

BRUNO RIBEIRO

October 14, 2021

RESEARCH INTERESTS

My research focuses on endowing *machine learning* algorithms with the ability to learn invariant representations for relational and temporal data for both correlational and counterfactual tasks. Invariances are one of the most consequential forms of prior knowledge in mathematics, physics, statistics, and also machine learning. An object is invariant if it is the same under some pre-specified set of transformations. For instance, graphs and tensors often are the same under index permutations (isomorphism), while temporal and spatial invariances often imply stationarity, and robustness to adversarial attacks is an invariance to perturbations. Whether designing recommender systems for social networks, building robots that can reason about the relationships between objects in their surrounding environment, learning to generate new drugs, learning causal relationships, or extracting logical rules from data examples, invariances are essential for robust automated learning. The outcome of my research has had far-reaching implications, from a principled framework to perform counterfactual tasks (predicting what-if scenarios), to endowing machine learning models with world-knowledge.

ACADEMIC EXPERIENCE

- 2015 – present Assistant Professor of Computer Science
Purdue University
- 2013 – 2015 Postdoctoral Fellow, Carnegie Mellon University
Mentor: Christos Faloutsos
- 2012 – 2013 Visiting Researcher, Laboratory for the Modeling of Biological and Socio-technical
Systems, Department of Physics, Northeastern University
- 2010 – 2013 Postdoctoral Researcher, University of Massachusetts Amherst
Mentor: Don Towsley

EDUCATION

- 2010 Ph.D. Computer Science, University of Massachusetts Amherst. Adviser: Don Towsley
- 2003 M.S. Computer Engineering, Federal University of Rio de Janeiro
- 2001 B.S. Computer Science, Federal University of Rio de Janeiro

AWARDS AND HONORS

| | |
|-----------|---|
| 2021 | Best Paper Award ACM CODASPY 2021 |
| 2020 | NSF CAREER Award |
| 2020 | Purdue Seed for Success Award |
| 2020 | Purdue College of Science Team Award |
| 2016 | Best Paper Award ACM SIGMETRICS'16 |
| 2014 | Best Paper Award IEEE NetSciCom'14 |
| 2004–2009 | Brazilian National Scholarship Board (CAPES) Full Fellowship [Ph.D.] |
| 2002–2003 | Brazilian National Science Foundation (CNPq) M.S. Scholarship [M.Sc.] |
| 2000–2001 | Brazilian National Science Foundation (CNPq) Undergraduate Scholarship. |

PUBLICATIONS

Journal Papers

- Pintu Adhikari^G, Kevin Xia^U, Garrett Shaffer, **Bruno Ribeiro**, Dimitrios Peroulis, “Automated Microwave Filter Tuning Using Machine Learning“ (*under review*)
- S Chandra Mouli^G, Leonardo Teixeira^G, Jennifer Neville, **Bruno Ribeiro**, “Deep Lifetime Clustering”, (in preparation)
- 2021 J15. Carlos H. C. Teixeira^G, Mayank Kakodkar^G, Vinícius Dias^G, Wagner Meira Jr., **Bruno Ribeiro**, Sequential Stratified Regeneration: MCMC for Large State Spaces with an Application to Subgraph Counting Estimation, Data Mining and Knowledge Discovery, 2021
- J14. Lisette Espín-Noboa^G, Fariba Karimi, **Bruno Ribeiro**, Kristina Lerman, Claudia Wagner, Quantifying Biases in Attribute Inference on Social Networks, *Applied Network Science*, Springer, 2021
- 2020 J13. Mandana Saebi, Jian Xu, Lance M Kaplan, **Bruno Ribeiro**, Nitesh V Chawla, “Efficient modeling of higher-order dependencies in networks: From algorithm to application for anomaly detection”, *EPJ Data Science*, 2020
- J12. Wo Jae Lee^G, Kevin Xia^U, Nancy Denton, **Bruno Ribeiro**, John W. Sutherland, “A Development of Speed Invariant Deep Learning Model with Application to Condition Monitoring of Rotating Machinery,“ *Journal of Intelligent Manufacturing*, 2020
- 2019 J11. Guangyu Li, Yong Liu, **Bruno Ribeiro**, Hao Ding, On group popularity prediction in event-based social networks, *IEEE Transactions on Network Science and Engineering*, 2019
- J10. Vladyslav Fedchenko, Giovanni Neglia, Bruno Ribeiro, Feedforward Neural Networks for Caching: Enough or Too Much?, *ACM SIGMETRICS Performance Evaluation Review*, 2019.
- 2018 J9. Fabricio Murai, **Bruno Ribeiro**, Don Towsley, Pinghui Wang, Characterizing Directed and Undirected Networks via Multidimensional Walks with Jumps, *Transactions on Knowledge Discovery from Data (TKDD)*, 2018.
- J8. Pinghui Wang, Junzhou Zhao, **Bruno Ribeiro**, John C.S. Lui, Don Towsley, and Xiaohong Guan, Practical Characterization of Large Networks Using Neighborhood Information, *Knowledge and Information Systems*, 2018.

- 2017 J7. Fabricio Murai^G, Diogo Renno^G, **Bruno Ribeiro**, Gisele L. Pappa, Don Towsley, Krista Gile, “Selective Harvesting over Networks,” *Data Mining and Knowledge Discovery (Journal)*, 2017.
- 2015 J6. Bo Jiang^G, Liyuan Sun^G, Daniel R. Figueiredo, **Bruno Ribeiro**, Don Towsley, “On the Duration and Intensity of Cumulative Advantage Competitions”, *JSTAT*, 2015.
- 2014 J5. P. Wang^G, J.C.S. Lui, **B. Ribeiro**, D. Towsley, J. Zhao, X. Guan, “Efficiently Estimating Motif Statistics of Large Networks,” *TKDD (Journal)*, 2014.
- J4. Peng Xia^G, Kun Tu^G, **Bruno Ribeiro**, Hua Jiang, Xiaodong Wang, Cindy Chen, Benyuan Liu, Don Towsley, “Who is Dating Whom: Characterizing User Behaviors of a Large Online Dating Site,” *Lecture Notes on Social Networks*, 2014.
- 2013 and earlier J3. F. Murai^G, **B. Ribeiro**, D. Towsley, P. Wang, “On Set Size Distribution Estimation and the Characterization of Large Networks via Sampling,” *IEEE JSAC Special Issue on Network Science*, 2013.
- J2. **Bruno Ribeiro**, Nicola Perra, and Andrea Baronchelli, “Quantifying the Effect of Temporal Resolution on Time-varying Networks,” *Scientific Reports*, 2013
- J1. S. C. Coutinho and Bruno Ribeiro. “On Holomorphic Foliations without Algebraic Solutions,” *Experimental Mathematics*, v.10, n.4, p.529 - 536, 2001.

