

# Xiaosu Tong

2101 Cumberland Ave. Apt. 3207,

West Lafayette, IN, 47906

PHONE 765-586-6016 EMAIL tongx@purdue.edu

## OBJECTIVE

To secure an internship in data scientist position with a focus on big and complex data analysis, machine learning.

## EDUCATION

**Purdue University, Department of Statistics, West Lafayette, IN, US**

**August 2012 - Present**

➤ **Ph.D. in Computational Statistics**

Working in the areas of differential privacy, big and complex data analysis, data visualization, and data mining.

➤ **Master of Science in Applied Statistics** GPA: 3.81/4.0.

**May 2012**

**Shenyang Jianzhu University, School of Science, SHENYANG, CHINA**

➤ **Bachelor of Science in Information and Calculation Science**

**July 2010**

## SKILL

- Strong skills in C, R, RHIPE/Hadoop, SAS, Matlab, SPSS, HTML, LINUX operation systems;
- Proficient in data visualization, statistical consulting, experimental design, and data analysis;
- Experience with Vim, Emacs, version control (cvs and svn), Latex.

## RESEARCH EXPERIENCE

**Research Assistant, Professor Bowei Xi**

**August 2012 – April 2013**

- Explored three different methods for modifying covariance matrix to solve non positive-definite matrix issue in LDA and QDA classification with sensitivity and differential privacy concepts.
- Formulated the distribution functions of test statistics with differential privacy to investigate the performance of type I and type II error of test using MATLAB.

**Research Assistant, Professor William Cleveland**

**May 2013 – Present**

- Visualized the critical attributes of Kiva dataset in XDATA project. Discovered two most significant anomalies about default loans in Kiva dataset by using RHIPE in R. Conducted a hierarchical cluster analysis to detect repeated borrowers in Kiva by using Divide and Recombine method with MapReduce model.
- Visualized the all data measurement status of NCAR spatial time series data about temperature and precipitation by using RHIPE in R.

## PUBLICATION

➤ **R/RHIPE/Hadoop for Divide and Recombine** (Technical Report)

**December 2012**

## RELATED EXPERIENCE

**Senior Consultant, Statistical Consulting Service, Purdue University**

**May 2011 – August 2013**

- Abstract statistical models from real life research problems in diverse areas to provide assistance to clients with statistics analysis or probability problems and running of a wide variety of statistical computing programs, including SAS, SPSS, and R. Worked with 33 clients totally.

**Teaching Assistant, STAT301, Department of Statistics, Purdue University**

**May 2011 – May 2012**

- Independently instructed the laboratory course 'Elementary Statistical Methods'.

**Teaching Assistant, STAT695V, Department of Statistics, Purdue University**

**August 2013 – Present**

- Designed all homework assignments for the course, and assisted Prof. William Cleveland with lecture slides of R learning.

## HONORS AND AWARDS

- **National Scholarship** awarded by Department of Education, Shenyang, China (October 2008)
- **2<sup>nd</sup> Prize** in National Mathematical Modeling Contest, Shenyang, China (October 2008)
- **Top 2% of class** Scholarship of Shenyang Jianzhu University, Shenyang, China (October 2007)