

Chenguang Zhu

Contact Information	EER, University of Texas at Austin 2501 Speedway, Austin, TX, 78712, USA	+1 (650) 229-9955 cgzhu@utexas.edu chenguang-zhu.github.io
Interests	My interests mainly focus on program analysis techniques with applications in automated software testing, software evolution, and the integration of software engineering and machine learning approaches. Recently, I am particularly interested in designing program analysis and transformation techniques to ensure the correctness of machine learning programs.	
Education	University of Texas at Austin , Austin, TX, USA Ph.D., Electrical and Computer Engineering Aug 2017 – Present <ul style="list-style-type: none">- Advisor: Sarfraz Khurshid- Area of Study: Software Engineering- GPA 3.8/4.0 University of Toronto , Toronto, ON, Canada M.Sc., Computer Science Sep 2015 – Apr 2017 <ul style="list-style-type: none">- Advisor: Marsha Chechik- Area of Study: Software Engineering- GPA 4.0/4.0 Harbin Institute of Technology , Harbin, China B.E., Software Engineering Aug 2011 – Jul 2015 <ul style="list-style-type: none">- Ranked 1st out of 123 students- GPA 93.2%	
Publications	<ol style="list-style-type: none">1. Chenguang Zhu, Ripon Saha, Mukul Prasad, and Sarfraz Khurshid. Restoring the Executability of Jupyter Notebooks by Automatic Upgrade of Deprecated APIs. In <i>Proceedings of the 36th IEEE/ACM International Conference on Automated Software Engineering (ASE)</i>, To appear, 2021.2. Xiuheng Wu, Chenguang Zhu, and Yi Li. Diffbase: A Differential Factbase for Effective Software Evolution Management. In <i>Proceedings of the 29th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (FSE)</i>, To appear, 2021.3. Chenguang Zhu, Yi Li, Julia Rubin, and Marsha Chechik. GenSlice: Generalized Semantic History Slicing. In <i>Proceedings of the 36th IEEE International Conference on Software Maintenance and Evolution (ICSME)</i>, pages 81–91, 2020.4. Xindong Zhang, Chenguang Zhu, Yi Li, Jianmei Guo, Lihua Liu, and Haobo Gu. Prefix: Large-Scale Patch Recommendation by Mining Defect-Patch Pairs. <i>Proceedings of the 42nd International Conference on Software Engineering (ICSE): Software Engineering in Practice</i>, pages 41–50, 2020.5. Yi Li, Chenguang Zhu, Milos Gligoric, Julia Rubin, and Marsha Chechik. Precise Semantic History Slicing Through Dynamic Delta Refinement. <i>Automated Software Engineering (ASEJ)</i>, 26(4), pages 757–793, 2019.6. Chenguang Zhu, Owolabi Legunsen, August Shi, and Milos Gligoric. A Framework for Checking Regression Test Selection Tools. In <i>Proceedings of the 41st International Conference on Software Engineering (ICSE)</i>, pages 430–441, 2019.	

7. Yi Li, **Chenguang Zhu**, Julia Rubin, and Marsha Chechik. CSlicerCloud: A Web-Based Semantic History Slicing Framework. In *Proceedings of the 40th International Conference on Software Engineering: Companion Proceedings (ICSE DEMO)*, pages 57–60, 2018.
8. Kaiyuan Wang, **Chenguang Zhu**, Ahmet Celik, Jongwook Kim, Don Batory, and Milos Gligoric. Towards Refactoring-Aware Regression Test Selection. In *Proceedings of the 40th International Conference on Software Engineering (ICSE)*, pages 233–244, 2018.
9. Yi Li, **Chenguang Zhu**, Julia Rubin, and Marsha Chechik. Semantic Slicing of Software Version Histories. *IEEE Transactions on Software Engineering (TSE)*, 44(2), pages 182–201, 2018.
10. Yi Li, **Chenguang Zhu**, Julia Rubin, and Marsha Chechik. FHistorian: Locating Features in Version Histories. In *Proceedings of the 21st International Systems and Software Product Line Conference (SPLC)*, pages 49–58, 2017.
11. **Chenguang Zhu**, Yi Li, Julia Rubin, and Marsha Chechik. A Dataset for Dynamic Discovery of Semantic Changes in Version Controlled Software Histories. In *Proceedings of the 14th International Conference on Mining Software Repositories (MSR)*, pages 523–526, 2017.
12. Yi Li, **Chenguang Zhu**, Julia Rubin, and Marsha Chechik. Precise Semantic History Slicing Through Dynamic Delta Refinement. In *Proceedings of the 31st IEEE/ACM International Conference on Automated Software Engineering (ASE)*, pages 495–506, 2016.