Conference Papers

- Leonardo Teixeira^G, Brian Jalaian, **Bruno Ribeiro**, Reducing Overconfidence Against Adversaries through Graph Algorithms, (under review)
- 2022 C41. Yun Seong Nam^G, Jianfei Gao^G, Sanjay Rao, **Bruno Ribeiro**, “Xatu: Exploiting a richer throughput model for video streaming through neural networks”, *SIGMETRICS 2022 (accepted, acceptance rate 23%)*
- 2021 C40. Leonardo Cotta^G, Chris Morris, **Bruno Ribeiro**, Reconstruction for Powerful Graph Representations, *Neural Information Processing Systems (NeurIPS), 2021* (8 pages, acceptance rate 26%)
- C39. Beatrice Bevilacqua^G, Yangze Zhou^G, **Bruno Ribeiro**, Graph Representations for Counterfactual Size-Invariant Graph Classification, International Conference on Machine Learning (ICML) 2021 (long talk, acceptance rate 3%)
- C38. Mengyue Hang^G, Jennifer Neville, **Bruno Ribeiro**, A Collective Learning Framework to Boost GNN Expressiveness, International Conference on Machine Learning (ICML) 2021 (short talk, acceptance rate 21%)
- C37. S Chandra Mouli^G, **Bruno Ribeiro**, Neural Network Extrapolations with G-invariances from a Single Environment, *International Conference on Learning Representations (ICLR), 2020* (9 pages, acceptance rate 28%)

- C36. Mohsen Minaei^G, S Chandra Mouli^G, Mainack Mondal, **Bruno Ribeiro**, Aniket Kate, “Deceptive Deletions for Protection of Withdrawn Posts on Social Media”, In *Network and Distributed System Security Symposium (NDSS)*, 2021 (18 pages, acceptance rate (2020) 17.4%)
- C35. Jiacheng Li^G, Ninghui Li, **Bruno Ribeiro**, “Membership Inference Attacks and Defenses in Classification Models”, *CODASPY 2021*
- C34. Pintu Adhikari^G, Kevin Xia^U, Garrett Shaffer, Bruno Ribeiro, Dimitrios Peroulis, An S-band Automatically Tunable Bandpass Filter Based on a Machine Learning Approach, IEEE 21st Annual Wireless and Microwave Technology Conference, 2021
- 2020 C33. Leonardo Cotta^G, Carlos H. C. Teixeira, Ananthram Swami, **Bruno Ribeiro**, “Un-supervised Joint k-node Graph Representations with Compositional Energy-Based Models”, *Neural Information Processing Systems (NeurIPS)*, 2020 (8 pages, acceptance rate 21%)
- C32. Amit Sheoran^G, Sonia Fahmy, Matthew Osinski, Chunyi Peng, **Bruno Ribeiro**, Jia Wang, “Experience: Towards Automated Customer Issue Resolution in Cellular Networks”, *MOBICOM 2020* (13 pages, acceptance rate 16%)
- C31. Balasubramaniam Srinivasan^G, **Bruno Ribeiro**, “On the Equivalence between Node Embeddings and Structural Graph Representations”, *International Conference on Learning Representations (ICLR)*, 2020 (9 pages, acceptance rate 26%)
- C30. Jianfei Gao^G, Mohamed A. Zahran^G, Amit Sheoran^G, Sonia Fahmy, **Bruno Ribeiro**, *Infinity Learning: Learning Markov Chains from Aggregate Steady-State Observations*, AAI 2020 (7 pages, acceptance rate 20.6%).
- 2019 C29. Ryan L. Murphy^G, Balasubramaniam Srinivasan^G, Vinayak Rao, **Bruno Ribeiro**, Relational Pooling for Graph Representations, International Conference on Machine Learning (ICML), 2019 (8 pages, acceptance rate 22.6%)
- C28. Changping Meng^G, Jiasen Yang^G, **Bruno Ribeiro**, Jennifer Neville, HATS: A Hierarchical Sequence-Attention Framework for Inductive Set-of-Sets Embeddings, KDD (Oral), 2019 (9 pages, acceptance rate (oral) 9%)
- C27. Ryan L. Murphy^G, Balasubramaniam Srinivasan^G, Vinayak Rao, **Bruno Ribeiro**, “Janossy Pooling: Learning Deep Permutation-Invariant Functions for Variable-Size Inputs”, *International Conference on Learning Representations (ICLR)*, 2019 (10 pages, acceptance rate 31%)
- C26. Huangyi Ge^G, Sze Yiu Chau^G, Bruno Ribeiro, Ninghui Li, “Random Spiking and Systematic Evaluation of Defenses Against Adversarial Examples”, *CODASPY 2019*
- 2018 C25. Jason Meng^G, Chandra S. Mouli^G, **Bruno Ribeiro**, Jennifer Neville, “Subgraph Pattern Neural Networks for High-Order Graph Evolution Prediction”, *AAAI*, 2018 (8 pages, acceptance rate 25%)
- C24. Pedro Savarese^G, Mayank Kakodar^G, **Bruno Ribeiro**, “From Monte Carlo to Las Vegas: Improving Restricted Boltzmann Machine Training through Stopping Sets”, *AAAI*, 2018 (8 pages, acceptance rate 25%)

