Prashant Pandey

Contact Information

72 Central Campus Drive
Salt Lake City, UT - 84112
E-mail: prashant.pandey@utah.edu
Phone: (+1) 631-949-6948
Website https://prashantpandey.github.io
Google Scholar https://goo.gl/Fz82hB
Github https://github.com/prashantpandey/

Work Experience

University of Utah, Salt Lake City, UT — August 2022 - Present
Assistant Professor

VMware Research, Palo Alto, CA — August 2021 - July 2022
Research Scientist

Lawrence Berkeley Lab/UC Berkeley, Berkeley, CA — December 2019 - July 2021
Postdoctoral Research Fellow, Computational Research Division
Advisors: Prof. Kathy Yelick & Prof. Aydin Buluc

Carnegie Mellon University, Pittsburgh, PA — December 2018 - November 2019
Postdoctoral Associate, School of Computer Science
Advisor: Prof. Carl Kingsford

TIBCO Inc., Pune, India — July 2011 - June 2013
Software Developer, Cloud Platform

Education

Stony Brook University, Stony Brook, NY — December 2018
Ph.D. Computer Science
Thesis: Fast and Space-Efficient Maps: Shrinking Big Data Down to Size
Advisors: Prof. Michael Bender & Prof. Rob Johnson

University of Pune, Pune, India — August 2007 - June 2011
Bachelor of Engineering (BE), Information Technology

Internships

Google, Manhattan, NY — May 2017 - August 2017
Research Intern, Google Spanner

Google, Kirkland, WA — May 2016 - August 2016
Research Intern, Google Could Infrastructure

Intel Labs, Portland, OR — May 2015 - August 2015
Research Intern, Security and Privacy Lab

Intel Labs, Portland, OR — May 2014 - August 2014
Research Intern, Security and Privacy Lab

Awards and Achievements

- 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
  Ranked 1st in my college and 7^{th} across the University ($\sim$ 2000 students)
- **Academic Excellence Scholarship**, University of Pune. — 2009, 2010, 2011
- **Travel Fellowships**
  FAST 2015, FAST 2016, SIGMOD 2017, ISMB 2017, AlgoPARC 2017, RECOMB 2018, ESA 2018, Dagstuhl 2019

Conference Publications

BP-tree: Overcoming the Point-Range Operation Tradeoff for In-Memory B-trees *VLDB 2023*
Helen Xu, Amanda Li, Brian Wheatman, Manoj Marneni, **Prashant Pandey**

Singleton Sieving: Overcoming the Memory/Speed Trade-Off in Exascale k-mer Analysis *ACDA 2023*
Hunter McCoy, Steven Hofmeyr, Katherine Yelick, **Prashant Pandey**

Communication Optimization for Distributed Execution of Graph Neural Networks *IPDPS 2023*
Süreyya Emre Kurt, Jinghua Yan, Aravind Sukumaran-Rajam, **Prashant Pandey**, P. Sadayappan

High-Performance Filters for GPUs *PPOPP 2023*
Hunter McCoy, Steven Hofmeyr, Katherine Yelick, **Prashant Pandey**

Distance and Time Sensitive Filters for Similarity Search in Trajectory Datasets *APOCS 2023*
Madhav Narayan Bhat, Paul Cesaretti, Mayank Goswami, **Prashant Pandey**

IcebergHT: High Performance PMEM Hash Tables Through Stability and Low Associativity *SIGMOD 2023*
Prashant Pandey, Michael Bender, Alex Conway, Martin Farach-Colton, William Kuszmaul, Guido Tagliavini, Rob Johnson

Terrace: A Hierarchical Graph Container for Skewed Dynamic Graphs *SIGMOD 2021*
Prashant Pandey, Brian Wheatman, Helen Xu, Aydin Buluc

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

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

An Efficient, Scalable, and Exact Representation of High-Dimensional Color Information Enabled Using de Bruijn Graph Search *RECOMB 2019*
Fatemeh Almodaresi, **Prashant Pandey**, Michael Ferdman, Rob Johnson, Rob Patro

Locality Sensitive Hashing for the Edit Distance *ISMB 2019*
Guillaume Marçais, 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

[*]**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*
Prashant Pandey, Fatemeh Almodaresi, Michael A. Bender, Michael Ferdman, Rob Johnson, and Rob Patro

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

---

[*] Author names in alphabetical order

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**]

BϵtrFS: 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**]

Underlined - Utah student advisee.
All papers in 2023 are with Utah affiliation.

