Zhengjie Miao

D307 LSRC Building, 308 Research Drive, Durham, NC 27708
zjmiao@cs.duke.edu +1 (718) 916-6276 http://www.cs.duke.edu/~zjmiao

EDUCATION

Duke University, Durham, NC, U.S.A

- Ph.D. in Computer Science — Aug 2017 – Present
  - Advisor: Prof. Sudeepa Roy
  - Research Initial Project: Explaining Wrong Queries Using Small Examples
  - GPA: 3.85 / 4.00
  - Courses: Distributed Information Systems, Computer Vision, Decision Making at Scale

Columbia University, New York, NY, U.S.A

- M.S. in Computer Science — Aug 2015 – Dec 2016
  - GPA: 4.00 / 4.00
  - Courses: Database System Implementation, Advanced Database Systems, Visual Databases, Data Mining, Natural Language Processing, Machine Learning, Advanced Machine Learning, Analysis of Algorithm

Peking University, Beijing, P.R. China

- B.S. in Computer Science and Technology — Sep 2011 – Jul 2015
  - GPA: 3.50 / 4.00
- B.A. in Economics — Sep 2012 – Jul 2015

RESEARCH INTERESTS

Databases, Data Provenance, Data Mining, Visual Analytics

PUBLICATIONS

Explaining Wrong Queries Using Small Examples
Zhengjie Miao, Sudeepa Roy, and Jun Yang
in *ACM SIGMOD International Conference on Management of Data (SIGMOD)*, Jun 2019 (to appear)

Going Beyond Provenance: Explaining Query Answers with Pattern-based Counterbalances
Zhengjie Miao*, Qitian Zeng*, Boris Glavic, and Sudeepa Roy
in *ACM SIGMOD International Conference on Management of Data (SIGMOD)*, Jun 2019 (to appear, * denotes equal contribution)

Combining Design and Performance in a Data Visualization Management System
Eugene Wu, Fotis Psallidas, Zhengjie Miao, Haoci Zhang, Laura Rettig, Yifan Wu, Thibault Sellam
in *Conference on Innovative Data Systems Research (CIDR)*, Jan 2017

Story Explorer: A Visual Analysis Tool for Heterogeneous Text Data
Chenglong Wang, Zhengjie Miao, Siming Chen, Zipeng Liu, Zuchao Wang, Zhenhuang Wang, and Xiaoru Yuan
in *IEEE Conference on Visual Analytics Science and Technology (VAST)*, Nov 2014

MovementFinder: A Multi-filter Visual Analytics Design for Movement Data Investigation
Siming Chen, Zuchao Wang, Zipeng Liu, Zhenhuang Wang, Chenglong Wang, Zhengjie Miao, and Xiaoru Yuan
in *IEEE Conference on Visual Analytics Science and Technology (VAST)*, Nov 2014

A Platform For Collaborative Visual Analysis on Streaming Tweets
Zipeng Liu, Zhenhuang Wang, Siming Chen, Zuchao Wang, Zhengjie Miao, Xiaoru Yuan
in *IEEE Conference on Visual Analytics Science and Technology (VAST)*, Nov 2014

RESEARCH EXPERIENCE

Database Research Group, Duke University
Research Assistant, advised by Prof. Sudeepa Roy and Prof. Jun Yang Aug 2017 – present

- **Explaining Surprising Query Answers Using Patterns**
  - A framework that provides explanations for surprising outcomes in the results of an aggregate query by finding underlying pattern that holds over the data but is violated by the result
  - Formalized the concept of aggregate regression patterns and the definition of counterbalancing explanations using aggregate regression patterns
  - Designed and implemented the explanation generating algorithm
  - This work has been accepted for publication in SIGMOD'19
- **Explaining Wrong Queries Using Small Examples**
  - A debugging tool that takes a database instance and two queries as input and outputs the smallest subinstance of the database over which the input queries return different results
  - Designed and implemented algorithms to find the smallest subinstance using data provenance and SMT solvers
  - Integrated the algorithm into a web-based tool and deployed the tool in an undergraduate database course to help students debug their queries
  - This work has been accepted for publication in SIGMOD'19

Database Research Group, Columbia University
Research Assistant, advised by Prof. Eugene Wu Sep 2015 – Dec 2016

- **Data Visualization Management System**
  - Designed an aggressive prefetching system to improve the response time of networked data visualizations, by allowing the server to push anticipated data to the client
  - Built a predictive model for user mouse interations on web-browsers and implemented a JavaScript library for the model
  - Developed experiments using a Chrome extension to collect user interaction data

Visualization and Visual Analytics Group, Peking University
Research Assistant, advised by Prof. Xiaoru Yuan Apr 2014 – May 2015

- **Visual Analytics in Ancient Chinese Literature**
  - Researched topic on text visualization and collaborative sense making
  - Designed and implemented the platform and interface for viewing and annotating entities and events in the literature, and visualized the storylines, using D3.js and Python
- **Store Explorer: A visual analysis tool for heterogeneous text data**
  - Designed and implemented a flexible filter and a visual analytics interface for text dataset
  - Built up the Story Explorer system for analyzing social network and behavior log information

AWARDS & HONORS

- SIGMOD Travel Award Jun 2018
- 7th Place in ACM/ICPC 2015 Greater New York Regional Nov 2015
- Award for Excellent Detailed Analysis, VAST Challenge Mini-Challenge 1 Nov 2014
- Award for Excellent Comprehensive Visual Analysis System,

VAST Challenge Mini-Challenge 2 Nov 2014

- Award for Excellent Collaborative Streaming Analysis, VAST Challenge Mini-Challenge 3 Nov 2014
- The May Fourth Scholarship, Peking University May 2013
- Silver medal in ACM/ICPC 2012 Asia Regional Contest in Tianjin Oct 2012
- First Prize in National Olympiad in Informatics in Hunan Province Nov 2009

WORK EXPERIENCE

MUX Group, Yahoo Beijing R&D Center

Technical Intern Jul 2014 – Oct 2014

- Project: Mobile Catalogs
  - Optimized algorithms for extracting mobile device information and detecting user network environment
  - Implemented an algorithm for semi-automatically maintaining the index of mobile devices
  - Optimized the storage format of user information and reduced the storage space by 2/3

TEACHING EXPERIENCE

Teaching Assistant, Introduction to Database Systems (Duke CompSci 316) Spring 2019

- I assisted in writing and grading the assignments and projects.

Teaching Assistant, Everything Data (Duke CompSci 216) Spring 2018

- I assisted in writing and grading the assignments, labs, and projects.

SKILLS

- Programming Languages
  - Proficient in C, C++, JAVA, Python, JavaScript, R, LaTeX, HTML, SQL
  - Familar with Matlab, Scala, Scheme
- Frameworks
  - D3.js, Nodejs, Hadoop, SciKit-learn, Tensorflow