Prashant Pandey

Contact Information

440 Huntington Avenue
Boston, MA - 02115
p.pandey@northeastern.edu
(+1) 631-949-6948
Website https://prashantpandey.github.io
Google Scholar https://goo.gl/Fz82hB
Github https://github.com/prashantpandey/

Work Experience

Northeastern University, Boston, MA — January 2025 - Present
Assistant Professor

University of Utah, Salt Lake City, UT — August 2022 - December 2024
Assistant Professor

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

Education

UC Berkeley/Berkeley Lab, 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

Stony Brook University, Stony Brook, NY — August 2013 - December 2018
Ph.D. Computer Science
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

- IEEE CS TCHPC Early Career Researchers Award for Excellence in High Performance Computing [Supercomputing SC 2023] — 2023
- Catacosinos Fellowship for the most impactful research at Stony Brook University — 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

Funding

NSF: Elements: Real-Time, Incremental, and Sustainable Sequence Search over SRA July 2025
Role: Solo PI
Amount: $600,000

One Utah Data Science Hub Seed Award March 2024
Scalable and Information-Rich Sequence Search over SRA for Advanced Biological Analyses
Role: Solo PI
Amount: $50,000

NSF: CAREER: Practical Adaptive Filters and Applications June 2024
Role: Solo PI
Amount: $607,746

DOE: Exascale Computing Project: High Performance GPU Filters October 2022
Role: Utah PI (Joint with UC Berkeley/Lawrence Berkeley National Lab)
Amount: $250,000

Conference Publications

FaSTCC: Fast Sparse Tensor Contractions on CPUs *SC 2025*
Saurabh Raje, Hunter McCoy, Atanas Rountev, **Prashant Pandey**, P. Sadayappan

A Locality-Optimized In-Memory B-Skiplist *ICPP 2025*
Yicong Luo, Senhe Hao, Brian Wheatman, **Prashant Pandey**, Helen Xu

Evaluating Learned Indexes for External-Memory Joins *ACDA 2025*
Yuvaraj Chesetti, **Prashant Pandey**

Zombie Hashing: Reanimating Tombstones in a Graveyard *SIGMOD 2025*
Yuvaraj Chesetti, Benwei Shi, Jeff M. Phillips, **Prashant Pandey**

Adaptive Quotient Filters *SIGMOD 2025*
Richard Wen, Hunter McCoy, David Tench, Guido Tagliavini, Michael A. Bender, Alex Conway, Martin Farach-Colton, Rob Johnson, **Prashant Pandey**

BYO: A Unified Framework for Benchmarking Large-Scale Graph Containers *VLDB 2024*
Brian Wheatman, Xiaojun Dong, Zheqi Shen, Laxman Dhulipala, Jakub Łącki, **Prashant Pandey**, Helen Xu

Beyond Bloom: A Tutorial on Future Feature-Rich Filters *SIGMOD 2024*
Prashant Pandey, Martin Farach-Colton, Niv Dayan, Huanchen Zhang

IONIA: Efficient Replication for SSD-based Write-Optimized KV Stores *FAST 2024*
Yi Xu, Henry Zhu, **Prashant Pandey**, Alex Conway, Rob Johnson, Ramnatthan Alagappan, Aishwarya Ganesan

Gallatin: A vEB Tree-Based GPU Memory Manager *PPOPP 2024*
Hunter McCoy, **Prashant Pandey**

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**

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

High-Performance Filters for GPUs *PPOPP 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

Singleton Sieving: Overcoming the Memory/Speed Trade-Off in Exascale k-mer Analysis *ACDA 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**

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

deBGR: An Efficient and Near-Exact Representation of the Weighted de Bruijn Graph *ISMB 2017*

Prashant Pandey, Michael A. Bender, 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: Most Reproducible Paper]**

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 - Direct student advisee.

Author names in alphabetical order. I am lead author.

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

Press Articles on Research

Dangling Pointers Blog covered ZombieHT paper from SIGMOD 2025 — August 2025
Link: `https://shorturl.at/UE0JN`

NSF Career award was recently covered by Khoury News — March 2025
Link: `https://shorturl.at/Ze7aO`

Our IcebergHT paper from SIGMOD 2023 featured in Quanta Magazine — February 2024
Link: `https://shorturl.at/dyAYZ`

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

Scaling Computational Biology at VMware. (Link: `https://shorturl.at/lpLR6`) — April 2018

Finding a Needle in a Field of Haystacks. Cell Systems publishes research on Mantis — July 2018
Link: `https://goo.gl/LJopwR`

Invited Talks

Zombie hashing — August 2025
"Software Performance Engineering", Dagstuhl, Germany 2025

Vector Search for Large-Scale Genomic Discovery — July 2025
Invited Keynote *at the 1st Workshop on Vector Databases (VecDB) at ICML 2025*

From Hash Tables to B-trees: Rethinking Core Data Structures for Scalable Performance May 2025
IISc Banglore and AMD Research India