Journal Publications

Using Advanced Data Structures to Enable Responsive Security Monitoring *Cluster Computing 2022*
Janet Vorobyeva, Daniel R. Delayo, Michael A. Bender, Martin Farach-Colton, **Prashant Pandey**, Cynthia A. Phillips, Shikha Singh, Eric D. Thomas, Thomas M. Kroeger

An Incrementally-Updatable and Scalable System for Large-Scale Sequence Search using LSM-Trees *BIOINFORMATICS 2022*
Fatemeh Almodaresi, Jamshed Khan, Sergey Madaminov, Michael Ferdman, Rob Johnson, **Prashant Pandey**, and Rob Patro

VariantStore: an index for large-scale genomic variant search *Genome Biology 2021*
Prashant Pandey, Yinjie Gao, Carl Kingsford

*__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

Timely Reporting of Heavy Hitters using External Memory *TODS 2021*
Shikha Singh, **Prashant Pandey**, Michael A. Bender, Jonathan W. Berry, Martin Farach-Colton, Rob Johnson, Thomas M. Kroeger, Cynthia A. Phillips

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

Locality Sensitive Hashing for the Edit Distance *BIOINFORMATICS 2019*
Guillaume Marçais, Dan DeBlasio, **Prashant Pandey**, and Carl Kingsford

Mantis: A Fast, Small, and Exact Large-Scale Sequence-Search Index *Cell Systems 2018*
Prashant Pandey, Fatemeh Almodaresi, Michael A. Bender, Michael Ferdman, Rob Johnson, and Rob Patro

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

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

BεtrFS: 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

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

IcebergHT: High Performance PMEM Hash Tables Through Stability and Low Associativity
"From Big Data Theory to Big Data Practice" February 2023, Dagstuhl, Germany

Designing High-Performance and Feature Rich GPU Filters For Exascale Computing
"Joint PNNL-Utah Weekly HPC Seminar", September 2022

Scalability Challenges in Large-Scale Sequence Search
"Utah Center of Data Science (UCDS) Seminar Series", September 2022

Vector Quotient Filters: Overcoming the Time/Space Trade-Off in Filter Design
"Applied and Computational Discrete Algorithms (ACDA)", September 2022 Aussois, France

Time to Change Your Filter
Boston University, February 2022

Locality Sensitive Hashing for the Edit Distance
Northeastern University, February 2021

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

Timely Reporting of Heavy Hitters using External Memory
"Theoretical Foundations of Storage Systems" March 2019, Dagstuhl, Germany

Scheduling Problems in Write-Optimized Key-Value Stores
"New Challenges in Scheduling Theory" March 2018, Aussois, France

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

IcebergHT: High Performance PMEM Hash Tables Through Stability and Low Associativity
SIGMOD 2023, Seattle, USA

Terrace: A Hierarchical Graph Container for Skewed Dynamic Graphs
SIGMOD 2021, Xi'an, China

Vector Quotient Filters: Overcoming the Time/Space Trade-Off in Filter Design
SIGMOD 2021, Xi'an, China

VariantStore: A Space-Efficient and Fast Variant Search Index
ISMB 2020 Virtual conference

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

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

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

Funding

- **DOE: Exascle Computing Project: High Performance GPU Filters**
  - Role: Utah PI (Joint with UC Berkeley/Lawrence Berkeley National Lab)
  - Utah portion: $250,000.00
  - Duration: October 2022 – September 2023
- **NSF: PPoSS: Large: A Compact, Dynamic, and Distributed Data Structure Library for Computational Biology** [Under Submission]
  - Role: Utah PI (Joint with UC Davis/Rutgers/StonyBrook)
  - Utah portion: $1,250,000.00
  - Duration: October 2023 – September 2028

