

# MUHAMMAD ALI GULZAR

gulzar@cs.vt.edu • <https://people.cs.vt.edu/~gulzar>  
Computer Science • Virginia Tech • H-Index 8 • Citations 264

## RESEARCH STATEMENT

---

The vision of my research is to build systems that improve developer productivity through automated debugging and testing of big data analytics. Broadly, I am interested in designing novel tool support for data-centric software development. My approach combines insights from software engineering, distributed systems, and databases. I have also been awarded the Google Ph.D. Fellowship in the sub-field of software engineering.

## EDUCATION

---

|  |             |
|--|-------------|
| <b>Ph.D. in Computer Science</b><br>University of California, Los Angeles<br><i>Thesis: Interactive and Automated Debugging for Big Data Analytics</i> | 2014 – 2020 |
| <b>B.S in Computer Science</b><br>Lahore University of Management Sciences   | 2010 – 2014 |

## EMPLOYMENT

---

|   |                   |
|---|-------------------|
| <b>Assistant Professor in Computer Science</b><br>Virginia Polytechnic Institute and State University, Blacksburg, VA | AUG '20 – PRESENT |
| <b>Software Engineer, Intern</b><br>Google Inc., Mountain View  | JUN '19 – SEPT 19 |
| <b>Software Engineer Tools and Infrastructure Intern</b><br>Google Inc., Mountain View                                | JUN '18 – SEPT 18 |
| <b>Summer Research Assistant</b><br>NEC Labs America, Princeton NJ  | JUN '16 – SEP '16 |
| <b>Research Intern</b><br>Koç University, Turkey  | MAY '12 – AUG '12 |

## AWARDS AND GRANTS

---

2019-20 Outstanding Computer Science Graduating Ph.D. Student  
2019-20 Northrop Grumman Outstanding Computer Science Graduate Student Research Award  
Google Ph.D. Fellowship 2017-20 (with 3rd-year extension)  
NSF I-Corps Grant (Entrepreneurial Lead) 2018  
Gold medal at ACM Student Research Competition at ICSE 2018  
SIGMOD Student Travel Award 2017  
SoCC Student Travel Award 2017 and 2019  
Graduation with Distinction, Dean's Honor List Award, National Mathematics Olympiad Finalist

## RESEARCH TRACK PUBLICATIONS

---

|      |   |   |
|------|---|---|
| [P1] | <b>HeteroRefactor: Refactoring for Heterogeneous Computing with FPGA</b><br>Jason Lau*, Aishwarya Sivaraman*, Qian Zhang*, <b>Muhammad Ali Gulzar</b> ,<br>Jason Cong, Miryung Kim<br><i>The proceedings of the 42nd International Conference on Software Engineering. 12 Pages.</i><br>(* are equal co-first authors ordered alphabetically by their last names) | ICSE'20<br><br>Acceptance Rate: 20.9%     |
| [P2] | <b>White-Box Testing of Big Data Analytics with Complex User-Defined Functions</b><br><b>Muhammad Ali Gulzar</b> , Shaghayegh Mardani,<br>Madan Musuvathi, Miryung Kim<br><i>The 27th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering. 12 Pages.</i>  | ESEC/FSE'19<br><br>Acceptance Rate: 24.4% |
| [P3] | <b>PerfDebug: Performance Debugging of Computation Skew in Dataflow Systems</b><br>Jason Teoh, <b>Muhammad Ali Gulzar</b> , Harry Xu, Miryung Kim.<br><i>ACM Symposium on Cloud Computing. 12 Pages.</i>  | SoCC'19<br>Acceptance Rate: 24.8%         |
| [P4] | <b>LogLens: A Real-Time Log Analysis System</b>   | ICDCS'18                                  |