- C23. Carlos Teixeira^G, Leonardo Cotta^G, Bruno Ribeiro, and Wagner Meira Jr., Graph Pattern Mining and Learning through User-defined Relations, *ICDM 2018*. (6 pages, acceptance rate 19%)
- C22. Yun Seong Nam^G, Zahaib Akhtar^G, Ramesh Govindan, Sanjay Rao, Jessica Chen, Ethan Katz-Bassett, **Bruno Ribeiro**, Jibin Zhan, Hui Zhang, “Oboe: Auto-tuning Video ABR Algorithms to Network Conditions”, *SIGCOMM 2018*. (12 pages, acceptance rate 18%)
- C21. Guangyu Li^G, Yong Liu, **Bruno Ribeiro** and Hao Ding, On Group Popularity Prediction in Event-Based Social Networks, *ICWSM 2018*. (4 pages, acceptance rate ~20%)
- C20. Mohamed S Hassan^G, Bruno Ribeiro, Walid G Aref, SBG-sketch: a self-balanced sketch for labeled-graph stream summarization, *ICSSDM 2018*. (12 pages)
- 2017 C19. Jiasen Yang^G, **Bruno Ribeiro**, Jennifer Neville, “Should We Be Confident in Peer Effects Estimated From Partial Crawls of Social Networks?,” *AAAI ICWSM 2017* (4 pages, acceptance rate 20%).
- 2016 C18. Konstantin Avrachenkov, **Bruno Ribeiro**, Jithin Sreedharan^G, “Inference of Online Social Network Statistics via Lightweight Random Walk Crawls,” *SIGMETRICS’16* (12 pages, alphabetical order, acceptance rate 13.5%)
- C17. Bo Jiang^G, Daniel Figueiredo, **Bruno Ribeiro**, Don Towsley, “On the Duration and Intensity of Competitions in Nonlinear Pólya Urn Processes with Fitness,” *SIGMETRICS’16* (12 pages, acceptance rate 13.5%) (**Best Paper Award**)
- C16. Flavio Figueiredo^G, **Bruno Ribeiro**, Jussara Almeida, Christos Faloutsos, “TribeFlow: Mining & Predicting User Trajectories,” *WWW’16* (12 pages, acceptance rate 17%)
- 2015 C15. **Bruno Ribeiro**, Minh Hoang^G, Ambuj Singh, “Beyond Models: Forecasting Complex Network Processes Directly from Data,” *WWW’15* (12 pages, acceptance rate 16%)
- C14. **Bruno Ribeiro**, Christos Faloutsos, “Modeling Website Popularity Competition in the Attention-Activity Marketplace”, *WSDM’15* (12 pages, acceptance rate 16.4%)
- 2014 C13. **Bruno Ribeiro**, “Modeling and Predicting the Growth and Death of Membership-based Websites,” *WWW’14* (12 pages, acceptance rate 16%)
- 2013 & Earlier work C12. Peng Xia^G, **Bruno Ribeiro**, Cindy Chen, Benyuan Liu, Don Towsley, “A Study of User Behavior on an Online Dating Site,” *ASONAM’13*.
- C11. Ting-Kai Huang^G, Md Sazzadur Rahman, Harsha V. Madhyastha, Michalis Faloutsos, and **Bruno Ribeiro**, “An Analysis of Socware Cascades in Online Social Networks,” *WWW’13*.
- C10. Daniel Figueiredo, Philippe Nain, **Bruno Ribeiro***, Edmundo de Souza e Silva, and Don Towsley, “Characterizing Continuous Time Random Walks on Time-varying Graphs,” *SIGMETRICS’12*. *Corresponding author.
- C9. **Bruno Ribeiro**, Don Towsley, “On the Estimation Accuracy of Degree Distributions from Graph Sampling,” *IEEE Conference on Decision and Control, 2012*.

- C8. Bruno Ribeiro, Pinghui Wang, Fabricio Murai, and Don Towsley, Sampling Directed Graphs with Random Walks, *IEEE INFOCOM 2012*.
- C7. **Bruno Ribeiro**, Daniel Figueiredo, Edmundo de Souza e Silva, and Don Towsley, Characterizing Dynamic Graphs with Continuous-time Random Walks (short paper), *SIGMETRICS 2011*.
- C6. **Bruno Ribeiro** and Don Towsley, “Estimating and Sampling Graphs with Multidimensional Random Walks,” *IMC 2010*.
- C5. William Gauvin, Bruno Ribeiro, Ben Liu, Don Towsley, and Jei Wang, “Measurement and Gender-Specific Analysis of User Publishing Characteristics on MySpace,” *IEEE Network Magazine (Special Edition on Social Networking)*, 2010.
- C4. Bruno Ribeiro, Tao Ye, and Don Towsley, “A Resource-minimalist Flow Size Histogram Estimator,” *IMC 2008*.
- C3. Bruno Ribeiro, Weifeng Chen, Gerome Miklau, and Don Towsley, “Analyzing Privacy in Enterprise Packet Trace Anonymization,” *NDSS 2008*.
- C2. Bruno Ribeiro, Don Towsley, Tao Ye, and Jean Bolot, “Fisher Information of Sampled Packets: an Application to Flow Size Estimation,” *IMC 2006*.
- C1. Weifeng Chen, Yong Huang, Bruno Ribeiro, Kyoungwon Suh, Honggang Zhang, Edmundo de Souza e Silva, Jim Kurose, and Don Towsley, “Exploiting the IPID field to infer network path and end-system characteristics.” *Passive and Active Measurement Workshop (PAM) 2005*.