Students

- **Advising**
  - Hunter McCoy Ph.D. CS — Started Fall 2022
  - Yuvraj Chaesetti Ph.D. CS — Started Fall 2022
  - Benwei Shi Ph.D. CS (Co-advise with Prof. Jeff) — Started Fall 2023
  - Jinghua Yan Ph.D. CS (Co-advise with Prof. Saday) — Started Fall 2023
  - Susmitha Raja MS CS (Research Assistant) — Started Fall 2022
  - Medha Kalkur MS CS (Research Assistant) — Started Fall 2022
  - Manoj Marneni MS CS (Research Assistant) — Started Fall 2022
  - Pranjal Patil MS CS (Independent Study) — Started Fall 2022
  - Alex Tokita BS CS (UROP Scholar) — Started Fall 2022
- **Committee Member**
  - Ankit Bhardwaj — Ph.D. CS
  - Mahesh Lakshminarasimhan — Ph.D. CS
  - AnanthKrishna Prasad — Ph.D. CS
  - Amit Samanta — Ph.D. CS
  - LeAnn Lindsey — Ph.D. CS
  - Todd Thornley — M.S CS

Professional Service

- **Workshop Organiser**:
  2023: Workshop on Filter Data Structures SPAA 2023
- **Program Committee**:
  2024: SIGMOD, PPOPP
  2023: VLDB, SIGMOD ARC, SPAA, IPDPS, ESA, IEEE BigData
  2022: IEEE BigData, ACM BCB, APOCS, IPDPS
  2021: ACDA, RECOMB-Seq, IPDPS, ALENEX
  2020: EURO-PAR, RECOMB-Seq
  2019: ESA
- **Journals**:
  Transactions on Parallel and Distributed Systems (TPDS) 2020
  Transactions on Databases (TODS) 2018
  Journal of Experimental Algorithms (JEA) 2019
  IEEE Access 2019, 2021
  Oxford BIOINFORMATICS (2018, 2019, 2020)
  Journal of Computational Biology (JCB) 2021, 2022, 2023
  Transactions on Knowledge and Data Engineering (TKDE) 2021, 2022
- **Subreviewer**:
  SC 2023
  SODA 2023
  FAST 2022
  ISMB 2021, STACS 2021
  RECOMB 2020
  WABI 2019, CIAC 2019
- **Session chair**: ALENEX 2021
- **Judge**: Poster session RECOMB 2019

Department Service

- **Director:** Data Science Graduate Certificate Program — Spring 2023 –
- **Organizer:** Data Science Seminar Series — Fall 2023 –
- **Organiser**: KSoC Annual Sports Event — Spring 2023 –
- Graduate Admissions Committee, School of Computing, University of Utah — 2023
- **Organizer CS Colloquium Series**: — Fall 2022 –
  - Harsha Vardhan Simhadri (Microsoft Research) — March 2023
  - Guillaume Marçais (Carnegie Mellon University) — February 2023
  - Pedro Pedreira (Meta Velox) — November 2022
  - David Shue (Google SpannerX) — October 2022
  - Rob Johnson (VMware Research) — October 2022

Teaching

Assistant Professor, School of Computing, University of Utah

- CS 6530: Adv. Database Systems — *Fall 2023*
- CS 6968/5968: Data Str & Alg Scalable Comp — *Spring 2023*
- CS 6530: Adv. Database Systems — *Fall 2022*

TA 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*

Press Articles on Research

A general purpose counting filter: making every bit count. The Morning Paper. *August 2017*
Link: `https://goo.gl/nReGcF`

Scaling Computational Biology at VMware. VMware Research Blog *April 2018*
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. *July 2018*
Link: `https://goo.gl/LJopwR`