Aritra Bose

Lawson Building
Department of Computer Science
Purdue University
305 N, University Street
West Lafayette, IN 47906

Email: bose6@purdue.edu, bosearitra08@gmail.com
Web: `http://web.ics.purdue.edu/~bose6/`
Phone: +1(518) 522-7975

Research Interests

Population Genetics, Machine Learning, Bioinformatics, Network Science, Data Mining, Randomized Numerical Linear Algebra, Biological Networks

Education

Purdue University — West Lafayette, IN, USA
Ph.D. in *Computer Science* — Aug 2016 - Present
Adviser: Prof. Petros Drineas

Rensselaer Polytechnic Institute — Troy, NY, USA
M.S. in *Computer Science* — Aug 2014 - Jul 2016

West Bengal University of Technology — Kolkata, WB, India
B.Tech in *Information Technology* — Aug 2009 - Jun 2013

Experience

Research Assistant, Purdue University — West Lafayette, IN, USA
Aug 2016 - Present

- Developing novel techniques to understand complex genetic structure of human populations and reconstructing population history.
- Applying statistical machine learning on neurodegenerative and neurodevelopmental disorders as a part of the Psychiatric Genomics Consortium.
- Detecting de novo variants with Whole Exome Sequencing data for Tourette Syndrome.

Summer Research Intern, IBM T.J. Watson Research Center — Yorktown Heights, NY, USA
May 2018 - Jul 2018

- Worked in the Computational Biology group under the supervision of Dr. Laxmi Parida and Dr. Daniel E. Platt, on population genetic models.

Summer Research Intern, IBM T.J. Watson Research Center — Yorktown Heights, NY, USA
May 2017 - Aug 2017

- Developed algorithms to modulate Ancestral Recombination Graphs under selection.

Summer Research Intern, IBM T.J. Watson Research Center — Yorktown Heights, NY, USA
May 2016 - Aug 2016

- Worked in the Computational Biology group on sampling Ancestral Recombination Graphs of multiple populations.
- Worked on developing a novel method to infer on sociolinguistic stratification of Indian subcontinent and to highlights on its demographic history.

Research Assistant, Rensselaer Polytechnic Institute — Troy, NY, USA
Aug 2014 - May 2016

- Applying data mining techniques to analyze population genetic networks and infer on relationships between various world populations.
- Performing noise reduction and feature selection to model the game of chess for predictive analysis using Machine Learning.

Teaching Assistant, Rensselaer Polytechnic Institute — Troy, NY, USA
Aug 2014 - May 2014

- Teaching assistant of Data Structures (CSCI 1200), helped instructor, Prof. Barbara Cutler, supervising lab classes of up to 100 students, guiding them during office hours and grading homework and tests.

Analyst-Intern, Teradata Corporation — Hyderabad, India
Oct 2013 - April 2014

Research Trainee, CoE in Bioinformatics, Bose Institute — Kolkata, India
Sep 2012 - Oct 2013

- Construction and Analysis of Differentially Regulated Gene Networks to identify sets of key gene which regulates divergent host responses to any diseased system.

Final Project, Meghnad Saha Institute of Technology — Kolkata, India
Aug 2012 - May 2013

- Developed a software package to synthesize Quantum Computing Circuits.

Summer Intern, Indian Institute of Technology — Guwahati, India
May 2012 - Jul 2012

- Construction and Analysis of Integrated Regulatory Network derived from High Throughput Sequencing data and studying ChIP-Seq analysis techniques.
- Studied PeakSeq, a scoring methodology for ChIP-Seq experiments.

Winter Intern, Indian Statistical Institute — Kolkata, India
Dec 2011 - Mar 2012

- A quantitative model for Human Olfactory Receptors using Fractal parameters.

Publications

1. **A. Bose**, P. Paschou, P. Drineas, *Meta-analyses and identification of Disease Informative Markers from genome-wide association studies*, In Progress.
2. **A. Bose**, V. Kalantzis, E. Kontopoulou, M. Elkady, P. Paschou, P. Drineas, *TeraPCA: a fast and scalable software package to study genetic variation in tera-scale genotypes*, to appear in Bioinformatics (2019).
3. **A. Bose**, F. Utro, D.E. Platt, L. Parida, *Multiple Loci Selection with Multi-way Epistasisin Coalescence with Recombinations*, Under review.
4. **A. Bose**, D.E. Platt, L. Parida, P. Paschou, P. Drineas, *Integrating linguistics, social structure, and geograpy to model gene flow in India*, Under review.
5. G. Stamatoyannopoulos, **A. Bose**, A. Teodasiadis, F. Tsetsos, A. Plantiga, N. Psatha, N. Zogas, E. Yannaki, P. Zalloua, K.K. Kidd, B.L. Browning, J. Stamatoyannopoulos, P. Paschou, P. Drineas, *Genetics of the Peloponnesean populations and the theory of the extinction of the medieval Peloponnesean Greeks*, European Journal of Human Genetics (EJHG), 25(5), pp. 637-645, 2017..
6. S.Hassan, P. Pal Choudhury and **A. Bose**,(2011), *A Quantitative model for Human Olfactory Receptors*, Nature Precedings, npre20126967-2

