Zhongjun Jin (Mark)

CONTACT INFORMATION

4945 Bob and Betty Beyster Building
2260 Hayward Street
Ann Arbor, MI 48109, USA

Phone: (765) 421-5014
E-mail: markjin@umich.edu, markjin1990@gmail.com
Website: `https://markjin1990.github.io/`

OBJECTIVES

Applying for an industrial research scientist / applied scientist / research engineer / software engineer position starting in summer 2020.

RESEARCH INTERESTS

Build interactive systems for data scientists/analysts and programmers using a combination of AI, HCI and PL techniques.

EDUCATION

University of Michigan, Ann Arbor, MI, USA **Aug. 2014 - May 2020**

Ph.D. Candidate, Computer Science and Engineering

- Advisor: Prof. Michael Cafarella and Prof. H. V. Jagadish

Purdue University, West Lafayette, IN, USA **Aug. 2011 - May 2014**

B.S. in Computer Science, Mathematics

Tianjin University, Tianjin, China **Aug. 2009 - Jul. 2011**

Electronic Information Science

PROFESSIONAL EXPERIENCE

Microsoft Research, Redmond, WA **Feb 2019 - May 2019**

Research Intern (Mentored by Yeye He)
Designed and implemented the prototype system AUTOTRANSFORM in C♯ for data format standardization.

Trifacta, San Francisco, CA **May 2017 - Sep. 2017**

Software Engineering Intern (Mentored by Sean Kandel, Michael Minar, and Joe Hellerstein)
Designed and implemented the prototype system CLX for data format standardization in Python and Javascript. The work was integrated to Trifacta Cloud Wrangler as a main feature in Aug 2018.

Qualcomm, San Diego, CA **May 2013 - Aug. 2013**

Software Engineering Intern
Integrated functional tests into Qualcomm test automation system using Perl.

RESEARCH PROJECTS

- FOOFAH. A system synthesizing data transformation programs using user-provided positive examples.
- CLX. A system profiling string data patterns and suggesting transformation programs for data in non-standard patterns.
- AUTOTRANSFORM. A system recommending string data transformation programs learned offline from a large corpus of web data.
- PRISM. A system synthesizing SQL queries using imprecise tuple examples the end user provides.
- DEEPWRANGLER. A neural-guided program synthesis framework targeting at reducing the amount of examples end users need to provide in examples-driven data transformations.

Conference and Workshop Papers

1. Christopher Baik, **Zhongjun Jin**, Michael Cafarella, and H. V. Jagadish, "Constructing Expressive Relational Queries with Dual-Specification Synthesis", in *CIDR* 2020.
2. **Zhongjun Jin**, Michael Cafarella, H. V. Jagadish, Sean Kandel, Michael Minar, and Joseph M. Hellerstein, "CLX: Towards verifiable PBE data transformation", in *EDBT* 2019.
3. **Zhongjun Jin**, Christopher Baik, Michael Cafarella, H. V. Jagadish, and Yuze Lou, "Demonstration of a Schema Mapping System Using Multiresolution Constraints", in *CIDR* 2019.
4. Christopher Baik, **Zhongjun Jin**, and Michael Cafarella, "Disambiguating Queries in Conversational Interface", in *CAST @ VLDB* 2019.
5. Abolfazl Asudeh, **Zhongjun Jin**, and H. V. Jagadish, "Assessing and Remedying Coverage for a Given Dataset", in *ICDE* 2019.
6. **Zhongjun Jin**, Christopher Baik, Michael Cafarella, and H. V. Jagadish, "Beaver: Towards a Declarative Schema Mapping", in *HILDA @ SIGMOD* 2018.
7. **Zhongjun Jin**, Michael R Anderson, Michael Cafarella, and H. V. Jagadish, "Foofah: Data Transformation By Example", in *SIGMOD* 2017.
8. **Zhongjun Jin**, Michael R Anderson, Michael Cafarella, and H. V. Jagadish, "Foofah: A Programming-By-Example System for Synthesizing Data Transformation Programs", in *SIGMOD* 2017.
9. Rohit Ranchal, Bharat K. Bhargava, Ruchith Fernando, Hui Lei, and **Zhongjun Jin**, "Privacy Preserving Access Control in Service-Oriented Architecture", in *ICWS* 2016.
10. Pelin Angin, Bharat Bhargava, and **Zhongjun Jin**, "A Self-Cloning Agents Based Model for High-Performance Mobile-Cloud Computing", in *CLOUD* 2015.
11. **Zhongjun Jin**, Mengjing Xu, Chenkai Sun, Abolfazl Asudeh, and H. V. Jagadish, "Mithra-Coverage: A System for Investigating Population Bias for Intersectional Fairness", under review.
12. Christopher Baik, **Zhongjun Jin**, Michael Cafarella, and H. V. Jagadish, "A Dual-Specification System for Inferring SQL Queries", under review.
13. **Zhongjun Jin**, Christopher Baik, Michael Cafarella, H. V. Jagadish, and Yuze Lou, "Schema Mapping with Multiresolution Constraints", under review.

Honors and Awards

- 1st Prize in "Systems, Software Engineering and Computer Science" session in *Michigan Engineering Graduate Symposium 2017 (EGS 2017)*, 2017.
- Selected as "Best of Demos" at SIGMOD 2017.
- Sigmod Travel Award, 2017.
- University of Michigan Departmental PhD Fellowship, 2014.
- Outstanding Undergraduate Research Endeavor Award, Purdue Computer Science Dept, 2014
- Purdue Computer Science Neel Memorial Scholarship, 2013
- Purdue Computer Science Departmental Scholarship, 2012

Invited Talks

- "Intelligent Self-service Data Preparation: Problems and Solutions", 11/15/2018, Llamasoft Inc., USA.

Service

- External Reviewer: SoCC'19