Benjamin (Min) Wesley Priest (they/them)

CONTACT INFORMATION

Postdoctoral Researcher
Center for Applied Scientific Computing
Lawrence Livermore National Laboratory

Cell: +1-937-681-1935
E-mail: priest2@llnl.gov

RESEARCH INTERESTS

Efficient analysis of large, dynamic datasets: sketching, streaming algorithms, machine learning, high performance computing, graph algorithms, numerical linear algebra, compressed sensing, graph theory, optimization, network analysis, and theory of deep learning.

EDUCATION

Thayer School of Engineering at **Dartmouth College**, Hanover, VT, USA
Ph.D., Engineering **09/2015 – 02/2019**
- Advisor: Professor George Cybenko
- Thesis: Sublinear Approximations of Vertex Centrality in Evolving Graphs

The Ohio State University, Columbus, OH, USA
B.S., Mathematics **09/2007 – 06/2011**
B.S., Computer and Information Science **09/2007 – 06/2011**

RESEARCH EXPERIENCE

Lawrence Livermore National Laboratory, Livermore, CA, USA
Center for Applied Scientific Computing. Supervisors: Geoff Sanders, Michael Schneider and Roger Pearce

Computing Scientist **02/2021 – present**
PI and Co-I of multiple research projects investigating scalable graph analytics, machine learning, and statistical modeling on High-Performance Computing (HPC) systems. Supervised 1 postdoc and more than a dozen graduate students. Selected research contributions include novel algorithms and software for scalable Gaussian process (GP) estimation [1], cosmology, climate, and space domain modeling [2], distributed subspace embedding and sketches [3], and distributed K nearest neighbors.

Postdoctoral Researcher **04/2019 – 02/2021**
Developed novel sketching algorithms to cluster [3] and perform local query approximation [4] massive graphs on HPC. Solved reinforcement learning [5], image classification [6], and quantum machine learning [7] problems using GPs and neural kernels.

Computation Student Intern **05/2018 – 01/2019**
Designed novel HPC communication library to accelerate non-traditional communications [8]. Used cardinality sketches to estimate local triangle counts in distributed graphs [9].

Dartmouth College, Hanover, NH, USA
Thayer School of Engineering. Advisor: Professor George Cybenko

Research and Teaching Assistant **09/2015 – 02/2019**
Invented streaming approximation algorithms for several centrality indices on massive graphs using sketches. Designed game and graph-theoretic models for advanced persistent threats in cyber defense. Taught courses in machine learning and lead a team of TAs.

MIT Lincoln Laboratory, Lexington, MA, USA
Cyber Analytics and Decision Systems. Supervisor: Dr. Kevin M. Carter

Assistant Research Scientist **08/2011 – 07/2015**
Modeled computer networks using novel machine learning algorithms. Developed multi-agent systems for high-fidelity network simulations and cyber defense evaluation.

Air Force Institute of Technology, Wright-Patterson Air Force Base, OH, USA
Program Encryption Group. Supervisor: Professor J. Todd McDonald

Engineering Technician GS-05 **Summer, 2008 & 2009**
Developed encryption metrics for circuits using abstract interpretation semantic models

Technical Expertise

Mathematics
- Applied Mathematics
- Real Analysis
- Measure Theory
- Graph Theory
- Combinatorics

Computer Science and Engineering
- Distributed & parallel algorithms
- Streaming algorithms & sketching
- Data structures

Data Science and Processing
- Probability & Random Variables
- Statistics & Estimation
- Machine learning & deep learning
- Numerical Optimization
- Stochastic Processes
- Information Theory
- Communication Theory

Programming and Scripting Languages
- C/C++, Python, Bash,
- Julia, Java, R, MATLAB

Distributed Computing
- MPI, Hadoop MapReduce, Lustre

Analytical Software
- Keras, TensorFlow, PyTorch, Mathematica

Utility Software
- Git, GitHub/Gitlab/Bitbucket
- LaTeX, BibTeX
- Microsoft, LibreOffice, Google Suite

Operating Systems
- Apple OS X
- Linux, RedHat, and other UNIX variants

Interpersonal
- Teamwork and communication
- Leadership and mentoring
- Public and technical speaking

Peer-Reviewed Conference Publications

[10] Rafael Bidese, Chinedu Eleh, Yunli Zhang, Roberto Molinari, Nedret Billor, **Benjamin W Priest**, Imène R Goumiri, Amanda L Muyskens, and Alec M Dunton. Stellar blend image classification using computationally efficient Gaussian processes (MuyGPs). *arXiv preprint arXiv:2208.14592*, 2022

