John T. Halloran

Contact Information	Da Vinci Building 1632 Da Vinci Ct. Davis, CA 95618	https://jthalloran.bitbucket.io

Research Interests

Machine learning algorithms for HPC architectures, computational biology, large-scale convex optimization, dynamic graphical models, approximate inference in graphical models, discriminative and unsupervised learning, automatic speech recognition

Education

University of Washington, Seattle
Ph.D. in Electrical Engineering, March 18, 2016
Advisors: Prof. Jeff Bilmes and Prof. William Noble
Committee: Prof. Jeff Bilmes, Prof. William Noble, Prof. Katrin Kirchhoff, and Prof. Ali Shojaie
Areas: Machine learning, proteomics, dynamic graphical models, computational biology, generative training, discriminative training

University of Hawaii at Manoa, Honolulu
M.S. in Electrical Engineering, August, 2010
Advisors: Prof. Yingbin Liang and Prof. Narayana Prasad Santhanam
Areas: Graphical models, convex optimization, wireless communications, lossy data compression

Seattle University, Seattle
B.S. in Electrical Engineering, June, 2008
B.S. in Mathematics, June, 2008

Work Experience

Postdoctoral Scholar, University of California, Davis (July 2016 - Present)
Working with Prof. David Rocke to develop machine learning methods for proteomics and genomics analysis.

Research Assistant, University of Washington (Dec. 2010 - March 2016)
Developed machine learning methods for protein analysis.

Research Assistant, University of Hawaii at Manoa (June 2009 - Aug. 2010)
Worked on machine learning methods for wireless communications and optimal information theoretic lossy compression.

Teaching Assistant, University of Hawaii at Manoa (Aug. 2008 - May 2009)
Provided lectures, administered labs, and graded assignments/exams for courses in the Electrical Engineering Department.

Undergraduate Research Assistant, Seattle University (June 2007 - Sept. 2007)
Investigated machine learning methods for image compression.

Software Skills

C/C++, Python, Matlab, Caffe, Keras, CUDA, OpenMP, Blas, Git, Subversion, Mercurial, Bash, Univa Grid Engine (UGE/SGE) cluster software, HTCondor cluster software, LaTeX, HTML, Microsoft Office, Linux, Windows, Cygwin

Grants

- Nvidia Hardware Grant, August 2016 (Nvidia Tesla K40 GPU). *Hardware awarded in perpetuity to explore the use of deep learning for computational biology problems.*

- Genome Training Grant, August 2011-August 2013 (tuition and stipend, two years). *Awarded to nine Ph.D. students doing genomics/proteomics-related research at UW each year.*

Conferences

[1] **JT Halloran** and DM Rocke. "Learning Concave Conditional Likelihood Models for Improved Analysis of Tandem Mass Spectra." *Advances in Neural Information Processing Systems (NIPS)*, 2018.

[2] **JT Halloran** and DM Rocke. "Gradients of Generative Models for Improved Discriminative Analysis of Tandem Mass Spectra." *Advances in Neural Information Processing Systems (NIPS)*, 2017. **Spotlight Presentation**, 3.5% Acceptance rate.

[3] S Wang, **JT Halloran**, JA Bilmes, and WS Noble. "Faster and more accurate graphical model identification of tandem mass spectra using trellises." *Bioinformatics (Proc. of the ISMB)*, 2016.

[4] **JT Halloran**, JA Bilmes, and WS Noble. "Learning Peptide-Spectrum Alignment Models for Tandem Mass Spectrometry." *Uncertainty in Artificial Intelligence (UAI)*, 2014.

[5] AP Singh, **J Halloran**, JA Bilmes, K Kirchhoff, and WS Noble. "Spectrum Identification using a Dynamic Bayesian Network Model of Tandem Mass Spectra." *Uncertainty in Artificial Intelligence (UAI)*, 2012.

[6] Y Liang, L Lai, and **J Halloran**. "Distributed algorithm for collaborative detection in cognitive radio networks." *The Allerton Conference on Communication, Control, and Computing*, 2009.

Journals

[7] **JT Halloran** and DM Rocke. "A Matter of Time: Faster Percolator Analysis via Efficient SVM Learning for Large-Scale Proteomics." *Journal of Proteome Research (JPR)*, 2018.

