

Prashant Pandey

RESEARCH INTEREST	My research interests lie at the intersection of Systems and Algorithms. I design and build theoretically well-founded data structures for big data problems in computational biology, storage, and streaming.	
CONTACT INFORMATION	2000 Walnut Avenue, Apt O301 Fremont, CA - 94538 Website Google Scholar Github	E-mail: ppandey@berkeley.edu Phone: (+1) 631-949-6948 https://prashantpandey.github.io https://goo.gl/Fz82hB https://github.com/prashantpandey/
EDUCATION	Stony Brook University , Stony Brook, NY <i>Ph.D. Computer Science</i> <i>Thesis: Fast and Space-Efficient Maps: Shrinking Big Data Down to Size</i> <i>Advisors: Prof. Michael Bender & Prof. Rob Johnson</i>	December 2018 <i>GPA (3.8/4.0)</i>
	University of Pune , Pune, India <i>Bachelor of Engineering (BE), Information Technology</i> Ranked 1st in college and 7th across University	August 2007 - June 2011 <i>First class with distinction</i>
WORK EXPERIENCE	Lawrence Berkeley Lab/UC Berkeley , Berkeley, CA <i>Postdoctoral Scholar, Computational Research Division</i> <i>Advisors: Prof. Kathy Yelick & Prof. Aydin Buluc</i>	December 2019 - Present
	Carnegie Mellon University , Pittsburgh, PA <i>Postdoctoral Scholar, School of Computer Science</i> <i>Advisor: Prof. Carl Kingsford</i>	December 2018 - November 2019
	Stony Brook University , Stony Brook, NY <i>Research Assistant, Applied Algorithms Lab</i>	August 2014 - October 2018
	TIBCO Inc. , Pune, India <i>Software Developer, Cloud Platform</i>	July 2011 - June 2013
INTERNSHIPS	Google , Manhattan, NY <i>Research Intern, Google Spanner</i>	May 2017 - August 2017
	Google , Kirkland, WA <i>Research Intern, Google Cloud Infrastructure</i>	May 2016 - August 2016
	Intel Labs , Portland, OR <i>Research Intern, Security and Privacy Lab</i>	May 2015 - August 2015
	Intel Labs , Portland, OR <i>Research Intern, Security and Privacy Lab</i>	May 2014 - August 2014
AWARDS AND ACHIEVEMENTS	<ul style="list-style-type: none">• Recipient of Catacosinos Fellowship for the most impactful research at SBU 2018• Best Paper Award FAST 2016 2016• Runner's Up to Best Paper FAST 2015 2015• A Special CS Department Chair Fellowship, Stony Brook University 2013• University Rank Holder, University of Pune 2011• Academic Excellence Scholarship, University of Pune. 2009, 2010, 2011	
PRESS ARTICLES ON RESEARCH	A general purpose counting filter: making every bit count. The Morning Paper. <i>August 2017</i> Link: https://goo.gl/nReGcF	
	Scaling Computational Biology at VMware. VMware Research Blog <i>April 2018</i> Link: https://blogs.vmware.com/research/2018/04/18/scaling-computational-biology-vmware/	
	Finding a Needle in a Field of Haystacks. Cell Systems publishes research on Mantis, a new sequencing search tool. <i>July 2018</i> Link: https://goo.gl/LJopwR	

MetaGNN: Metagenomic Reads Classification Using Graph Neural Networks

Prashant Pandey, Giulia Guidi, Alok Tripathy, Aydn Buluc, and Katherine Yelick

An Incrementally-Updatable and Scalable System for Large-Scale Sequence Search using LSM-Trees

Fatemeh Almodaresi, Jamshed Khan, Sergey Madaminov, **Prashant Pandey**, Michael Ferdman, Rob Johnson, and Rob Patro

Terrace: A Hierarchical Graph Container for Skewed Dynamic Graphs

Prashant Pandey, Brian Wheatman, Helen Xu, Aydin Buluc

VariantStore: A Space-Efficient and Fast Variant Search Index *bioRxiv 2020*

Prashant Pandey, Yinjie Gao, Carl Kingsford

Accepted to ISMB 2020 talk/poster

External-Memory Dictionaries in the Affine and PDAM Models *TOPC 2021*

Michael A. Bender, Alex Conway, Martin Farach-Colton, William Jannen, Yizheng Jiao, Rob Johnson, Eric Knorr, Sara McAllister, Nirjhar Mukherjee, **Prashant Pandey**, Donald E. Porter, Jun Yuan, Yang Zhan