[11] Killian Wood, Alec M Dunton, Amanda Muyskens, and **Benjamin W Priest**. Scalable Gaussian process hyperparameter optimization via coverage regularization. *arXiv preprint arXiv:2209.11280*, 2022

[12] Imène R Goumiri, Alec M Dunton, Amanda L Muyskens, **Benjamin W Priest**, and Robert E Armstrong. Light curve completion and forecasting using fast and scalable Gaussian processes (MuyGPs). *arXiv preprint arXiv:2208.14592*, 2022

[13] Alec M Dunton, **Benjamin W Priest**, and Amanda Muyskens. Fast Gaussian process posterior mean prediction via local cross validation and precomputation. *arXiv preprint arXiv:2205.10879*, 2022

[14] Trevor Steil, Tahsin Reza, Keita Iwabuchi, **Benjamin W Priest**, Geoffrey Sanders, and Roger Pearce. TriPoll: computing surveys of triangles in massive-scale temporal graphs with metadata. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 1–12, 2021

[6] Imène R Goumiri, Amanda L Muyskens, Michael D Schneider, **Benjamin W Priest**, and Robert E Armstrong. Star-galaxy separation via Gaussian processes with model reduction. In *2020 Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS)*, 2020

[3] **Benjamin W Priest**, Alec Dunton, and Geoffrey Sanders. Scaling graph clustering with distributed sketches. In *2020 IEEE High Performance Extreme Computing Conference (HPEC)*, pages 1–7. IEEE, 2020

[5] Imène R Goumiri, **Benjamin W Priest**, and Michael D Schneider. Reinforcement learning via Gaussian processes with neural network dual kernels. In *2020 IEEE Conference on Games (CoG)*, pages 1–8. IEEE, 2020

[15] Trevor Steil, Scott McMillan, Geoffrey Sanders, Roger Pearce, and **Benjamin W Priest**. Kronecker graph generation with ground truth for 4-cycles and dense structure in bipartite graphs. In *2020 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, pages 237–246. IEEE, 2020

[16] Roger Pearce, Trevor Steil, **Benjamin W Priest**, and Geoffrey Sanders. One quadrillion triangles queried on one million processors. In *2019 IEEE High Performance Extreme Computing Conference (HPEC)*, pages 1–5. IEEE, 2019

[8] **Benjamin W Priest**, Trevor Steil, Geoffrey Sanders, and Roger Pearce. You've got mail (ygm): Building missing asynchronous communication primitives. In *2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, pages 221–230. IEEE, 2019

[17] Trevor Steil, **Benjamin W Priest**, Geoffrey Sanders, Roger Pearce, Timothy La Fond, and Keita Iwabuchi. Distributed kronecker graph generation with ground truth of many graph properties. In *2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, pages 251–260. IEEE, 2019

[9] **Benjamin W Priest**, Roger Pearce, and Geoffrey Sanders. Estimating edge-local triangle count heavy hitters in edge-linear time and almost-vertex-linear space. In *2018 IEEE High Performance extreme Computing Conference (HPEC)*, pages 1–7. IEEE, 2018

[18] Luan Huy Pham, Massimiliano Albanese, and **Benjamin W Priest**. A quantitative framework to model advanced persistent threats. In *ICETE (2)*, pages 448–459, 2018

[19] **Benjamin W Priest**, Era Vuksani, Neal Wagner, Brady Tello, Kevin M Carter, and William W Streilein. Agent-based simulation in support of moving target cyber defense technology development and evaluation. In *Proceedings of the 18th Symposium on Communications & Networking*, pages 16–23, 2015

[20] Kevin M Carter, Rajmonda S Caceres, and **Benjamin W Priest**. Characterization of latent social networks discovered through computer network logs. In *Networks in the Social and Information Sciences workshop of the 29th Annual Conference on Neural Information Processing Systems*, 2015

[21] **Benjamin W Priest**and Kevin M Carter. Characterizing latent user interests on enterprise networks. In *The Twenty-Seventh International Flairs Conference*, 2014

[22] Kevin M Carter, Rajmonda S Caceres, and **Benjamin W Priest**. Latent community discovery through enterprise user search query modeling. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*, pages 871–874, 2014

[23] Kevin Gold, Zachary J Weber, **Benjamin W Priest**, Josh Ziegler, Karen Sittig, William W Streilein, and Mark Mazumder. Modeling how thinking about the past and future impacts network traffic with the gosmr architecture. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 127–134. Citeseer, 2013

[24] **Benjamin W Priest**and Kevin Gold. Utility discounting explains informational website traffic patterns before a hurricane. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 53–54, 2013