Talks and Posters in Conferences, Workshops or Seminars Attended

- Preliminary exam talk on *Computational methods in Population Genetics*, given to the PhD committee in the Computer Science Department, Purdue University, IN.
- **Platform** presentation on *Algorithms to modulate ARG by Selection* at the **RECOMB-Genetics** meeting held in Paris on April, 2018. (This talk is given by Dr. Laxmi Parida)
- Poster presentation on *TeraPCA: A fast and scalable method to study genetic variation in tera-scale genotypes* at **Conference on Scientific Computing and Approximation**, held at Purdue University, IN.
- Poster presentation on *TeraPCA: A fast and scalable method to study genetic variation in tera-scale genotypes* at **ASHG 2017** meeting held at Orlando, FL.
- Poster presentation on *COGG* at Summer Intern Showcase held in IBM T.J. Watson Research Center, NY.
- Poster presentation on *COGG* at **Biology of Genomes 2017** meeting held at Cold Spring Harbor Labs, NY.
- **Platform** talk on *COGG (Correlation Optimization of Genetics and Geodemographics)*, a novel method to model sociolinguistic stratification in India at **American Society for Human Genetics (ASHG) 2016**, Annual Meeting in Vancouver, BC, Canada. Abstract selected in top 8% out of over 6000 submissions.
- Presented a poster at Student Research Showcase at Computer Science Department, Purdue University, IN.
- Presented a poster at **Biology of Genomes 2016** conference held in Cold Spring Harbor Labs in Cold Spring, NY.
- Presented posters in Student Research Symposium at Computer Science Department, Rensselaer Polytechnic Insitute.
- **Summer school** on "Mathematics of Data", organized by **Park City Mathematics Institute (PCMI)** and the Institute for Advanced Study (IAS), held in, Midway, Utah, USA.
- American Society for Human Genetics, 2015, Annual Meeting in Baltimore, MD, USA as a trainee researcher.
- **Gene Golub SIAM Summer School 2015**, held in, Delphi, Greece.

Computer Skills

- ***Languages:*** C, C++, Python, MATLAB, Java, R, AWK, PL/SQL, Scripting(bash,etc), Perl, HTML, LaTeX
- ***Operating Systems:*** GNU/Linux, Unix, Windows
- ***Computational Biology:*** GATK, PLINK, GSEA, MeV, samtools, bcftools, PeakSeq, Cytoscape, EIGENSOFT, ADMIXTURE, BEAGLE, GERMLINE, IMPUTE and other computational biology tools.
- ***Databases:*** MySQL, TERADATA, Oracle, DB2

Graduate Coursework (ongoing)

Machine Learning, Computational Linear Algebra, Parallel Computing, Computer Operating Systems, Foundations of Data Science, Design and Analysis of Algorithms, Programming Languages, Frontiers of Network Science, Distributed Systems, Randomized Algorithms, Randomized Algorithms in Numerical Linear Algebra, Theory of Computability and Complexity, Data Communications and Networking

Independent Coursework

Coursera.org: Algorithms: Design and Analysis, by Stanford University; Bioinformatis I and II, by University of Toronto, Python for Genomic Data Science by Johns Hopkins University.

Activities & Achievements

- Reviewer for the following:
  Journals
  American Journal of Medical Genetics(2); BMC Bioinformatics (3)
  Conferences
  KDD(2); NIPS(1); RECOMB(1); WABI(1)
- Received an "early offer" given to the best intern in the Computational Genomics group at IBM T.J. Watson Research Center for an internship in Summer of 2017.
- Got selected as a Research Intern in GE Global Research, NY for summer 2016.
- Got selected as a Graduate Summer Intern for summer 2016, in the Text&Graph Analytics group of Xerox Research Center, India.
- Student member: American Society of Human Genetics.
- Peer Adviser to incoming graduate students in the Computer Science Department in Rensselaer Polytechnic Institute as well as in Purdue University.
- Co-Founder of the Robotics club of Meghnad Saha Institute of Technology which has over 400 students now.
- Presented a paper,"Automated Hybrid Bot" in Student's Convention, 2011 at Indian Statistical Institute, Kolkata, India organized by, Computer Society of India.

Awards of Merit

- NSF Travel Grant to the following conferences:
  - Biology of Genomes: 2016 and 2017 meeting held at Cold Spring Harbor Labs, NY.
  - American Society of Human Genetics (ASHG) 2015 at Baltimore, MD.; 2016 at Vancouver, BC, Canada; 2017 at Orlando, FL.
  - International Conference for Distributed Computing and Internet Technologies (ICDCIT) 2017 meeting held at Bhubaneswar, Odisha, India.
- Received a 4 year fellowship from Ministry of Human Resource Development (M.H.R.D), Government of India for significant achievement in Higher Secondary Examination
- Ranked 1 out of 65 students in Department of Information Technology in Meghnad Saha Institute of Technology, Kolkata.
- Won several school, college and corporate level Quiz competitions across India.

References

Prof. Petros Drineas
Associate Professor
Department of Computer Science
Purdue University
West Lafayette, IN, USA
drineas@gmail.com

Prof. Peristera Paschou
Associate Professor
Department of Biological Sciences
Purdue University
West Lafayette, IN, USA
ppaschou@gmail.com

Dr. Laxmi Parida
Distinguished RSM &
Manager, Computational Genomics, IBM T.J. Watson Research Center
Yorktown Heights, NY, USA
parida@us.ibm.com

Additional Information

- *Date of Birth*: August 8, 1990
- *Marital Status*: Single
- *Citizenship*: Indian