YUCHEN WANG

+1 (573) 239-5737
ycwang0712@gmail.com
30 Gardner Road, Apartment 4F
Brookline, Massachusetts, 02445

EDUCATION

***Master of Arts* in Statistics**, University of Missouri, Columbia, Missouri **July 2015**

- Cumulative GPA 3.9 with A+ in Probability Theory, Advanced Linear Models and Bayesian Statistics.
- Research work in spatial point patterns and numerical problems in gene expression analysis.
- Authored R packages MethyBayes (CRAN) and sppmix (GitHub).

***Bachelor of Science* in Statistics**, East China Normal University, Shanghai, China **June 2013**

***Minor* in Computer Science**, Shanghai Jiao Tong University, Shanghai, China **June 2013**

RELEVENT EXPERIENCE

Analyst, Marketing Science, In4mation Insights, Needham, Massachusetts **September 2015 – Present**

- Developed a scoring system to define product superiority for tobacco products.
- Participated in several projects including topics like market structure, nested logit models, segmentation and partial least square path models.
- Took part in the development of company's internal Bayesian software (in *C++*) and developed an R interface and shiny front-end. Developed several R packages to facilitate the analysis and management of projects.

Research Assistant, University of Missouri, Columbia, Missouri **August 2014 – May 2015**

- Optimized a Bayesian MCMC algorithm in Fortran for gene data. Reduced computation time from more than 3 hours to less than 20 seconds.
- Parallelized a large-scale Bayes factor computation program in Fortran using both *OpenMP* and *MPI*; Tuned for distributed computing on different kinds of high performance computing clusters.

Research Assistant, East China Normal University, Shanghai, China **July 2012 - June 2014**

- Joined a research team of Statistical Process Control (SPC) in tobacco manufacturing as a part-time research assistant.
  - The only undergraduate student in a research team of 7 members consisting graduate students and faculties.
  - Experienced in analyzing real-world big data (GBs per hour) from observational studies on Amazon EC2 using parallelized algorithms.
  - Designed a control scheme for controlling tobacco dehumidification process, which is the most critical and complicated control unit among all tobacco manufacturing processes.
  - Collaborated with both quality control team at the tobacco manufacturing factory and software engineering team of an IT company. The final quality control product could precisely detect abnormalities of the manufacturing in advance so as to minimize the cost of failure products.
- Hired as a full-time research assistant in charge of the SPC project after graduation. Led a team of 21 members.
  - Developed an R package for the project containing raw data, data cleaning routines, statistical algorithms and visualization tools. Programmed over 2,500 lines of R code for this project.

RELATED SKILLS

Statistical Computing Highly skilled in *R*; Developer of 15+ *R* packages (GitHub: @wangyuchen); Mostly used packages including *MCMCpack* for Bayesian computing; *spatstat* and *splancs* for spatial point pattern analysis; *ggplot2* and *knitr* for graphics and reports; *shiny* user. Proficent in *SAS* procedures including *glm*, *nlin*, *genmod* and *mixed*.

Database Everyday user of *dplyr*, not a fan of *data.table*, but fluent in both; Familiar with relational database and SQL concepts.

General Purpose Programming Solid understanding of object-oriented programming. Experience in integrating *C++* and *Fortran* into *R* packages. Extensive parallel computing experience in *OpenMP* and *MPI* with *Fortran*. Proficient in *MATLAB*.

Specialities in Statistics Solid foundation in mathematical statistics and multivariate analysis; Extensive experience in Bayesian analysis; Spatial analysis with emphasis in point pattern models and spatio-temporal models; Statistical Process Control; Data mining and machine learning.

Operating Systems Everyday Linux (Cent OS) user; *VIM* user. Responsible for HPC cluster management at ECNU.