[25] Kevin Gold, **Benjamin W Priest**, and Kevin M Carter. An expectation maximization approach to detecting compromised remote access accounts. In *The Twenty-Sixth International FLAIRS Conference*, 2013

Peer-Reviewed Journal Publications

[26] James J Buchanan, Michael D Schneider, Robert E Armstrong, Amanda L Muyskens, **Benjamin W Priest**, and Ryan J Dana. Gaussian process classification for galaxy blend identification in LSST. *The Astrophysical Journal*, 924(2):94, 2022

[2] Amanda L Muyskens, Imène R Goumiri, **Benjamin W Priest**, Michael D Schneider, Robert E Armstrong, Jason Bernstein, and Ryan Dana. Star–galaxy image separation with computationally efficient Gaussian process classification. *The Astronomical Journal*, 163(4):148, 2022

ArXiv Papers

[1] Amanda Muyskens, **Benjamin W Priest**, Imène Goumiri, and Michael Schneider. MuyGPs: Scalable Gaussian process hyperparameter estimation using local cross-validation. *arXiv preprint arXiv:2104.14581*, 2021

[7] Matthew Otten, Imène R Goumiri, **Benjamin W Priest**, George F Chapline, and Michael D Schneider. Quantum machine learning using Gaussian processes with performant quantum kernels. *arXiv preprint arXiv:2004.11280*, 2020

[4] **Benjamin W Priest**. Degreesketch: Distributed cardinality sketches on massive graphs with applications. *arXiv preprint arXiv:2004.04289*, 2020

Working Papers

[27] **Benjamin W Priest**, Trevor Steil, Geoffrey Sanders, and Keita Iwabuchi. Fast approximate k-nearest neighbors recovery of graphs using subspace embeddings on graph exponents. **In preparation**

[28] **Benjamin W Priest**, Robert E Armstrong, and Gregory Sallaberry. Modeling weak lensing shear fast approximate k-nearest neighbors recovery of graphs using subspace embeddings on graph exponents. **In preparation**

[29] Imène R Goumiri, Amanda L Muyskens, **Benjamin W Priest**, and Robert E Armstrong. Light curve forecasting and anomaly detection using scalable, anisotropic, and heteroscedastic Gaussian process models. **In preparation**

[30] Amanda Muyskens, **Benjamin W Priest**, Imène R Goumiri, and Michael D Schneider. An analysis of the sensitivity of kernel hyperparameters on the kriging weights. **In preparation**

Other Conference Publications

[31] **Benjamin W Priest**and George Cybenko. Approximating centrality in evolving graphs: toward sublinearity. In *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security, Defense, and Law Enforcement Applications XVI*, volume 10184, pages 58–66. SPIE, 2017

[32] **Benjamin W Priest**and George Cybenko. Efficient inference of hidden markov models from large observation sequences. In *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security, Defense, and Law Enforcement Applications XV*, volume 9825, pages 179–187. SPIE, 2016

Book Chapters

[33] Geoffrey Sanders, Roger Pearce, **Benjamin W Priest**, and Trevor Steil. Massive-scale distributed triangle enumeration and applications. **In preparation**

[34] **Benjamin W Priest**, George Cybenko, Satinder Singh, Massimiliano Albanese, and Peng Liu. Online and scalable adaptive cyber defense. In *Adversarial and Uncertain Reasoning for Adaptive Cyber Defense*, pages 232–261. Springer, 2019

Conference Talks

Benjamin W Priest. Scaling Graph Clustering with Distributed Sketches. At: *2020 High Performance Extreme Computing Conference*, HPEC 2020. Waltham, CA, USA (virtual conference), 21–25 September 2020.

Benjamin W Priest. Approximating centrality in evolving graphs: toward sublinearity. At: *2017 SPIE Defense + Security Conference*, SPIE D+S. Anaheim, CA, USA, 9–13 April 2017.

Benjamin W Priest. Efficient Inference of hidden Markov models from large observations sequences. At: *2016 SPIE Defense + Security Conference*, SPIE D+S. Anaheim, CA, USA, 17–21 April 2016.

Benjamin W Priest. Agent-based simulation in support of moving target cyber defense technology development and evaluation. At: *18th Symposium on Communications & Networking, 2015 ACM Spring Simulation Multi-Conference*, CNS/SpringSim. Alexandria, VA, USA, 12–15 April 2015.

Benjamin W Priest. Characterizing latent user interests on enterprise networks. At: *2014 International Florida Artificial Intelligence Research Society Conference*, FLAIRS. Pensacola Beach, FL, USA, 21–23 May 2014.