- Biplob Debnath, Mohiuddin Solaimani, **Muhammad Ali Gulzar**,  
Nipun Arora, Cristian Lumezanu, Jianwu Xu, Bo Zong, Hui Zhang,  
Guofei Jiang, Latifur Khan. *Acceptance Rate: 20.6%*  
*IEEE 38th International Conference on Distributed Computing Systems. 11 Pages.*
- [P5] **Adding data provenance support to Apache Spark** VLDB JOURNAL'18  
Matteo Interlandi, Ari Ekmekji, Kshitij Shah, **Muhammad Ali Gulzar**,  
Sai Tetali, Miryung Kim, Todd Millstein, Tyson Condie.  
*The VLDB Journal (Special Issue on "The Best Papers of" VLDB 2018). 21 Pages.*
- [P6] **Automated Debugging in Data-Intensive Scalable Computing** SoCC'17  
**Muhammad Ali Gulzar**, Matteo Interlandi, Xueyuan Han, Mingda Li,  
Tyson Condie, Miryung Kim. *Acceptance Rate: 23.6%*  
*ACM Symposium on Cloud Computing. 13 Pages.*
- [P7] **BigDebug: Debugging Primitives for Interactive Big Data Processing in Spark** ICSE'16  
**Muhammad Ali Gulzar**, Matteo Interlandi, Seunghyun Yoo, Sai Tetali,  
Tyson Condie, Todd Millstein, Miryung Kim. *Acceptance Rate: 19%*  
*The proceedings of the 38th International Conference on Software  
Engineering. 12 Pages.*
- [P8] **Optimizing Interactive Development of Data-Intensive Applications** SoCC'16  
Matteo Interlandi, Sai Tetali, **Muhammad Ali Gulzar**, Joseph Noor,  
Tyson Condie, Miryung Kim, Todd Millstein. *Acceptance Rate: 25.1%*  
*ACM Symposium on Cloud Computing. 13 Pages.*
- [P9] **Titian: Data Provenance Support in Spark** VLDB'16  
Matteo Interlandi, Kshitij Shah, Sai Tetali, **Muhammad Ali Gulzar**,  
Seunghyun Yoo, Miryung Kim, Todd Millstein, Tyson Condie. *Acceptance Rate: 21.2%*  
*Proc. VLDB Endow., 9(3):216-227, Nov. 2015. 12 Pages.*

#### DEMONSTRATION TRACK PUBLICATIONS

---

- [D1] **BigTest: Symbolic Execution Based Systematic Test Generation for Apache Spark** ICSE'20  
**Muhammad Ali Gulzar**, Madan Musuvathi, Miryung Kim. *Acceptance Rate: 33.3%*  
*The 42nd International Conference on Software Engineering.  
Research Demonstration Track. 4 Pages.*
- [D2] **BigSift: Automated Debugging of Big Data Analytics in Data-intensive  
Scalable Computing** FSE'18  
**Muhammad Ali Gulzar**, Siman Wang, Miryung Kim. *Acceptance Rate: 38.8%*  
*The 26th ACM Joint European Software Engineering Conference and Symposium  
on the Foundations of Software Engineering. Research Demonstration Track. 4 Pages.*
- [D3] **Debugging Big Data Analytics in Spark with BigDebug** SIGMOD'17  
**Muhammad Ali Gulzar**, Matteo Interlandi, Tyson Condie, Miryung Kim. *Acceptance Rate: 34%*  
*The Proceedings of The 2017 ACM SIGMOD/PODS Conference  
Demonstration Track. 4 Pages.*
- [D4] **BigDebug: Interactive Debugger for Big Data Analytics in Apache Spark** FSE'16  
**Muhammad Ali Gulzar**, Matteo Interlandi, Tyson Condie, Miryung Kim. *Acceptance Rate: 40%*  
*The proceedings of the 24th Joint European Software Engineering Conference  
and Symposium on the Foundations of Software Engineering.  
Research Demonstration Track. 5 Pages.*

