012.

Reviewer (Journals): IEEE Transactions on Knowledge Discovery; Internet Mathematics; Operational Research Letters, Algorithmica, Transactions on Algorithms.

Reviewer (Conferences): FOCS, SODA, STOC, ESA, KDD, SDM, VLDB, PODS, ISAAC, APPROX, RANDOM etc.

Awards and Scholarships

- **Dean's PhD Fellowship Award (2010-2011)**, for excellence in Ph.D. Research, University of Maryland Computer Science Department, College Park.
- **Travel Grant (2010)**, from Foundation of Computer Science (FOCS), Las Vegas.
- **Travel Grant (2010)**, from Innovations in Computer Science (ICS), Beijing, China.
- **Best Paper Award: 35th International Conference on Very Large Data Bases (VLDB)**, (2009).
- **Goldhaber Travel Award** (2009), from University of Maryland Graduate School.
- **Graduate Fellowship award** (2007-2009), from University of Maryland Computer Science Department, College Park.
- **Best Presentation Award at IRISS (Inter Research Institute Student Seminar, co-event conducted along with IJCAI 2007)**, 2007 in Foundations and Algorithms Session.
- **Academic Excellence Award**, for the year 2004-05, IIT Kanpur.
- **Ministry of Human Resource and Development Scholarship**, India (2004–06), as a GATE (Graduate Aptitude Test for Engineers)-qualified student.
- Awarded by **State Government of West Bengal, Governor of West Bengal** and **State Bank of India**, for outstanding performance in Higher Secondary Examination (+12 examination) (rank 2nd in the state out of 400,000), 2000.
- **Jagdish Bose National Science Talent Search (JBNSTS) Scholarship** (1999-2000).
- **State Government of West Bengal Award** for securing highest marks in English in Higher Secondary Examination, 2000.
- **State Government of West Bengal Award and State Bank of India Award** for outstanding performance in Secondary Examination (+10 examination)(8th in the state out of 600,000), 1998.

Research Experience**AT&T Shannon Research Laboratory**

Senior Member of Technical Staff, Research, August 2011-Current.

IBM T.J. Watson Research Center, Yorktown Heights, NY

Research Intern, Jun–Aug 2010.

AT&T Shannon Research Laboratory, Florham Park, New Jersey,

Research Intern, Jun–Aug 2009.

IBM Almaden Research Center, San Jose, CA

Research Intern, Jun–Aug 2008.

University of Maryland, College Park, MD

Graduate Research Assistant, Fall 2007-Summer 2008, Fall 2009-Summer 2011.

IIT Kanpur, India

Master's Thesis Research, Fall 2005–Spring 2006.

Indian Statistical Institute, Kolkata, India

Undergraduate Intern, Summer 2003, Summer 2004.

Mentoring Experience**Student Interns**

1. Donatella Firmani , Sapienza University of Rome, (April 2012-July 2012),
Project: Map Reduce Algorithms for Graph Clustering.
2. Kook Jin Ahn, University of Pennsylvania, (June 2012-August 2012),
Project: Resource Allocation Problems on Cloud.
3. Harmeet Jandu, Rutgers University, (September 2012-December 2012),
Project: Completed her Master's thesis under my supervision on Optimization on Hadoop.

Teaching Experience

I am visiting University of Minnesota Twin Cities, Computer Science Department in Fall 2013 and will be offering a new advanced graduate course on ALGORITHMIC TECHNIQUES FOR BIG DATA ANALYSIS

University of Maryland, College Park, MD

Teaching Assistant, Fall 2008, Spring 2009

I was a TA for the undergraduate courses on Cryptography and Algorithm Design.

IIT Kanpur, India

Teaching Assistant, Fall 2004–Spring 2006

I was a TA for the graduate courses on Parallel Numerical and Semi-numerical Algorithms and Data-streaming Algorithms and Systems; and undergraduate courses on Artificial Intelligence and Theory of Computation.

Selected Invited Talks

Information Theory Workshop and Applications (ITA), Invited Talk, “Resource Replication Problems,” February, 2013.

University of Minnesota Twin Cities, “A Unified Approach to Resource Allocation Problems,” September, 2012 .

AT&T Shannon Research Laboratory, Florham Park, “AdCell-Ad Allocation in Cellular Networks,” February, 2012.

University of Maryland, “Discovering Conservation Rules,” , March, 2012.

Google Research, Mountain View, CA, “Allocating Resources & Processing Top-k Queries: The Power of Probabilistic Methods,” March, 2011.

Purdue University, “Allocating Resources & Processing Top-k Queries: The Power of Probabilistic Methods,” March, 2011.

AT&T Shannon Research Laboratory, Florham Park, “How to Approximate Better:

Computing Dense Subgraphs and Processing Top-K Queries,” March, 2011.

Yahoo! Research, Santa Clara “How to Approximate Better: Computing Dense Subgraphs and Allocating Resources,” February, 2011.

University of Wisconsin Madison, CS Theory Seminar Series, “New Constructive Aspects of The Lovász Local Lemma,” October, 2010.

IBM T.J. Watson Research Center, “The Matroid Median Problem,” August, 2010.

IBM T.J. Watson Research Center, IP (Integer Programming) Talk, “New Constructive Aspects of The Lovász Local Lemma,” June, 2010.

Center for Bioinformatics & Computational Biology, University of Maryland, “Dense Subgraphs with Restrictions and Applications to Gene Annotation Graphs,” April, 2010.

AT&T Shannon Labs, Florham Park, NJ, “Conservation Dependency in Data-Quality,” August, 2009.

References References will be provided upon request.