Invited Talks

Benjamin W Priest. High-fidelity enterprise network emulation using the GOSMR architecture. In: *2014 MIT Lincoln Laboratory Cyber and Net-Centric Workshop*, CNW. June, 2014.

Poster Presentations

Benjamin W Priest. Estimating edge-local triangle count heavy hitters in edge-linear time and almost-vertex-linear space. At: *GraphChallenge Workshop at the IEEE High Performance Extreme Computing Conference*, HPEC. 25–27 September 2018.

Benjamin W Priest. Efficient Sublinear Estimation of Local Triangle Count Heavy Hitters. At: *2018 Summer Student Poster Symposium at Lawrence Livermore National Laboratory*. 9 August 2018.

Benjamin W Priest. Characterization of latent social networks discovered through computer network logs. At: *Networks in the Social and Information Sciences workshop of the 29th Annual Conference on Neural Information Processing Systems*, NIPS. Montreal, Canada, 12 December 2015.

Benjamin W Priest. Utility discounting explains informational website traffic patterns before a hurricane. At: *22nd International World Wide Web Conference*, WWW. 2013. Rio de Janeiro, Brazil, 13–17 May 2013.

Grants

In Preparation

- Co-I, "Science with LSST Year 1 Data", LLNL LDRD SI, $2,000,000/yr. FY 25-27.

Awarded

- PI, "Hierarchical Graph-Based Clustering in Distributed Memory", LLNL LDRD 23-ERD-044, $550,000/yr. October 1, 2023 to September 30, 2026.
- Co-I, "HPC-Enabled Detection System for Petabyte Scale Astronomy Surveys", LLNL LDRD 23-ERD-044, $700,000/yr. October 1, 2022 to September 30, 2025.
- Co-I, "MuyGPs: Non-Stationary Gaussian Processes at HPC Scales", LLNL LDRD 23-ERD-028, $850,000/year. October 1, 2021 to September 30, 2024.

- Co-I, "Scalable Uncertainty Quantification Using Gaussian Processes Surrogate Models", LLNL LDRD 21-FS-037, $100,000. January 1, 2021 to September 30, 2021.
- Co-I, "Interactive Exploratory Graph-Enabled Data Analytics at HPC Scales", LLNL LDRD 21-ERD-020, $500,000/year. October 1, 2020 to September 30, 2022.
- Co-PI, "Scalable Approximate Graph Clustering", LLNL LDRD 20-FS-037, $150,000. February 1, 2020 to September 30, 2020.

Declined

- PI, "Scalable and Highly Concurrent Network Science via Distributed Sketching", DOE ASCR DE-FOA-0002722, $800,000/yr. FY 23-25.
- Co-I, "Probabilistic AI Pipeline Modules for Rubin LSST Dark Energy", DOE ASCR DE-FOA-0002705, $1,000,000/yr. FY 23-25.
- Co-PI, "Scalable Single-Pass Compressive Autoencoders", LLNL LDRD Feasibility Study, $150,000. FY 23.
- Co-I, "EpochGrafts: Relational Data Fusion via Dynamic Graph Analysis", LLNL LDRD ER, $500,000/year. FY 22-24.
- PI, "Scalable Non-stationary Approximate Gaussian Processes", DOE ASCR DE-FOA-0002493, $800,000/year. FY 22-24.
- PI, "Distributed Memory Sketching Algorithms at HPC Scales", DOE ASCR DE-FOA-0002497, $400,000/year. FY 22-23.
- PI, "`croquis`: Distributed Subspace Embeddings for High Performance Computing", LLNL Tech Base, $100,000. FY 21-22.

MENTORING

Postdocs

Alec Dunton, Graduate student in Applied Mathematics, University of Colorado Boulder. Fast and scalable Gaussian process approximation in distributed memory. 2021-2023.

Students

Juliette Mukangango, PhD student in Statistics, Colorado School of Mines. Novel loss and objective functions for outlier robustness in training sparse MuyGPs models. Primary Advisor: Douglas Nychka. Summer 2023.

Akil Andrews, PhD student in Computer Science, University of New Mexico. Adaptive Bayesian optimization under changing data representations. Primary Advisor: Melanie E Moses. Summer 2023.

Ian McGovern, PhD student in Statistics, University of California, Los Angeles. Uncertainty analysis of hybrid deep neural network and Gaussian process predictions. Primary Advisor: Frederic Schoenberg. Summer 2023.

Abiodun Sumonu, PhD student in Mathematics, University of Alabama. Survey of Biclustering Algorithms. Summer 2023.

