Sara Mohammad Taheri

October 31, 2019

Phone: (781) 363-5446
Email: mohammadtaheri.s@husky.neu.edu

Khoury College of Computer Sciences
Northeastern University, Boston MA

Research interests

- Causal inference and its applications especially in biology.
- Causal structure learning from observational biomolecular data
- Probabilistic graphical models (PGMs)
- Statistical and machine-learning methods.

Education

2016-current PhD student, Computer Science, Northeastern University, USA
Advisor: O. Vitek

2014 MS, Mathematics, Sharif University of Technology, Iran.
Thesis: *"On mixing time for some Markov Chain Monte Carlo"*
Advisor: K. Alishahi

2010 BS. in Mathematics, Sharif University of Technology, Iran.
Diploma: *"Roman Domination in Graph Theory"*
Advisor: S. Akbari

Professional Experience

09/2016-present Graduate Assistant. Northeastern University, USA.

- Developed an online open-source R-based software for systems suitability and statistical process control in mass spectrometry-based quantitative proteomics. It is hosted under `<https://eralpdogu.shinyapps.io/msstatsqc/>` which is currently used by researchers for statistical analysis and quality control of their experiments.
- Developed two R packages, MSstatsQC and MSstatsQCgui package and submitted to bioconductor.
- Current research is on finding causal regulatory networks from observational biomolecular data. The nodes represent genes, proteins, transcripts or metabolites, and directed edges represent causal regulatory relationships.

06/2018-09/2018 Precision Medicine Research Intern . GNS Healthcare Company, USA.

- Enhanced constraint-based structure learning algorithms algorithms in bnlearn (R package) by adding an optional input to the algorithms to use for mitigating the effect of measurement error on network inference. This modification caused significant improvement in the results of the algorithm.This work was a follow up on the internship in 2017 with the same company.

06/2017-09/2017 Precision Medicine Research Intern . GNS Healthcare Company, USA.

- Improved constraint-based structure learning algorithms by mitigating the effect of measurement error on network inference. The idea was to correct the estimate of Covariance matrix from observational data by estimating the variance of measurement error for each variable.

05/2019 Instructor for Github and RMarkdown Northeastern University, USA.

05/2018

05/2017

- "May Institute - Computation and statistics for mass spectrometry and proteomics."
- Supported by the 1R25EB023929-01 award from National Institute of Health (NIH), and by German network for bioinformatics infrastructure.

09/2009-11/2017 Teaching assistant. Northeastern University, USA.

- MS course "Introduction to Data Management and Processing". Grading, Office hours. 2017
- MS course "Machine Learning". Grading, Office hours. 2017
- MS course "Collecting, Storing, and Retrieving Data". Office hours. 2016
- MS course "Topics in Statistics and Data Analysis". Grading, office hours. 2016
- BS course "Number Theory". Grading, office hours. 2009

02/2009-03/2010 Volunteer researcher. Institute for Theoretical Physics & Mathematics, Iran.

- Research on Roman Domination in graph theory.
  Advisor: S.Akbari

02/2009-08/2015 K-12 Mathematics teacher, Tehran, Iran.

Programming skills

R (including `Shiny`, `tidyr`, `dplyr`, `ggplot2`, package development, markdown documentation), Python, LaTeX, Bash scripting

Awards

2006 Fellowship for B.Sc. in Mathematics, Sharif University of Technology, Iran

Publications

1. E. Dogu, S. Mohammad-Taheri, S. E. Abbatiello, M. S.Bereman, B. MacLean, B. Schilling, O. Vitek, "MSstatsQC: Longitudinal System Suitability Monitoring and Quality Control for Targeted Proteomic Experiments". *Molecular and Cellular Proteomics.* 2017

2. E. Dogu, S. Mohammad Taheri, R. Olivelia, F. Marty, Ian Lienert, L. Reiter, E. Sabid, O. Vitek, "MSstatsQC 2.0: R/Bioconductor package for statistical quality control of mass spectrometry-based proteomic experiments". *J. Proteome Res.*2018

Posters

1. Mohammad Taheri Sara, Furchtgott Leon, Hayete Boris, Vitek Olga, 2019. "Improving Structure Learning of Bayesian Network in Experiments with Complex Designs." SAMSI 2019

2. Dogu Eralp, Mohammed Taheri Sara, Pujol Roger, Sabido Eduard, Vtek Olga, 2018. "New developments in MSstatsQC: a new R/Bioconductor package MSstatsQCgui, quality control for targeted and discovery proteomics workflows, machine learning based quality monitoring." ASMS 2018

3. Dogu Eralp, Mohammed Taheri Sara, Abbatiello Susan, Bereman Michael, Maclean Brendan, Schillng Birgit, Vitek Olga, 2017. "MSstatsQC: Longitudinal system suitability and quality control for proteomic experiments." ASMS Conference on Mass Spectrometry and Allied Topics.

Presentations

2019

- Improving structure learning of Bayesian network in experiments with complex designs, SAMSI Conference, North Carolina, USA
- How to infer network structure in presence of measurement error, Princeton University, Department of molecular biology, USA