

Prashant Pandey

CONTACT INFORMATION

72 Central Campus Drive
Salt Lake City, UT - 84112

Website

Google Scholar

Github

prashant.prashn@gmail.com

(+1) 631-949-6948

<https://prashantpandey.github.io>

<https://goo.gl/Fz82hB>

<https://github.com/prashantpandey/>

WORK EXPERIENCE

University of Utah, Salt Lake City, UT
Assistant Professor

August 2022 - Present

VMware Research, Palo Alto, CA
Research Scientist

August 2021 - July 2022

TIBCO Inc., Pune, India

July 2011 - June 2013

Software Developer, Cloud Platform

EDUCATION

UC Berkeley/Berkeley Lab, Berkeley, CA
Postdoctoral Research Fellow, Computational Research Division
Advisors: Prof. Kathy Yelick & Prof. Aydin Buluc

December 2019 - July 2021

Carnegie Mellon University, Pittsburgh, PA
Postdoctoral Associate, School of Computer Science
Advisor: Prof. Carl Kingsford

December 2018 - November 2019

Stony Brook University, Stony Brook, NY
Ph.D. Computer Science

August 2013 - December 2018

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
Research Intern, Google Spanner

May 2017 - August 2017

Google, Kirkland, WA
Research Intern, Google Cloud Infrastructure

May 2016 - August 2016

Intel Labs, Portland, OR
Research Intern, Security and Privacy Lab

May 2015 - August 2015

Intel Labs, Portland, OR
Research Intern, Security and Privacy Lab

May 2014 - August 2014

AWARDS AND ACHIEVEMENTS

- **IEEE CS TCHPC Early Career Researchers Award for Excellence in High Performance Computing [SC 2023]** 2023
- **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 7th across the University (~ 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: CAREER: Practical Adaptive Filters and Applications
Utah portion: \$607,746

June 2024

One Utah Data Science Hub Seed Award

March 2024

Utah portion: \$50,000

DOE: Exascale Computing Project: High Performance GPU Filters

October 2022

Role: Utah PI (Joint with UC Berkeley/Lawrence Berkeley National Lab)

Utah portion: \$250,000

CONFERENCE PUBLICATIONS

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 Olikar, 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. Kroegeer, 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**

Author names in alphabetical order. I am lead author.

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**]

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**]

Underlined - Utah student advisee.

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

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

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

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 <i>IT University of Copenhagen, Copenhagen, Denmark</i>	September 2019
Timely Reporting of Heavy Hitters using External Memory <i>"Theoretical Foundations of Storage Systems", Dagstuhl, Germany</i>	March 2019
Scheduling Problems in Write-Optimized Key-Value Stores <i>"New Challenges in Scheduling Theory", Aussois, France</i>	March 2018
Compact Representation of Annotated de Bruijn Graphs <i>Berkeley Lab, Berkeley CA</i>	January 2018
deBGR: An Efficient Representation of the Weighted de Bruijn Graph <i>Google Research, NY VMware Research, Palo Alto CA</i>	September 2017
Intel Software Guard Extensions (SGX) <i>Sandia National Laboratories, Livermore CA</i>	August 2015

CONFERENCE TALKS

IcebergHT: High Performance Hash Tables Through Stability and Low Associativity <i>Seattle, USA</i>	SIGMOD 2023
Terrace: A Hierarchical Graph Container for Skewed Dynamic Graphs <i>Xi'an, China</i>	SIGMOD 2021
Vector Quotient Filters: Overcoming the Time/Space Trade-Off in Filter Design <i>Xi'an, China</i>	SIGMOD 2021
VariantStore: A Space-Efficient and Fast Variant Search Index <i>Virtual conference</i>	ISMB 2020
Timely Reporting of Heavy Hitters using External Memory <i>Portland, OR</i>	SIGMOD 2020
Small Refinements to the DAM Can Have Big Consequences for Data-Structure Design <i>Phoenix, AZ</i>	SPAA 2019
Buffered Count-Min Sketch on SSD: Theory and Experiments <i>Helsinki, Finland</i>	ESA 2018
Mantis: A Fast, Small, and Exact Large-Scale Sequence-Search Index <i>Paris, France</i>	RECOMB 2018
deBGR: An Efficient Representation of the Weighted de Bruijn Graph <i>Prague, Czech Republic</i>	ISMB 2017
A General-Purpose Counting Filter: Making Every Bit Count <i>Chicago, IL</i>	SIGMOD 2017

STUDENTS

- **Advising**

• Hunter McCoy Ph.D. CS	Started Fall 2022
• Yuvraj Chaesetti Ph.D. CS	Started Fall 2022
• 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

PROFESSIONAL SERVICE

- **Workshop Organiser:**
 - Workshop on Filter Data Structures SPAA (FCRC 2023) 2023
- **Program Committee:**
 - SIGMOD, VLDB, EDBT 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
- **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 –

TEACHING

Assistant Professor, School of Computing, University of Utah

- 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

TA EXPERIENCE

Teaching Assistant, CS Dept, Stony Brook University

- CSE 548: Analysis of Algorithms
- CSE 535: Asynchronous Systems
- CSE 110: Introduction to Computer Science (Advanced Java)

Fall 2015
Fall 2015
Spring 2014, Fall 2023

PRESS ARTICLES ON RESEARCH

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/1pLR6>)

April 2018

Finding a Needle in a Field of Haystacks. Cell Systems publishes research on Mantis

July 2018

Link: <https://goo.gl/LJopwR>