[8] J Liu, **J Halloran**, J Bilmes, R Daza, C Lee, E Mahen, D Prunkard, C Song, S Blau, M Dorschner, V Gadi, J Shendure, A Blau, and W Noble. "Comprehensive statistical inference of the clonal structure of cancer from multiple biopsies." *Scientific Reports*, 2017.

[9] **JT Halloran**, WS Noble, and JA Bilmes. "A dynamic Bayesian network for accurate detection of peptides from tandem mass spectra." *Journal of Proteome Research (JPR)*, 2016.

[10] Y Liang, L Lai, and **J Halloran**. "Distributed cognitive radio network management via algorithms in probabilistic graphical models." *IEEE JSAC, Special Issue on Advances in Cognitive Radio Networking and Communications*, 2011.

Book Chapters

[11] **JT Halloran**. "Analyzing Tandem Mass Spectra using the DRIP Toolkit: Training, Searching, and Post-Processing." In *Data Mining for Systems Biology: Methods and Protocols*, Ed. by H Mamitsuka. Humana Press, 2018.

Theses

[12] **JT Halloran**. "Graphical Models for Peptide Identification of Tandem Mass Spectra." Ph.D. Thesis, University of Washington, Department of Electrical Engineering, 2016.

[13] **J Halloran**. "Probabilistic Graphical Models and Random Graphs with Applications to Wireless Communications and Data Compression." M.Sc. Thesis, University of Hawaii at Manoa, Department of Electrical Engineering, 2010.

Contributed Abstracts and Workshops

[14] **JT Halloran**. "Invited Talk: Analyzing Tandem Mass Spectra: A Graphical Models Perspective." *Proceedings of Machine Learning Research vol 73* (2017): 6-6.

[15] **JT Halloran** and DM Rocke. "Fisher Kernels for Improved Analysis of Tandem Mass Spectra." *NIPS Workshop on Machine Learning in Computational Biology (MLCB)*, 2016.

[16] **JT Halloran**, AP Singh, JA Bilmes, and WS Noble. "Peptide Identification of Tandem Mass Spectra via Spectrum Alignment using a Dynamic Bayesian Network." *NIPS Workshop on Machine Learning in Computational Biology (MLCB)*, 2012.

[17] AP Singh, **J Halloran**, JA Bilmes, K Kirchhoff, and WS Noble. "Spectrum Identification with a Dynamic Bayesian Network model of Tandem Mass Spectra." *RECOMB Satellite Conference on Computational Proteomics*, 2012.

Preprints

[18] RK Iyer, **JT Halloran**, and K Wei. "Jensen: An Easily-Extensible C++ Toolkit for Production-Level Machine Learning and Convex Optimization." arXiv:1807.06574, 2018.

Other Works

[19] R Rogers, JA Bilmes, and **JT Halloran**. "GMTK Tutorial on Dynamic Graphical Model Training with Gaussian Mixture Unaries, using TIMIT."

[20] **J Halloran**. "Integer Solutions to the Equation $a^n + b^n = c^n$, for $n > 2$."

SOFTWARE

All of the software listed below are available with source code at the URLs specified.

[1] RK Iyer, **JT Halloran**, and K Wei. JENSEN is an easily-customizable C++ toolkit for massive-scale machine learning (ML) and convex optimization. Jensen natively supports a large number of popular ML loss functions, state-of-the-art optimization algorithms, and ML applications. https://github.com/rishabhk108/jensen. 2018.

[2] **JT Halloran** and DM Rocke. PERCOLATOR is a popular post-processing for enriching the results of a shotgun proteomics database search. The following repository contains SVM solvers highly optimized for large-scale Percolator analysis. https://bitbucket.org/jthalloran/percolator_upgrade. 2018.

[3] **JT Halloran**, JA Bilmes, and WS Noble. The DRIP TOOLKIT (DTK) is a tandem mass spectrometry search engine that uses a dynamic Bayesian network. DTK supports parameter estimation, multithreading, utilities for cluster use, instantiating/decoding/plotting DRIP PSMs in the python shell, and in-browser analysis of identified spectra. https://jthalloran.bitbucket.io/dripToolkit. 2016-2018.

[4] J Liu, **JT Halloran**, JA Bilmes, and WS Noble. THEMIS infers the genotypes and relative frequencies of cancer clones from deep DNA sequence of multiple biopsies from a single cancer using a dynamic Bayesian network. https://github.com/jieliu6/THEMIS. 2017.