Vector Quotient Filters: Overcoming the Time/Space Trade-Off in Filter Design

SIGMOD 2021

Prashant Pandey, Alex Conway, Joe Durie, Michael Bender, Martin Farach-Colton, Rob Johnson

Distributed-Memory k-mer Counting on GPUs *IPDPS 2021*

Israt Nisa, **Prashant Pandey**, Marquita Ellis, Leonid Oliker, Aydin Buluc, Katherine Yelick

An Efficient, Scalable, and Exact Representation of High-Dimensional Color Information Enabled Using de Bruijn Graph Search *JCB 2020*

Fatemeh Almodaresi, **Prashant Pandey**, Michael Ferdman, Rob Johnson, Rob Patro

Timely Reporting of Heavy Hitters using External Memory *SIGMOD 2020*

***Prashant Pandey**, ***Shikha Singh**, Michael A. Bender, Jonathan W. Berry, Martin Farach-Colton, Rob Johnson, Thomas M. Kroeger, Cynthia A. Phillips

Locality Sensitive Hashing for the Edit Distance *ISMB 2019*

Guillaume Marais, Dan DeBlasio, **Prashant Pandey**, and Carl Kingsford

†Small Refinements to the DAM Can Have Big Consequences for Data-Structure Design *SPAA 2019*

Michael A. Bender, Alex Conway, Martin Farach-Colton, William Jannen, Yizheng Jiao, Rob Johnson, Eric Knorr, Sara McAllister, Nirjhar Mukherjee, **Prashant Pandey**, Donald E. Porter, Jun Yuan, Yang Zhan

†The Online Event-Detection Problem *arXiv 2019*

Michael A. Bender, Jonathan W. Berry, Martin Farach-Colton, Rob Johnson, Thomas M. Kroeger, **Prashant Pandey**, Cynthia A. Phillips, Shikha Singh

An Efficient, Scalable and Exact Representation of High-Dimensional Color Information Enabled via de Bruijn Graph Search Problem *RECOMB 2019*

Fatemeh Almodaresi, **Prashant Pandey**, Michael Ferdman, Rob Johnson, Rob Patro

†Buffered Count-Min Sketch on SSD: Theory and Experiments *ESA 2018*

Mayank Goswami, Dzejla Medjedovic, Emina Mekic, **Prashant Pandey**

Mantis: A Fast, Small, and Exact Large-Scale Sequence-Search Index *RECOMB 2018 Cell Systems 2018*

Prashant Pandey, Fatemeh Almodaresi, Michael A. Bender, Michael Ferdman, Rob Johnson, and

*Joint first authors

†Author names in alphabetical order

Rob Patro

Rainbowfish: A Succinct Colored de Bruijn Graph Representation *WABI 2017*
Fatemeh Almodaresi, **Prashant Pandey**, and Rob Patro

deBGR: An Efficient and Near-Exact Representation of the Weighted de Bruijn Graph *ISMB 2017 BIOINFORMATICS 2017*
Prashant Pandey, Michael A. Bender, Rob Johnson, and Rob Patro

Squeakr: An Exact and Approximate k-mer Counting System *BIOINFORMATICS 2017*
Prashant Pandey, Michael A. Bender, Rob Johnson, and Rob Patro

A General-Purpose Counting Filter: Making Every Bit Count *SIGMOD 2017*
Prashant Pandey, Michael A. Bender, Rob Johnson, and Rob Patro
[Finalist for the most reproducible paper]
<http://db-reproducibility.seas.harvard.edu/papers/index.html>

A Fast x86 Implementation of Select *arXiv 2017*
Prashant Pandey, Michael A. Bender, and Rob Johnson

Writes Wrought Right, and Other Adventures in File System Optimization *TOS 2016*
Jun Yuan, Yang Zhan, William Jannen, **Prashant Pandey**, Amogh Akshintala, Kanchan Chandnani, Pooja Deo, Zardosht Kasheff, Michael Bender, Martin Farach-Colton, Rob Johnson, Bradley C. Kuszmaul, and Donald E. Porter

Optimizing Every Operation in a Write-Optimized File System *FAST 2016*
Jun Yuan, Yang Zhan, William Jannen, **Prashant Pandey**, Amogh Akshintala, Kanchan Chandnani, Pooja Deo, Zardosht Kasheff, Michael Bender, Martin Farach-Colton, Rob Johnson, Bradley C. Kuszmaul, and Donald E. Porter [Best Paper Award]

BetrFS: Write-Optimization in a Kernel File System *TOS 2015*
William Jannen, Jun Yuan, Yang Zhan, Amogh Akshintala, John Esmet, Yizheng Jiao, Ankur Mittal, **Prashant Pandey**, Phaneendra Reddy, Leif Walsh, Michael A. Bender, Martin Farach-Colton, Rob Johnson, Bradley C. Kuszmaul, and Donald E. Porter