Refereed Workshop Papers

- 2020 W14. PC Sruthi, Sanjay Rao, **Bruno Ribeiro**, “Pitfalls of data-driven networking: A case study of latent causal confounders in video streaming”, *ACM SIGCOMM Proceedings of the Workshop on Network Meets AI & ML*, 2020.
- 2019 W13. Leonardo Teixeira^G, Brian Jalaian, Bruno Ribeiro, Are Graph Neural Networks Miscalibrated?, Learning and Reasoning with Graph-Structured Representations Workshop @ ICML (Spotlight), 2019
- 2017 W12. Jiasen Yang^G, **Bruno Ribeiro**, Jennifer Neville, “Stochastic Gradient Descent for Relational Logistic Regression via Partial Network Crawls,” *UAI Seventh International Workshop on Statistical Relational AI*, 2017
- W11. Kun Tu^G, Bruno Ribeiro, Ananthram Swami, Don Towsley, “Temporal Clustering in Dynamic Networks with Tensor Decomposition,” *NIPS Time Series Workshop*, 2017.
- 2016 W10. Flavio Figueiredo^G, **Bruno Ribeiro**, Christos Faloutsos, Nazareno Andrade, Jussara M. Almeida, “Mining Online Music Listening Trajectories,” *ISMIR: 17th International Symposium/Conference on Music Information Retrieval*, 2016.

- 2014
- W9. Kun Tu^G, **Bruno Ribeiro**, H. Jiang, X. Wang, David Jensen, Benyuan Liu, Don Towsley, “Online Dating Recommendations: Matching Markets and Learning Preferences,” *5th International Workshop on Social Recommender Systems (SRS’14 @ WWW’14)*
- W8. Yeon-sup Lim^G, **Bruno Ribeiro**, Don Towsley, “Classifying Latent Infection States in Complex Networks,” *SIMPLEX’14*
- W7. K. Avrachenkov, P. Basu, G. Neglia, **B. Ribeiro***, and D. Towsley, “Pay Few, Influence Most: Online Myopic Network Covering,” *IEEE INFOCOM NetSciCom’14 (Best Paper Award)*. **Corresponding author.*
- W6. James Atwood^G, **Bruno Ribeiro**, Don Towsley, “Efficient Network Generation Under General Preferential Attachment,” *SIMPLEX’14*
- 2013 & Earlier work
- W5. Fabricio Murai^G, **Bruno Ribeiro**, Don Towsley, and Krista Gile, “Characterizing Branching Processes from Sampled Data,” *SIMPLEX’13*.
- W4. **Bruno Ribeiro**, Prithwish Basu, and Don Towsley, Multiple Random Walks to Uncover Short Paths in Power Law Networks, *IEEE INFOCOM NetSciCom 2012*.
- W3. Y. Lim, D. S. Menasché, **B. Ribeiro**, D. Towsley, P. Basu Online estimating the k central nodes of a network. *IEEE Network Science Workshop (NSW) 2011*.
- W2. Konstantin Avrachenkov, **Bruno Ribeiro**, and Don Towsley, Improving Random Walk Estimation Accuracy with Uniform Restarts, *7th Workshop on Algorithms and Models for the Web Graph (WAW 2010)*, 2010.
- W1. Bruno Ribeiro, William Gauvin, Benyuan Liu, and Don Towsley, “On MySpace Account Spans and Double Pareto-Like Distribution of Friends,” *IEEE INFOCOM NetSciCom 2010*.

STUDENTS

GRADUATED PhD STUDENTS

Changping (Jason) Meng [**C25,C28**] (Co-advisor Neville, Ph.D. Student, Purdue, Fall 2016 – May 2020). Working at Google. Thesis topic: *Predicting High-Order Associations and Labels in Massive Heterogeneous Graphs*.

Ryan Murphy [**C27,C29**] (Co-advisor Rao, Ph.D. Student, Purdue, Fall 2014 – May 2021). Thesis topic: *A Novel Framework for Invariant Neural Networks Applied to Graph and Set Data*.

CURRENT PhD STUDENTS

Summary: 4 RAs, 3 TAs (CS), 1 Fellowship

Leonardo V. Teixeira [**W13**] (Ph.D. Student, Purdue, Spring 2016 –). Supported as **33% RA, 67% TA**. Research topic: *Uncertainty Quantification in Adversarial Machine Learning*

Chandra Mouli [**C25,C36,C37**] (Ph.D. Student, Purdue, Fall 2016 –). Supported as **100% RA**. Research topic: *Physics-reinforced Deep Learning & Counterfactual Learning*.

Mayank Kakodkar [C24,J15] (Ph.D. Student, Purdue, Fall 2016 –). Supported as **90% RA, 10% TA**. Research topic: *Developing a scalable Las Vegas alternative to Markov Chain Monte Carlo methods*.

Jianfei Gao [C30] (Ph.D. Student, Purdue, Fall 2017 –). Supported as **100% RA**. Research topic: *Teaching Deep Neural Network Invariances in Sequences*.

Leonardo Cotta [C23,C33,C40] (Ph.D. Student, Purdue, Fall 2017 –). Supported as **20% Fellowship, 60% RA, 20% TA**. Research topic: *High-order Network Models, Moving Beyond Pairwise Link Prediction with Energy-based Models*.

Balasubramaniam Srinivasan [C27,C29,C31] (Ph.D. Student, Purdue, Fall 2018 –). Supported as **100% RA**. Research topic: *Novel Deep Neural Network Architectures for Relational Data*.

Yangze Zhou [C39] (Co-adviser Vinayak Rao, Ph.D. Student in Statistics, Purdue, Fall 2019 –). Supported as **100% RA by Ribeiro**. Research topic: *Counterfactual Graph Invariances in Machine Learning*.

Beatrice Bevilacqua [C39] [**Ph.D. Student, Fall 2020 –**] Recruited from **GoBoiler** program, Fellowship-supported. Research Topic: *Size-invariant Representations in Graph Classification Tasks*.

UNDERGRADUATE RESEARCH ASSISTANTS:

Ishaan Saxena. Purdue CS Undergraduate, January 2020 – March 2020 (COVID-19 impacted). Undergraduate research (with weekly meetings) Understanding Positional Graph Embeddings.

Jack Good (**Now a PhD student at CMU**). Purdue CS Undergraduate, August 2017 – December 2017. Undergraduate research (with weekly meetings) Understanding the Energy Landscape Curvature of Neural Networks.