[5] R Rogers, JA Bilmes, S Wang, and **JT Halloran**. The GRAPHICAL MODELS TOOLKIT (GMTK) is an open source, publicly available toolkit for rapidly prototyping statistical models using dynamic graphical models (DGMs) and dynamic Bayesian networks (DBNs). The discriminative training engine for general DBNs, based on the Maximum Mutual Information (MMI) criterion, was written by Shengjie Wang and I. https://melodi.ee.washington.edu/gmtk. 2016.

PROFESSIONAL ACTIVITIES

Reviewer for: BMC Genomics, IEEE Journal of Biomedical and Health Informatics, the Workshop on Algorithms in Bioinformatics (WABI), and the International Conference on Research in Computational Molecular Biology (RECOMB)

RELEVANT COURSES	Machine Learning, Probabilistic Graphical Models, Advanced Probabilistic Graphical Models, Dynamic Probabilistic Graphical Models, Supervised Learning and Optimization, Convex Optimization, Information Theory (I and II), Speech Recognition, Wavelets, Probability and Random Processes, Stochastic Processes, Applied Random Processes, Detection and Estimation Theory, Advanced Network Algorithms, Digital Signal Processing, Digital Image Processing, Numerical Methods, Linear Systems Theory, Wireless Communications, Nonlinear Dynamical Systems and Chaos, Abstract Algebra, Partial Differential Equations

TALKS

[1] "Gradients of Generative Models for Improved Discriminative Analysis of Tandem Mass Spectra." NIPS, Long Beach, CA. December, 2017.

[2] "Analyzing Tandem Mass Spectra: A Graphical Models Perspective." Keynote speaker, Advanced Methodologies for Bayesian Networks (AMBN), Kyoto, Japan. September, 2017.

[3] "Faster and More Accurate Graphical Models for Peptide Identification of Tandem Mass Spectra." Department of Biostatistics, UC Davis. May, 2016.

[4] "Graphical Models for Peptide Identification of Tandem Mass Spectra." MODE Laboratory, University of Washington. May, 2015.

[5] "Peptide Identification of Tandem Mass Spectra via Spectrum Alignment using a Dynamic Bayesian Network." NIPS Workshop on Machine Learning in Computational Biology (MLCB), Lake Tahoe, CA. December, 2012.

[6] "Peptide Identification in Tandem Mass Spectrometry using Dynamic Bayesian Networks." Department of Electrical and Computer Engineering, Seattle University. April, 2012.

[7] "Integer Solutions to the Equation $a^n + b^n = c^n$, for $n > 2$." Department of Mathematics, Seattle University. June, 2007.

MENTORING

- **Alex Ke**. Winter 2012 - Spring 2012. Ph.D. student, Electrical Engineering, University of Washington. Project: "Designing dynamic graphical models for vision-based activity recognition."

- **David Chen**. Winter 2012 - Spring 2012. Undergraduate student, Electrical Engineering, University of Washington. Project: "Designing dynamic graphical models for vision-based activity recognition."

- **Karthik Mohan**. Winter 2013 - Spring 2013. Ph.D. student, Electrical Engineering, University of Washington. Project: "Bipartite matchings of tandem mass spectra using submodular functions."

- **Wenruo Bai**. Fall 2013 - Winter 2014. Ph.D. student, Electrical Engineering, University of Washington. Project: "Bipartite matchings of tandem mass spectra using submodular functions."

- **Shengjie Wang**. Fall 2013 - Winter 2016. Ph.D. student, Computer Science and Engineering, University of Washington. Project: "Discriminative training graphical models for analysis of tandem mass spectra."

TEACHING EXPERIENCE

- Teaching Assistant for EE160: Programming for Engineers (in C), 2 sections, Fall 2008 at University of Hawaii at Manoa, Honolulu

- Teaching Assistant for EE213: Circuit Analysis II, Spring 2009 at University of Hawaii at Manoa, Honolulu

- Teaching Assistant for EE367: Computer Data Structures and Algorithms, Spring 2009 at University of Hawaii at Manoa, Honolulu

HOBBIES   Cooking noodle dishes, reading paperbacks, road trips, basketball, racquetball, and hiking

REFERENCES   Available upon request.