BetrFS: A Right-Optimized Write-Optimized File System *FAST 2015*
William Jannen, Jun Yuan, Yang Zhan, Amogh Akshintala, John Esmet, Yizheng Jiao, Ankur Mittal, **Prashant Pandey**, Phaneendra Reddy, Leif Walsh, Michael A. Bender, Martin Farach-Colton, Rob Johnson, Bradley C. Kuszmaul, and Donald E. Porter [Runner up to Best Paper]

PATENTS

Instructions that Facilitate the Implementation of the Fork System Call in Processes using Software Guard Extensions *October 2018*
<https://patents.google.com/patent/US10089447B2/en>
Prashant Pandey, Mona Vij, Somnath Chakrabarti, Krystof C. Zmudzinski

Apparatus and Method For Implementing a Forked System Call in a System with a Protected Region *January 2018*
<https://patents.google.com/patent/US9870467B2/en>
Prashant Pandey, Mona Vij, Somnath Chakrabarti, Krystof C. Zmudzinski

INVITED TALKS

MetaGNN: Binning Metagenomic Contigs using GNN and Taxonomic Labelling
Workshop on DL for (meta)genomic sequence data, Lawrence Berkeley National Lab, July 2020

Timely Reporting of Heavy Hitters using External Memory
University of Maryland, College Park, MD, October 2019

Timely Reporting of Heavy Hitters using External Memory
IT University of Copenhagen, Copenhagen, Denmark, September 2019

Compact Representation of Annotated de Bruijn Graphs

Berkeley Lab, Berkeley CA, January 2018

deBGR: An Efficient and Near-Exact Representation of the Weighted de Bruijn Graph

Google Research, NY, September 2017

VMware Research, Palo Alto CA, Aug 2017

Intel Software Guard Extensions (SGX)

Sandia National Laboratories, Livermore CA, August 2015

CONFERENCE TALKS

VariantStore: A Space-Efficient and Fast Variant Search Index

ISMB 2020 talk/poster

Timely Reporting of Heavy Hitters using External Memory

SIGMOD 2020, Portland, OR

Small Refinements to the DAM Can Have Big Consequences for Data-Structure Design

SPAA 2019, Phoenix, AZ

Timely Reporting of Heavy Hitters using External Memory

Theoretical Foundations of Storage Systems 2019, Dagstuhl, Germany

Buffered Count-Min Sketch on SSD: Theory and Experiments

ESA 2018, Helsinki, Finland

Mantis: A Fast, Small, and Exact Large-Scale Sequence-Search Index

RECOMB 2018, Paris, France

Scheduling Problems in Write-Optimized Key-Value Stores

New Challenges in Scheduling Theory 20018, Aussois, France

deBGR: An Efficient and Near-Exact Representation of the Weighted de Bruijn Graph

ISMB 2017, Prague, Czech Republic

A General-Purpose Counting Filter: Making Every Bit Count

SIGMOD 2017, Chicago, IL

PROFESSIONAL SERVICE

- **Session chair:** *ALLENEX 2021*
- **Program Committee:** *RECOMB-Seq 2021, IPDPS 2021, ALLENEX 2021, EURO-PAR 2020, RECOMB-Seq 2020, ESA 2019*
- **Reviewer:** *Transactions on Parallel and Distributed Systems (TPDS), Transactions on Databases (TODS), Journal of Experimental Algorithms (JEA), IEEE Access, Oxford BIOINFORMATICS (2018, 2019, 2020)*
- **Subreviewer:** *ISMB 2021, STACS 2021, RECOMB 2020, WABI 2019, CIAC 2019*
- **Judge:** *Poster session RECOMB 2019*

TEACHING EXPERIENCE

Teaching Assistant, CS Dept, Stony Brook University

- CSE 548: Analysis of Algorithms *Fall 2015*
- CSE 535: Asynchronous Systems *Fall 2015*
- CSE 110: Introduction to Computer Science (Advanced Java) *Spring 2014*
- CSE 110: Introduction to Computer Science (Advanced Java) *Fall 2013*

REFERENCES

Reference letters can be requested via email.

- Prof. Michael A. Bender *Stony Brook University, NY*
- Senior Staff Researcher Rob Johnson *VMware Research, CA*
- Prof. Rob Patro *University of Maryland, College Park, MD*
- Prof. Carl Kingsford *Carnegie Mellon University, PA*
- Prof. Kathy Yelick *University of California Berkeley, CA*
- Prof. Martin Farach-Colton *Rutgers University, NJ*