Kevin Xia [J12,C34] (**Now a PhD student at Columbia**). Purdue CS Undergraduate, January 2018 – December 2019. Undergraduate and MSc research (with weekly meetings) Adding Mechanistic Models Priors to Neural Networks

VISITING STUDENTS:

Beatrice Bevilacqua [C39] [**Starting PhD at Purdue, Fall 2020**] (Visiting MSc student from Spapienza University in Rome, part of **GoBoiler** program), Jul-Sept, 2019

Janaina Gomide (Visiting PhD student from Federal University of Rio de Janeiro, adviser Daniel Figueiredo), March – October, 2017 (**Ph.D. granted in 2019**)

Carlos Teixeira [J15,C33,C23] (Visiting PhD student from Federal University of Minas Gerais, adviser Wagner Meira), July – September, 2016 (**Ph.D. will be defended Fall 2020**).

Jian Xu [J13] (Visiting PhD student from Notre Dame, adviser Nitesh Chawla), May – June, 2016

Bhavana Jain (Foreign Undergraduate Student, IIT Madras, Summer 2017). Supported as part of Purdue's PURE program. Research topic: *Understanding the Energy Landscape of Restricted Boltzmann Machines*.

Abhishek Naik (Foreign Undergraduate Student, IIT Madras, Summer 2016). Supported as part of Purdue's PURE program. Research topic: *Predicting User Lifetime on Social Networks*.

TEACHING EXPERIENCE

- Fall 2021 CS37300 –Undergraduate *Data Mining and Machine Learning* (**210 students**)
Purdue University & Computer Science .
- Spring 2021 CS69000-DPL –Graduate level *Deep Learning* (**33 students**)
Purdue University & Computer Science (**developed as a permanent course**).
- Fall 2020 CS49000-LDA – Undergraduate *Large Scale Data Analytics* (**52 students**)
Purdue University & Computer Science (**developed as a permanent course**).
- Spring 2020 CS69000-DPL –Graduate level *Deep Learning* (**36 students [full]**)
Purdue University & Computer Science (**developed as a permanent course**).
This is an in-depth course on representation learning covering the mathematical foundations and practical hands-on tools, specifically: (a) Statistical tasks and how they relate to representation learning through simple measure theory and causality concepts; (b) Stochastic optimization; (c) Invariant representations of exchangeable random variables, set representations, graph representations (e.g., Graph Neural Networks), and node embeddings (e.g. word2vec, Glove); (d) Backpropagation and backpropagation-through-time; (e) Feedforward and convolutional networks; (f) Adversarial machine learning; (g) Noise outsourcing and variational auto-encoders; (h) Multi-task learning, transfer learning, and meta learning (via invariant theory); (i) Extrapolations and causal mechanisms; (j) Markov chains and sequence learning (including recurrent networks and transformers); (k) Rejection sampling and generative adversarial networks; (l) Evaluating the performance of neural networks, as well as formulating and testing hypotheses; (m) Understanding how algorithmic elements interact to impact performance.
- Fall 2019 CS49000-LSDA (**Developed for the Data Science Major**)
Undergraduate level *Large Scale Data Analytics* (**15 students**)
Purdue University & Computer Science (**developed as a permanent course**).
This course provides an introduction to the field of large scale data science. The course focuses on using computational methods and statistical techniques to analyze massive amounts of data and to extract knowledge. It provides an overview of the foundational computational and statistical tools for storing, manipulating, processing, and learning statistical models on big datasets. The course provides students hands-on experience with tools and methods.
- Fall 2018 CS69000-DPL – Graduate level *Deep Learning* (**27 students**)
Purdue University & Computer Science (**developed as a permanent course**).
- Spring 2018 CS57300 – Graduate level *Data Mining* (**71 students**)
Purdue University & Computer Science.
- Fall 2017 CS37300 – Undergraduate level *Data Mining & Machine Learning* (**87 students**)
Purdue University & Computer Science.

- Fall 2016 CS57300 – Graduate level “Data Mining” (**50 students**)
Purdue University & Computer Science.
- Spring 2016 CS57300 – Graduate level *Data Mining* (**84 students**)
Purdue University & Computer Science.
- Fall 2015 Graduate level *Topics In Data Mining* (10 students)
(CS69000-DM1) Purdue University & Computer Science.

SELECTED INVITED TALKS

- 2022 (**Invited Speaker**) Dagstuhl Seminar: Theory and Practice of Graph Embeddings (was November 2021 but now is March 2022)
(Keynote) AAAI Workshop
- 2021 (**Invited Speaker**) Unearthing the Relationship Between Graph Neural Networks and Matrix Factorization, Twitter ML Group, September 2nd.
(Invited Speaker) Unearthing the Relationship Between Graph Neural Networks and Matrix Factorization, MIT Graph Exploitation Symposium, May 17-18th.
- 2020 (**Invited Speaker / Virtual**) Unearthing the Relationship Between Graph Neural Networks and Matrix Factorization, Argonne National Laboratory, October 14th.
(Invited Lecture / Canceled COVID-19) Machine Learning in Networks (20 hours) (co-teaching with Jure Leskovec, Stanford), *First Summer Institute in Network Science*, Boston. <https://www.sinsa2020.org>. Other instructors include: Albert-László Barabási, Alex Vespignani, Noshir Contractor, Yamir Moreno, Tiago Peixoto, Santo Fortunato, Trey Ideker, Dima Krioukov, Olaf Sporns, and Jure Leskovec. June 29–July 1st, 2020
(Invited Speaker) Unearthing the Relationship Between Graph Neural Networks and Matrix Factorization, Caltech, Information Science and Technology Lunch Bunch Seminar Series, March 10th.
(Invited Speaker) Unearthing the Relationship Between Graph Neural Networks and Matrix Factorization, University of South California, Center for Systems and Control Seminar Series, March 9th.
(Invited Speaker) Unearthing the Relationship Between Graph Isomorphism, Graph Neural Networks, and Matrix Factorization, SIAM Workshop on Combinatorial Scientific Computing (CSC20), Feb 11th.
- 2019 (**Keynote**) Unearthing the Relationship Between Graph Neural Networks and Matrix Factorization. Winedale Workshop (UT Austin, Rice, Texas A& M), 2019. Past keynotes: Antonio Ortega (2019), Animashree Anandkumar (2016), John Langford (2015), Robert Kleinberg & Tim Roughgarden (2014), Piotr Indyk & Sanjeev Arora (2013), Stephen Wright & Alex Smola (2012), Ramesh Johari & Asu Ozdaglar (2011)