#### SHORT/WORKSHOP/INDUSTRY TRACK PUBLICATIONS

---

- [SWI1] **Perception and Practices of Differential Testing** ICSE SEIP'19  
**Muhammad Ali Gulzar**, Yongkong Zhu, Xiaofeng Han. *Acceptance Rate: 22.2%*  
*The proceedings of the 41th International Conference on Software Engineering  
Software Engineering in Practice Track. 10 Pages.*
- [SWI2] **Interactive Debugging for Big Data Analytics** HOTCLOUD'16  
**Muhammad Ali Gulzar**, Xueyuan Han, Matteo Interlandi, Shaghayegh  
Mardani, Sai Tetali, Tyson Condie, Todd Millstein, Miryung Kim. *Acceptance Rate: 30.8%*  
*The 8th USENIX Workshop on Hot Topics in Cloud Computing. 7 Pages.*

## TEACHING EXPERIENCE

---

|  |                         |
|--|-------------------------|
| <b>CS130 Software Engineering</b><br>University of California, Los Angeles         | SPRING '17 AND FALL '15 |
| <b>CS282 Network Centric Computing</b><br>Lahore University of Management Sciences | SPRING '14              |
| <b>CS310 Algorithms</b><br>Lahore University of Management Sciences                | SPRING '13              |
| <b>CS340 Databases</b><br>Lahore University of Management Sciences                 | FALL '13                |

## MENTORING EXPERIENCE

---

|   |                    |
|---|--------------------|
| Xueyuan Han ( <i>Undergraduate at UCLA. Currently a Ph.D. student at Harvard University</i> ) | JAN '16 - MAY '17  |
| Simon Wang ( <i>Research Intern (CSST) at UCLA. Currently a Masters student at UCSD</i> )     | MAY '17 - SEP '17  |
| Shaghayegh Mardani ( <i>Ph.D. student at UCLA</i> )   | SEP '17 - JUN '17  |
| Aishwarya Sivaraman ( <i>Ph.D. student at UCLA</i> )  | SEP '17 - JUN '17  |
| Jason Teoh ( <i>Ph.D. student at UCLA</i> )   | JAN '18 - PRESENT  |
| Fabrice Harel-Canada ( <i>M.S. student at UCLA</i> )  | APR '19 - PRESENT  |
| Usama Hameed ( <i>Ph.D. student at UCLA</i> )   | SEPT '19 - PRESENT |

## SERVICE

---

**Journal Reviewer:** Transaction on Software Engineering (TSE 2018)  
**Student Volunteer:** Co-lead Student Volunteer ICSE 2016  
**Mentor:** Incoming M.S. Computer Science Class in 2015  
**Panelist:** UCLA Computer Science Ph.D. Open-house year 2017 and 2018  
**Research Host:** UCLA Cross-disciplinary Scholars in Science and Technology 2017

## INVITED TALKS

---

|   |                                  |
|---|----------------------------------|
| White-Box Testing of Big Data Analytics                                     | Google PhD Fellowship Summit '19 |
| Automated Debugging and Testing of Big Data Analytics                       | ICSE' 18                         |
| Automated Debugging in Data-Intensive Scalable Computing                    | FSE' 18 , SoCC' 17               |
| Debugging Big Data Analytics in Spark                                       | Spark Summit'17                  |
| Debugging Big Data Analytics with BigDebug                                  | SIGMOD'17                        |
| Interactive Debugger for Big Data Analytics                                 | FSE'16                           |
| BigDebug: Debugging Primitives for Interactive Big Data Processing in Spark | ICSE'16                          |
| Interactive Debugging for Big Data Analytics                                | NEC Labs America, 2016           |
| Towards Big Data Debugging in Apache Spark                                  | Databricks Inc., 2015            |

## ACADEMIC REFERENCES

---

- **Miryung Kim**  
Professor of Computer Science  
University of California, Los Angeles  
✉: miryung @ cs.ucla.edu  
☎: +1 (310) 825-2858
- **Harry Xu**  
Associate Professor of Computer Science  
University of California, Los Angeles  
✉: harryxu @ cs.ucla.edu  
☎: +1 (310) 794-7145
- **Todd Millstein**  
Professor of Computer Science  
University of California, Los Angeles  
✉: todd @ cs.ucla.edu  
☎: +1 (310) 825-5942
- **Madan Musuvathi**  
Partner Research Manager  
Microsoft Research  
✉: madanm @ microsoft.com  
☎: +1 (425) 706-5946