

Cong Mu

cmu2@jhu.edu

<https://congm.github.io>

Education

Johns Hopkins University (Baltimore, MD)	08/2017 – 05/2023
Doctor of Philosophy Major: Applied Mathematics and Statistics	(Expected)
Master of Science in Engineering Major: Applied Mathematics and Statistics	
Master of Science in Engineering Major: Computer Science	
 Sun Yat-Sen University (Guangzhou, China)	08/2013 – 06/2017
Bachelor of Science Major: Statistics Minor: Finance	

Research Experience

Community Detection for SBM with Vertex Covariates | Johns Hopkins University

- Develop model-based spectral algorithms for clustering vertices in SBM with vertex covariates; assess effects of observed and unobserved vertex heterogeneity on block recovery; employ Chernoff information to analytically compare the performance and derive the Chernoff ratio expression for certain models of interest.
- Key words: **Spectral Graph Inference, Community Detection, Stochastic Blockmodel, Vertex Covariates, Chernoff Ratio.**
- **Mu, C., Mele, A., Hao, L., Cape, J., Athreya, A., & Priebe, C. E. (2020).** On spectral algorithms for community detection in stochastic blockmodel graphs with vertex covariates. arXiv preprint arXiv:2007.02156. [[Preprint](#)]

Statistical Models for Large Networks | Johns Hopkins University

- Built statistical models and algorithms that could be scaled to analyze large networks; estimated and simulated network formation models using high performance computing; developed R package with research objectives such as identifying the community structure.
- Key words: **Hierarchical Exponential-Family Random Graph Models, (Generalized) Random Dot Product Graph, SBM with Covariates, Variational Generalized EM Algorithms, Minorize-Maximization, Parallel Computation.**
- Dynamic Network Structure [[Shiny App](#)]
- (G)RDPG with Covariates [[R Package](#)]

Automatic Tools for Dash Cam Video | Johns Hopkins University

- Developed automatic tools for analyzing and annotating video stream with relevant information such as timing, speed, traffic, accidents, objects and etc.
- Key words: **Structural Similarity Index, Oriented FAST and Rotated BRIEF, Image Hashing, Robust Image Similarity Measure, Deep Neural Networks.**
- **Mu, C., & Budavári, T. (2018).** Dash Cam Video Analysis: Laptimes and Beyond. Poster presented at 2018 IDIES Annual Symposium, Baltimore, MD. [[Poster](#)]

Therapy Functional Measures | Johns Hopkins University & Johns Hopkins Hospital

- Identified patterns in patient functional trajectories; measured causal effect of different physical therapy dosage regimes on patient functional status; constructed features and built model to predict AMPAC score to optimize physical therapy in the hospital.
- Key words: **Linear Mixed-Effect Model, ARIMA.**
- Crockett M., Mu, C., & Dahbura, A. T. (2018). Predictive Analytics for Patient Mobility Using AM-PAC. Poster presented at 2018 Johns Hopkins Research Symposium on Engineering in Healthcare, Baltimore, MD. [[Poster](#)]

Professional Experience

Data & Applied Scientist Intern | Microsoft (Virtual) 06/2021 – 08/2021

- Pre-trained language models for a few selected languages using Microsoft Turing technology. Fine-tuned on different downstream tasks and built subsequent applications. (**Transformer, DeBERTa, DeepSpeed**)
- Created a demo using pre-trained and fine-tuned models. Presented in Azure Data ML Talk Series. (**Python Dash**)

Analyst Intern | GF Fund Management (Guangzhou, China) 11/2016 – 04/2017

- Selected features to build market emotional indicators and developed model to predict market, achieved 92% accuracy. (**XGBoost, Random Forest, Lasso**)
- Mined key business data and constructed data reporting system; analyzed and visualized product and user data to provide decision support. (**R Markdown, R Shiny**)

Data Scientist Intern | Research Center of Statistical Science (Guangzhou, China) 02/2016 – 10/2016

- Classified users to optimize delivery of advertisements and constructed program recommendation system; predicted whether user will be secondary loans to explore potential customers and evaluate risk in advance. (**Collaborative Filtering, Clustering**)
- Presented in 9th China-R Conference and Regional Data Science Conference on how to use R to interact and share ideas by using Shiny in R to make an interactive interface rapidly. (**R Shiny**)

Teaching Experience

Teaching Assistant | Johns Hopkins University

- Data Mining (Spring 2018, Spring 2019, Fall 2019)
- Applied Statistics and Data Analysis (Fall 2019)
- Applied Statistics and Data Analysis II (Spring 2020)
- Statistical Theory (Fall 2020, Fall 2021)

Honors & Awards

MINDS Data Science Fellowship, Johns Hopkins University	2021
Creel Family Engineering Fellowship, Johns Hopkins University	2019 – 2020
University Scholarship for Outstanding Student, Sun Yat-Sen University	2014 – 2016

Skills

R, Python, Matlab, C/C++, SQL, TensorFlow, PyTorch, OpenCV, Data Visualization