- (**Invited Speaker**) Unearthing the Relationship Between Node Embeddings and Graph Representations, Rice University, October 2019.
- (**Invited Speaker**) Unearthing the Relationship Between Graph Neural Networks and Matrix Factorization, Northeastern University, September, 2019.
- (**Invited Speaker**) New Directions in Representation Learning of Graphs, Great Lakes Workshop, September, 2019.
- (**Invited Speaker**) Unearthing the Relationship Between Graph Neural Networks and Matrix Factorization, Microsoft, September 2019.
- (**Keynote**) Towards a Deeper Understanding of Graphs, *GrAPL/GABB Workshop @ IPDPS Conference*, 2019.
Past keynotes: John R. Gilbert(2018) Prof. UCSB, Ümit V. Çatalyürek(2017) Prof. GA Tech, David A. Bader(2016) Prof. & Chair GA Tech
- 2018 (**Invited Speaker**) Learnable Pooling Layers for Deep Neural Networks, *LinkedIn*, Sunnyvale, CA, 2018.
- (**Invited Speaker**) Learnable Pooling Layers for Deep Neural Networks, *Pandora*, Oakland, CA, 2018.
- (**Invited Speaker**) Deep Learning for Aviation Safety Analysis, *EUROCONTROL*, Paris, France, 2018.
- (**Invited Speaker**) Joint Temporal Link Prediction Requires Rethinking Network (Graph) Models, *High-order Networks Satellite Meeting @ NetSci*, June, 2018.
- (**Invited Speaker**) Joint Temporal Link Prediction Requires Rethinking Network (Graph) Models, *Leibniz Institute for the Social Sciences*, Cologne, Germany, 2018.
- (**Invited Speaker**) Joint Temporal Link Prediction Requires Rethinking Network (Graph) Models, *RWTH Aachen*, Germany, 2018.
- 2017 (**Invited Speaker**) Deep Learning and Survival Bias, *CERIAS Big Data Seminar*, March, 2017.
- (**Invited Lecturer**) Modeling and predicting heterogeneous user behavior, *Summer School Series on Methods for Computational Social Science*, Italy, July, 2017.
- 2015 (**Invited Speaker**) Challenges of Forecasting and Measuring a Complex Networked World, *Analysis in Motion (AIM) Lecture Series* at the Pacific Northeastern National Laboratory, December 13, 2015.
- (**Keynote**) Challenges of Forecasting and Measuring a Complex Networked World, *Keynote* at the 7th Annual Workshop On Simplifying Complex Networks For Practitioners, WWW 2015.
- (**Invited Speaker**) The Big Data Paradox, Invited speaker at the ICWSM'15 Workshop for Modeling and Mining Temporal Interactions, 2015.

| | |
|------|--|
| | (Invited Speaker) 4th Gavriel Salvendy Symposium on Information Engineering, Purdue University, 2015. |
| | Pre-Purdue: |
| 2014 | (Invited Speaker) Modeling and predicting the growth and death of membership-based Internet startups, The Chinese University of Hong Kong, Apr. 2014 |
| | (Invited Speaker) Modeling and predicting the growth and death of membership-based Internet startups, ACM Talk at Loyla University, Mar. 2014 |
| 2013 | (Invited Speaker) The Socio-monetary Incentives of Online Social Network Malware Campaigns, Carnegie Mellon University, Oct. 2013 |
| | (Invited Speaker) Modeling and predicting the growth and death of membership-based Internet startups, Carnegie Mellon University, Sep. 2013 |
| | (Invited Speaker) Understanding the growth and death of social networks: evidence and model, BBN Technologies, July 2013 |
| | (Invited Speaker) Understanding the growth and death of online social networks: evidence and model, Boston University, Jul. 2013 |
| | (Invited Speaker) Online Myopic Network Covering, Boston University, Mar. 2013 |
| 2012 | (Invited Speaker) Walking in a Changing World: Characterizing Random Walks on Dynamic Networks, UCSB, Oct. 2012 |
| | (Invited Speaker) Characterizing Continuous Time Random Walks on Time Varying Graphs. MoBS, Department of Physics, Northeastern University, Apr. 2012 |
| | (Invited Speaker) Searching on Complex Networks. INRIA, Sophia-Antipolis, Mar. 2012 |
| | (Invited Speaker) Exploring Networks with Random Walks. Technicolor Labs, Feb. 2012 |
| | (Invited Speaker) Exploring Networks with Random Walks. Bell Labs, Murray Hill, NJ. Sep. 2012 |
| | (Invited Speaker) Understanding Complex Networks through Incomplete Information: Mistakes, Myths, and Positive Steps. <i>Invited Speaker</i> . Brazilian Computing Society Conference (CSBC), the premier annual computer science event in Brazil, 2010 |

INTERDISCIPLINARY OUTREACH

Ph.D. Committee of Tianyang Hu (Statistics, Guang Cheng, 2019–2020)

Ph.D. Committee of Xinlin Tao (Statistics, Guang Cheng, 2019–)

Ph.D. Committee of Georgios Georgalis (AAE, Karen Marais, 2017–)

Ph.D. Committee of Nicoletta Fala (AAE, Karen Marais, 2016–2018 [graduated])

Research Collaborators: Karen Marais (AAE), Timothee Pourpoint (AAE), Dimitri Peroulis (ECE), John Sutherland (ME), M. Ashraf Alam (ECE).

Invited Speaker: (*SMART Industry day*) Understanding the role of machine learning for industrial applications, January 14, 2021

Invited Speaker: New Directions in Representation Learning of Logical Formulae, PurPI Workshop @ Purdue, Sept. 2019.

Invited Lecture: (*SMART Industry day*) Understanding the role of machine learning for industrial applications, May 15, 2019

Invited Speaker: New Directions in Representation Learning of Graphs, Complex Systems Workshop @ Purdue, May 14. 2019.

