Chong Wu

A460 Mayo Building, MMC 303
420 Delaware St SE, Minneapolis, MN, 55455
✉: chongwu@umn.edu; ☎: 612-516-2222
: http://wuchong.org

RESEARCH INTEREST

High-dimensional data, big data, machine learning, statistical genetics

EDUCATION

University of Minnesota (UMN) Minneapolis, MN
Ph.D., Biostatistics May 2018 (Expected)
Advisors: Prof. Weihua Guan & Prof. Wei Pan GPA: 3.8/4.0
Thesis: Statistical methods for high-dimensional genetic and genomic data

Huazhong University of Science & Technology Wuhan, China
B.S., Applied Mathematics Jun. 2013
Advisor: Prof. Xiaoyang Zhou GPA: 3.9/4.0; Ranking: 1/29

RESEARCH EXPERIENCE

Graduate Research Assistant, Advisor: Prof. Wei Pan Sept. 2014–Present
Research areas: high-dimensional genetic and genomic data, machine learning, statistical genetics
Division of Biostatistics, University of Minnesota

Graduate Research Assistant, Advisor: Prof. Weihua Guan Sept. 2013–May 2017
Research areas: statistical genetics, methylation data
Division of Biostatistics, University of Minnesota

Undergraduate Research Assistant, Advisor: Prof. Xiaoyang Zhou 2011–2013
Research areas: data mining on regional economic indicators
School of Mathematics and Statistics, Huazhong University of Science & Technology

TEACHING EXPERIENCE

Guest Instructor Jan. 2017–May 2017
Course Name: PUBH 7475/8475 Statistical Learning and Data Mining
Division of Biostatistics, University of Minnesota

Lab Instructor, Instructor: Prof. William Thomas Jan. 2014–May 2014
Course Name: PUBH 6451 Biostatistics II
Division of Biostatistics, University of Minnesota

Teaching Assistant, Instructor: Prof. Ann M. Brearley Sept. 2013–Dec. 2014
Course Name: PUBH 6414 Biostatistical Methods I (Online Section)

Division of Biostatistics, University of Minnesota

PEER-REVIEWED PUBLICATIONS

1. Xu, Z., **Wu, C.**, Wei, P., and Pan, W. (2017+). A powerful framework for integrating eQTL and GWAS summary data. Accepted by *Genetics*, early online. (IF: 4.6)

2. Liu, B., **Wu, C.**, Shen, X., and Pan, W. (2017). A novel and efficient algorithm for de novo discovery of mutated driver pathways. *Annals of Applied Statistics*, 17(3):1481–1512.

3. Xu, Z., **Wu, C.**, Pan, W., and Alzheimer's Disease Neuroimaging Initiative (ADNI). (2017). Imaging-wide association study: integrating imaging endophenotypes in GWAS. *NeuroImage*, 159:159–169.
(IF: 5.8. This paper won a platform presentation at the American Society of Human Genetics (ASHG) 2017 Annual Meeting.)

4. **Wu, C.***, Kwon, S.*, Shen, X., and Pan, W. (2016). A new algorithm and theory for penalized regression-based clustering. *Journal of Machine Learning Research*, 17(188):1–25.
(* Co-first author. IF: 5.0, a leading journal in machine learning area.)

5. **Wu, C.**, Chen, J., Kim, J., and Pan, W. (2016). An adaptive association test for microbiome data. *Genome Medicine*, 8(1):1–12.
(IF: 7.1. This paper won the 2016 Joint Statistical Meetings (JSM) Distinguished Student Paper Award on Statistics in Genomics and Genetics Section.)

6. **Wu, C.**, Demerath, E. W., Pankow, J. S., Bressler, J., Fornage, M., Grove, M. L., Chen, W., and Guan, W. (2016). Imputation of missing covariate values in epigenome-wide analysis of DNA methylation data. *Epigenetics*, 11(2):132–139.
(IF: 4.4, the official journal of the DNA Methylation Society.)

7. Bose, M., **Wu, C.**, Pankow, J. S., Demerath, E. W., Bressler, J., Fornage, M., Grove, M. L., Mosley, T. H., Hicks, C., North, K., Kao, W. H., Zhang, Y., Boerwinkle, E., and Guan, W. (2014). Evaluation of microarray-based DNA methylation measurement using technical replicates: the Atherosclerosis Risk In Communities (ARIC) Study. *BMC Bioinformatics*, 15(1):1–10. (IF: 2.4)

PEER-REVIEWED PROCEEDINGS

8. Park, J. Y., **Wu, C.**, and Pan, W. (2017+). An adaptive gene-level association test for pedigree data. Accepted by *BMC Proceedings*, a special issue for the Genetic Analysis Workshop (GAW 20).

9. **Wu, C.**, Park, J.Y., Guan, W., and Pan, W. (2017+). A powerful gene-based test for methylation data. Accepted by *BMC Proceedings*, a special issue for the Genetic Analysis Workshop (GAW 20).

SUBMITTED AND IN PREPARATION

10. **Wu, C.**, Xu, G., Shen, X., and Pan, W. (2017+). An adaptive test on a high-dimensional parameter in the presence of a high-dimensional nuisance parameter in GLM with application to detect gene-environment interactions. Manuscript.
(Job talk manuscript, to be submitted to *Journal of the American Statistical Association*.)