Data Systems at Scale: Scaling Up by Scaling Down and Out — March 2025
Computer Science Department Colloquium, Williams College, MA

Parallel Graph Processing and Future Challenges — March 2025
Invited Keynote *at Principles and Practice of Parallel Programming (PPoPP) 2025, Las Vegas US*

Fusing Theory and Practice of Graph Algorithms — February 2025
The Institute for Computational and Experimental Research in Mathematics, Brown University, US

Adaptive Quotient Filters — July 2024
Theoretical Foundations of Nonvolatile Memory, Shonan Japan

Designing High-Performance In-Memory Indexes — February 2024
Northwest Database Society Annual Meeting, Google Kirkland

Designing High-Performance In-Memory Indexes — October 2023
Database Seminar Series, Georgia Tech

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

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

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

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

Time to Change Your Filter — February 2022
Boston University

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

MetaGNN: Binning Metagenomic Contigs using GNN and Taxonomic Labelling — July 2020

"Workshop on DL for (Meta)Genomic Sequence Data", Lawrence Berkeley National Lab

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

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

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

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

Compact Representation of Annotated de Bruijn Graphs January 2018
Berkeley Lab, Berkeley CA

deBGR: An Efficient Representation of the Weighted de Bruijn Graph Spetember 2017
Google Research, NY VMware Research, Palo Alto CA

Intel Software Guard Extensions (SGX) August 2015
Sandia National Laboratories, Livermore CA

Conference Talks

IcebergHT: High Performance 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 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

- **Program Committee**:

SIGMOD, VLDB, ICDE, ALENEX	2026
SIGMOD, VLDB, EDBT, ICDE, PPoPP, ACM ICS, ACDA	2025
SIGMOD, VLDB, PPoPP, IPDPS, IEEE BigData	2024
VLDB, SIGMOD ARC, SPAA, IPDPS, ESA, IEEE BigData	2023
IEEE BigData, ACM BCB, APOCS, IPDPS	2022
ACDA, RECOMB-Seq, IPDPS, ALENEX	2021
EURO-PAR, RECOMB-Seq	2020
ESA	2019

- **Workshop Chair**:

Workshop on Filter Data Structures SPAA (FCRC 2023)	2023

- **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**:

SODA	2024
SC, SODA	2024
FAST	2022
ISMB, STACS	2021
RECOMB	2020
WABI, CIAC	2019

- **Session chair**: ALENEX 2021
- **Judge**: Poster session RECOMB 2019

Department Service

- **Director:** Data Science Graduate Certificate Program — Spring 2023 –
- **Organizer:** Utah Center for Data Science Lecture Series `https://datascience.utah.edu/seminar.html` — Fall 2023 –
- **Organiser**: KSoC Annual Sports Event `https://users.cs.utah.edu/~pandey/ksocsportsevent/2023/` — Spring 2023 –
- **Graduate Admissions Committee** — 2023, 2024
- **Organizer KSoC Colloquium Series** — Fall 2022 –

Students

- **Advising**
  - Diandre Sabale Ph.D. CS — Started Fall 2024
  - Zikun Wang Ph.D. CS — Started Fall 2025
  - Hunter McCoy Ph.D. CS — Started Fall 2022
  - Yuvraj Chaesetti Ph.D. CS — Started Fall 2023
  - Aaditya Rangarajan MS CS (Independent Study) — Started Spring 2024
  - Ang Li MS CS (Independent Study) — Started Spring 2024
  - Benwei Shi Ph.D. CS (Co-advise with Prof. Jeff) — Spring 2023 – Fall 2023
  - Jinghua Yan Ph.D. CS (Co-advise with Prof. Saday) — Started Fall 2023
  - Susmitha Raja MS CS (Research Assistant) — Fall 2022 – Summer 2023
  - Medha Kalkur MS CS (Research Assistant) — Fall 2022 – Summer 2023
  - Manoj Marneni MS CS (Research Assistant) — Fall 2022 – Spring 2023
  - Pranjal Patil MS CS (Independent Study) — Spring 2023
  - Alex Tokita BS CS (UROP Scholar) — Fall 2022
- **Committee Member**
  - Ankit Bhardwaj — Ph.D. CS
  - Sayef Azad Sakin — Ph.D. CS
  - Mahesh Lakshminarasimhan — Ph.D. CS
  - AnanthKrishna Prasad — Ph.D. CS
  - Amit Samanta — Ph.D. CS

- LeAnn Lindsey — Ph.D. CS
- Chris Harker — Ph.D. CS
- Todd Thornley — M.S CS

Teaching

Assistant Professor, School of Computing, University of Utah

- CS 7270/4973: Adv. Database Systems Seminar — Fall 2025
- CS 7280/4973: Data Str & Alg for Scalable Comp — Spring 2025
- CS 6530: Adv. Database Systems — Fall 2024
- CS 6530: Adv. Database Systems — Fall 2023
- CS 6968/5968: Data Str & Alg for Scalable Comp — Spring 2023
- CS 6530: Adv. Database Systems — Fall 2022