Keegan Kresge, Post Baccalaureate in Mathematics, DOD. K-Nearest Neighbors performance of exponentiated subspace embeddings on large graphs. Summer 2023.

Marina Dunn, Masters student in Data Science, University of California, Riverside. Visualizing sparse Gaussian process optimization. Summer 2022.

Killian Wood, PhD Student in Applied Mathematics, University of Colorado, Boulder. Multiscale Bayesian optimization of MuyGPs. Primary Advisor: Emiliano Dall'Anese. Summer 2022.

Michał Lisicki, PhD student in Computer Science, University of Guelph. Distributional reinforcement learning on gridworld environments. Primary Advisor: Graham Taylor. Summer 2022.

Sudharshan Srinivasan, PhD student in Computer Science, University of Oregon. Communication optimization for highly non-uniform distributed graph algorithms. Primary Advisor: Boyana Norris. Summer 2021.

Alec Dunton, PhD student in Applied Mathematics, University of Colorado, Boulder. Parameter sensitivity of stochastic block models under subspace embeddings. Primary Advi-

sor: Alireza Doostan. Summer 2020.

Teams and Challenges

Davy Walker, Ukamaka Nnyaba, and Hewan Shemtaga, PhD Students in Computer Science, Auburn University. Auburn Data Science Capstone Project. ECG Time Series Classification Using Computationally Efficient Gaussian Processes. Autumn 2023.

Rafael Bidese, Chinedu Eleh, and Yunli Zhang, PhD Students in Computer Science, Auburn University. Auburn Data Science Capstone Project. Stellar Blend Image Classification Using Computationally Efficient Gaussian Processes. Autumn 2022.

Jocelyn Ornelas, Alan Triano, Cristian Espinosa, Denylson Fuentes, and Rahul Ravi, A PhD Student (Jocelyn) and four undergraduates in Data Science programs, University of California, Merced. LLNL Data Science Challenge. Asteroid detection and orbit extraction from astronomy corpora. Spring 2021.

Professional Service

Conference Service

- Program Committee: SIAM International Conference on Data Mining (SDM24). Houston, Texas, USA. April 18-20, 2024.
- Program Committee: 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD23). Long Beach, California, USA. August 6-10, 2024.
- Program Committee: 28th International AAAI Florida Artificial Intelligence Research Symposium Conference, FLAIRS-28. Hollywood, Florida, USA. May 18-20, 2015.

Journal Service

- Reviewer: American Astronomical Society: The Astrophysical Journal. 2022.

Teaching Experience

Thayer School of Engineering at **Dartmouth College**, Hanover, VT, USA

Teaching Assistant

Instructor for ENGS/QBS 108: Applied Machine Learning **Autumn 2017**

- Collaborated with instructors to develop course curriculum aimed at graduate engineering and computer science students and taught $\sim 25\%$ of course lecture content.
- Led team of 4 teaching assistants
- Provided group and one-on-one assistance to students covering lecture topics
- Planned, wrote, and graded all student assignments

Instructor for ENGS 177: Decision Making Under Risk and Uncertainty **Winter 2017**

- Planned and taught a weekly recitation covering practical machine learning topics
- Provided ground and one-on-one assistance to students covering lecture topics
- Wrote student assignments with the assistance of the instructor and provided grading

The Ohio State University, Columbus, OH, USA

Teaching Assistant

Instructor for CSE 625: Automata and Formal Languages **Summer & Autumn 2010**

- Planned and taught a weekly recitation covering details and proofs of lecture topics
- Graded student assignments

Grader for CSE 560: System Software Design and Devlopment **Summer 2010**

- Graded student assignments and held office hours

Awards

LLNL Deputy Director Science & Technology Excellence in Publication Award

2023: Gaussian process classification for galaxy blend identification in LSST

2023: Light curve completion and forecasting using fast and scalable Gaussian processes (MuyGPs)

2021: Scaling Graph Clustering with Distributed Sketches

Advanced Maui Optical and Space Surveillance Technologies (AMOS) Conference

2022: Best Machine Learning Paper

High Performance Extreme Computing (HPEC) Graph Challenge
- 2020: Graph Challenge Champion
- 2019: Graph Challenge Champion

International Conference on Security and Cryptography (SECRYPT)
- 2018: Best Paper Award

MIT Lincoln Laboratory
- 2015: Lincoln Scholar Fellowship

The Ohio State University
- 2010: Phi Beta Kappa Inductee
- 2010: Bingham Award in Philosophy
- 2008: Kenneth Cummings Scholarship
- 2007-2011: Distinguished Merit Scholarship
- 2007-2011: Ohio Academic Scholarship

CITIZENSHIP USA