Invited Speaker: (*NSF-sponsored Workshop on Broadening Participation in Computer Science*), Machine Learning and Student Trajectories, Santa Clara University, CA, March 23, 2019

Invited Speaker: (*SMART Industry day*) Machine learning for industrial applications, May 16, 2019

Invited Speaker: New Directions in Representation Learning of Graphs, Quantum AI Workshop @ Purdue, Sept 5,. 2018.

Invited Speaker: (*SMART Industry day*) Machine learning for industrial applications, May 1st, 2018

Invited Speaker: (*SMART Industry day*) Machine learning for chemical sensing, May 10, 2017

Invited Speaker: On the Role of Neural Networks on Engineering Challenges, *NASA Workshop on Engineering Challenges*, March, 2017.

Invited Speaker: Purdue Engineering Advisory Committee. Embracing Complexity in Machine Learning without Black Boxes, October 2016.

UNIVERSITY ACTIVITIES

Data Science Major:

Undergraduate core course development for the Data Science Major: Large Scale Data Analytics (3 credits, required course) (Spring 2019)

Other Service:

Member, Data Science Major and Professional Degree Development Committee (Fall 2016, 2017)

Member, Faculty Search Committee (Fall 2019, Spring 2018)

Member, Ph.D. Admissions Committee (Fall 2021, Spring 2022, Fall 2020, Spring 2019, Fall & Spring 2017 , Fall & Spring 2016, Fall 2015)

Member, Graduate Study Committee (Fall 2020, Spring 2019, January 2018 – May 2018)

Purdue CS Colloquia Co-chair (Fall 2021, Spring 2022)

PhD Committee Member:

Firoze, Adnan (cs-graphics, Daniel G. Aliaga)

Gupta, Prerit (cs-ml, Dan Goldwasser)

He, Liu (cs-graphics, Daniel G. Aliaga)

Lee, Younghun (cs-ml, Dan Goldwasser)

An, Shengwei (cs-security, Xiangyu Zhang)

Das, Siddhartha shan (cs-cse, Alex Pothen)

Mohammad-Mohsen Minaei-Bidgoli (CS-Security, Aniket Kate)

Tianyang Hu (Statistics, Guang Cheng, 2019–2020)

Xinlin Tao (Statistics, Guang Cheng, 2019–)

Rania Ibrahim (CS-CSE, David Gleich)

Fei Wang (CS-PL, Tiark Rompf)

Mengyue Hang (CS-ML, Jennifer Neville)

Georgios Georgalis (AAE, Karen Marais, 2017–)

Yingqi Liu (CS-Security, Xiangyu Zhang)

Giselle Marie Zeno Torres (CS-ML, Jennifer Neville)

Nicoletta Fala (AAE, Karen Marais , 2016–2018 [graduated])

Jihwan Lee (CS-DB, Sunil Prabhakar)

Tao Wu (CS-CSE, David Gleich).

External Member: Ph.D. Defense Committee of Janaina Gomide (UFRJ), August 2019

External Member: Ph.D. Defense Committee of Fabricio Murai (UMass Amherst), August 2016

MSc Committee Member:

M.S. Committee of John Moore (ML, adviser Jennifer Neville (CS))

Student Engagement Talks:

Invited Speaker, *Purdue INFORMS Chapter*, October 2018.

Undergraduate Engagement Seminar Talk, (CS 397), Fall 2015, 2016

Industry Engagement Talks:

Industry Engagement Talk for CERIAS and Purdue University: Monsanto (2018), Infosys (2018), Alliance Insurance (2018), ARL (2018), DoD (2018), Eli Lilly (2017), Ford (2017), Elanco (2016)

Outreach:

Created the Purdue CS Minds Website to Advertise CS Data Science faculty (<http://minds.cs.purdue.edu>)

PROFESSIONAL ACTIVITIES

Conference Leadership/Organization (At Purdue):

Founding member and organization committee of Machine Learning and Network Science Symposium (@ NetSci 2020, 2019, 2018, 2017)

Selected Technical Program Committees (At Purdue):

ICLR SPC, 2022, 2021, 2020, 2019, 2018

AAAI SPC, 2022, 2021, 2020, 2019, 2018

ICML 2021, 2020, 2019, 2018, 2017, 2016

NeurIPS 2021, 2020, 2019, 2018, 2017

SDM SPC, 2020, 2019

SIGKDD 2021, 2020, 2019, 2018, 2017, 2016

AISTATS 2019, 2018

WWW 2019, 2018, 2017, 2016, 2015.

NSF Panels (At Purdue):

NSF Panelist: III (2021, 2020, 2019)

NSF Panelist: BIGDATA (2017)

Journal Reviewer (At Purdue):

Editorial Board (2021) Reviewer at Journal of Machine Learning Research (2021, 2020, 2019, 2017)

IEEE Transactions on Knowledge and Data Engineering (2019, 2018)

ACM Transactions on Knowledge Discovery from Data (2018, 2017, 2016)

Data Mining and Knowledge Discovery (2016)

PNAS (2015)

Outreach (At Purdue):

(2017–present) Founding member of the Curated Pathways to Innovation, a pilot program at the Ocala school district in San Jose, CA, with the goal of using technology to set women and underrepresented minorities on the path to STEM careers: www.curatedpathways.org

CURRENT & PAST GRANTS

Active grants (At Purdue): 5 grants (2x NSF, 1x private).

Total dollar amount in active grants under Ribeiro's control: \$1,094,542.00

Past grants (At Purdue): 2 grants (1x NSF, 2x DoD, 1x FAA).