11. **Wu, C.** and Pan, W. (2017+). Integrating eQTL data with GWAS summary statistics in pathway-based analysis. Submitted.
(This paper won a poster talk (top 24 posters among about 3000 posters) at ASHG 2017 Annual Meeting.)

12. **Wu, C.***, Xu, G., and Pan, W.* (2017+). An adaptive test on high dimensional parameters in generalized linear models. *Statistica Sinica*, revised and resubmitted. (* Corresponding author)

13. Zhu, L., Li, Y., Chen, Y., Carrera, C., **Wu, C.**, and Fork, A. (2017+). Comparison between two post-dentin bond strength measurement methods. *Scientific Reports*, revised and resubmitted.

14. Park, J.Y., **Wu, C.**, Basu, S., McGue, M., and Pan, W. (2017+). Adaptive SNP set association testing in generalized linear mixed models with application to family studies. *Behavior Genetics*, revised and resubmitted.

15. **Wu, C.**, Xu, G., and Pan, W. An adaptive and powerful test for high dimensional covariance matrices. In Preparation.

16. **Wu, C.** and Pan, W. Network-based support vector machines with a new penalty. In Preparation.

17. **Wu, C.**, Kim, J., Wei, P., and Pan, W. Adaptive test for meta-analysis of rare variants in sequencing association studies. In preparation.

SOFTWARE

- Owner and maintainer for the following R packages:
  - prclust: Penalized Regression-Based Clustering Methods;
  - MiSPU: Microbiome Based Sum of Powered Score (MiSPU) Tests;
  - aSPU2: A New Version of Adaptive Sum of Powered Score (aSPU) Test;
  - GLMaSPU: Adaptive Tests on High Dimensional Parameters in Generalized Linear Models;
  - glmtlp: Truncated Lasso Regularized Generalized Linear Models.
- Contributor: Theano (Pull request: # 6130).

PRESENTATIONS

- "An adaptive association test for microbiome data."
  - Presentation at Eastern North American Region (ENAR) 2016 Spring Meeting

Austin, TX Mar. 2016

– Poster presented at School of Public Health Research Day, University of Minnesota
Minneapolis, MN Apr. 2016

– Poster presented at Minnesota Supercomputing Institute (MSI) Research Exhibition
Minneapolis, MN Apr. 2016

– Presentation at 2016 Joint Statistical Meetings (JSM)
Chicago, IL Aug. 2016

- "Iterative PCA in epigenome-wide association studies."
  – Poster presented at American Society of Human Genetics (ASHG) 2016 Annual Meeting
  Vancouver, BC, Canada Oct. 2016

- "A gene-level adaptive association test for methylation data."
  – Presentation at Genetic Analysis Workshop (GAW) 20
  San Diego, CA Mar. 2017

- "An adaptive test on high dimensional parameters in GLMs."
  – Presentation at ENAR 2017 Spring Meeting
  Washington, DC Mar. 2017

  – Poster presented at MSI Research Exhibition
  Minneapolis, MN Apr. 2017

  – Presentation at 2017 JSM
  Baltimore, MD Aug. 2017

- "Imaging-wide association study: integrating imaging endophenotypes in GWAS"
  – Invited presentation at 2017 JSM (On Prof. Wei Pan's behalf)
  Baltimore, MD Aug. 2017

  – Invited presentation at Third Annual Kliakhandler Conference (On Prof. Wei Pan's behalf)
  Houghton, MI Aug. 2017

- "Integrating eQTL data with GWAS summary statistics in pathway-based analysis"
  – Poster talk at ASHG 2017 Annual Meeting
  Orlando, FL Oct. 2017

HONORS & AWARDS

- **Pre-Doctoral Trainee Award** Oct. 2017
  – Association of Chinese Geneticists in America (ACGA)
- **Poster Award** May 2017
  – University of Minnesota Chapter of Sigma Xi
- **Elected to Delta Omega (Public Health Honorary Society)** May 2017
  – University of Minnesota

- **Elected to Sigma Xi (The Scientific Research Society)** May 2017
  – University of Minnesota
- **Distinguished Student Paper Award**, Genomics and Genetics Section Aug. 2016
  – 2016 Joint Statistical Meetings
- **Doctoral Dissertation Fellowship** 2016–2017
  – University of Minnesota
- **Travel Award**, Computational Neuroscience Summer School Jul. 2015
  – Statistical and Applied Mathematical Sciences Institute
- **Dean's Ph.D. Scholarship** Sept. 2013
  – University of Minnesota
- **Honorable Mention in Mathematical Contest in Modeling** Apr. 2012
  – Consortium for Mathematics and Its Application
- **National Scholarship** Sept. 2011
  – Ministry of Education, China

WORK EXPERIENCE

Biostatistics Intern, Advisor: Dr. Jeff Budd May 2014–Aug. 2014
Roles: Developed programs in S+ for statistical analysis
Beckman Coulter, Chaska, MN

PROFESSIONAL ACTIVITIES

Manuscript Reviewer

- Journal: Statistics in Biosciences, Journal of Theoretical Biology, Biometrical Journal, Computational Statistics and Data Analysis, Genetic Epidemiology, Bioinformatics
- Conference: 2017 Joint Statistical Meetings Student Paper Award Reviewer

Professional Memberships

- Member, American Statistical Association 2014–Present
- Member, Eastern North American Region 2015–Present
- Member, The American Society of Human Genetics 2016–Present

Programming Skills

- R, Python, MATLAB, C++