Total dollar amount in past grants under Ribeiro's control: \$1,053,114.00

Total dollar amount (active + past) of external grants at Purdue under Ribeiro's control: \$2,147,656.00

| | |
|---------------------------------------|---|
| 1. Agency/Title of Grant: | NSF- CAREER (IIS: Robust Intelligence): A Novel Blueprint for Representation Learning of Relational Invariance |
| 2. Duration of Funding: | 05/01/20 - 04/30/25 |
| 3. Total Amount of Award: | \$491,161 |
| 4. Your Role: | (PI) |
| 5. Percentage of funding responsible: | 100% |

| | |
|---------------------------------------|--|
| 1. Agency/Title of Grant: | NSF- FMITF:Track I: Symbolic Reasoning with Graph Networks |
| 2. Duration of Funding: | 10/01/19 - 09/30/23 |
| 3. Total Amount of Award: | \$750,000 |
| 4. Your Role: | (Co-PI) |
| 5. Percentage of funding responsible: | 33% |

| | |
|---------------------------------------|---|
| 1. Agency/Title of Grant: | DoD - Discovering Network Processes in Multilayer Time-evolving Networks under Incomplete Information |
| 2. Duration of Funding: | 04/01/16 - 12/31/19 |
| 3. Total Amount of Award: | \$1,649,382 |
| 4. Your Role: | (Purdue PI, Overall Technical Lead) |
| 5. Percentage of funding responsible: | 33% |

| | |
|---------------------------------------|---|
| 1. Agency/Title of Grant: | DoD - Data Science Approaches to Improve System Reliability |
| 2. Duration of Funding: | 06/23/17 - 06/22/19 |
| 3. Total Amount of Award: | \$146,000 |
| 4. Your Role: | (CO-PI) (PI Karen Marais, Purdue AAE) |
| 5. Percentage of funding responsible: | 50% |

| | |
|---------------------------------------|--|
| 1. Agency/Title of Grant: | YWCA – Curated Pathways of Innovation to Help URM Middle Schoolers Choose STEM Careers |
| 2. Duration of Funding: | 01/01/17 - 12/31/21 |
| 3. Total Amount of Award: | \$353,381 |
| 4. Your Role: | (PI) |
| 5. Percentage of funding responsible: | 100% |

| | |
|---------------------------------------|---|
| 1. Agency/Title of Grant: | IDSI (Purdue) – Blazing Fast Chemical Sensing |
| 2. Duration of Funding: | 06/01/18 – 05/31/20 |
| 3. Total Amount of Award: | \$215,163 |
| 4. Your Role: | (PI) |
| 5. Percentage of funding responsible: | 50% |

| | |
|---------------------------------------|---|
| 1. Agency/Title of Grant: | IDSI (Purdue) – Robust Machine Learning |
| 2. Duration of Funding: | 06/01/18 – 05/31/20 |
| 3. Total Amount of Award: | \$265,000 |
| 4. Your Role: | (PI) |
| 5. Percentage of funding responsible: | 20% |

| | |
|---------------------------------------|---|
| 1. Agency/Title of Grant: | Federal Aviation Administration – Predicting safety events in general aviation from smartphone sensor trajectories. |
| 2. Duration of Funding: | 07/01/16 - 06/30/18 |
| 3. Total Amount of Award: | \$290,000 |
| 4. Your Role: | (CO-PI) |
| 5. Percentage of funding responsible: | 50% |

| | |
|---------------------------------------|---|
| 1. Agency/Title of Grant: | Lilly Endowment – Wabash Heartland Innovation Network Economic Ecosystem – Machine Learning component |
| 2. Duration of Funding: | 11/17/17 - 08/31/22 |
| 3. Total Amount of Award: | \$40,000,000 |
| 4. Your Role: | (CO-PI) (PI Ali Shakouri, Purdue ECE) |
| 5. Percentage of funding responsible: | 1% |

| | |
|---------------------------------------|--|
| 1. Agency/Title of Grant: | NSF – CICI: CE: Enhancing Cybersecurity for Broadening Data-Driven Research and Partnerships |
| 2. Duration of Funding: | 10/01/17 - 09/30/20 |
| 3. Total Amount of Award: | \$841,506 (841,506) |
| 4. Your Role: | (CO-PI) |
| 5. Percentage of funding responsible: | 8% |

| | |
|---------------------------------------|--|
| 1. Agency/Title of Grant: | NSF – XSEDE Extended Collaborative Support |
| 2. Duration of Funding: | 02/08/2018 – 02/07/2019 |
| 3. Total Amount of Award: | \$8,345 (computing support) |
| 4. Your Role: | PI |
| 5. Percentage of funding responsible: | 100% |

| | |
|---------------------------------------|--|
| 1. Agency/Title of Grant: | Purdue – Major Scientific Equipment Award |
| 2. Duration of Funding: | 05/01/17 - 12/31/17 |
| 3. Total Amount of Award: | \$267,735 |
| 4. Your Role: | (CO-PI, helped develop the requirements and justification) |
| 5. Percentage of funding responsible: | 25% |

| | |
|---------------------------------------|---|
| 1. Agency/Title of Grant: | (Prior to Purdue) NSF – NetSE Modeling and Analysis of Network Dynamics |
| 2. Duration of Funding: | 05/01/11 - 05/31/15 |
| 3. Total Amount of Award: | \$900,000 |
| 4. Your Role: | (CO-PI) |
| 5. Percentage of funding responsible: | 33% |

| | |
|---------------------------------------|---|
| 1. Agency/Title of Grant: | (Prior to Purdue) ARO – ExtraPro: Extrapolating Properties and Behaviors of Information Dissemination |
| 2. Duration of Funding: | 10/01/12 - 11/31/14 |
| 3. Total Amount of Award: | \$780,235 |
| 4. Your Role: | (CO-PI) |
| 5. Percentage of funding responsible: | 25% |

| | |
|---------------------------------------|---|
| 1. Agency/Title of Grant: | (Prior to Purdue) ARO – Extracting Network Knowledge: Exploring Large Graphs in a Dynamic and Uncertain World |
| 2. Duration of Funding: | 10/01/11 - 11/31/12 |
| 3. Total Amount of Award: | \$109,770 |
| 4. Your Role: | (CO-PI) |
| 5. Percentage of funding responsible: | 50% |

| | |
|---------------------------------------|--------------------------------------|
| 1. Agency/Title of Grant: | Industrial gifts – Google and Nvidia |
| 2. Duration of Funding: | 10/01/17 – |
| 3. Total Amount of Award: | \$5,749 |
| 4. Your Role: | (PI) |
| 5. Percentage of funding responsible: | 100% |

Last updated: October 14, 2021