Professional
Experience

Research Assistant

Jan 2020 – Present

University of Texas at Austin
Advisor: Sarfraz Khurshid

Austin, TX, USA

Conducting research on the integration of software engineering and artificial intelligence. Built RELANCER, an automatic technique that restores the executability of broken data science programs by upgrading deprecated APIs.

Research Intern

May 2020 – Aug 2020

Fujitsu Research of America, Inc.
Advisors: Ripon Saha, Mukul Prasad

Sunnyvale, CA, USA

Developed a program analysis framework for Python.

Research Intern

Jun 2019 – Jan 2020

Nanyang Technological University
Advisors: Yi Li, Yang Liu

Singapore, Singapore

Conducted research on software evolution and patch recommendation. Built GenSlice, a generalized semantic history slicing framework which integrates several history slicing techniques under a uniform lens and provides a systematic approach for comparisons and analyses. Co-developed DIFFBASE, a framework for extracting, storing, and supporting efficient querying and manipulation of differential facts to support various program analysis tasks. Contributed to PREFIX, a framework that mines bug-fix patterns from large scale codebases, then utilizes these patterns to detect defects and recommend patches.

Research Assistant

Aug 2017 – May 2019

University of Texas at Austin
Advisor: Milos Gligoric

Austin, TX, USA

Conducted research on software evolution and testing. Built RTSHECK, a framework for automatically finding bugs in regression test selection tools. Contributed to REKS, a refactoring-aware regression test selection technique.

Research Assistant Sep 2015 – May 2017
 University of Toronto Toronto, ON, Canada
 Advisor: Marsha Chechik

Conducted research on software version history analysis and feature engineering. Co-developed DEFINER, a technique relying on dynamic analysis to extract code changes related to a high-level functionality. Built DO SC, a dataset for benchmarking software analysis techniques that dynamically discover semantic changes. Co-developed FHISTORIAN, a technique that establishes feature relationship graphs based on software history.

Research Intern Jul 2016 – Sep 2016
 Carnegie Mellon University Mountain View, CA, USA
 Advisors: Arie Gurfinkel, Temesghen Kahsai

Conducted research on synthesizing invariants for program verification. Developed a machine-learning based invariant generation technique in the SeaHorn verification framework.

Software Engineer Intern Nov 2014 – Feb 2015
 Tencent Shenzhen, China

Developed a message-exchanging module of the QQMusic iOS App.

Undergraduate Research Assistant May 2014 – Jun 2015
 Harbin Institute of Technology Harbin, China
 Advisors: Tiantian Wang, Peijun Ma

Conducted research on code clone detection. Co-developed a system that relies on static analysis to identify copy-pasted code in software systems.

Press
 Diffbase: A Differential Factbase for Effective Software Evolution Management.
 - NTU News: <https://bit.ly/2Zde01B>
 GenSlice: Generalized Semantic History Slicing.
 - NTU News: <https://bit.ly/3dgrrC1>
 Prefix: Large-Scale Patch Recommendation by Mining Defect-Patch Pairs.
 - Alibaba Cloud: <https://bit.ly/3gWXWHT>
 - Sohu News (Chinese): <https://bit.ly/3dhUcyh>
 - Zhihu (Chinese): <https://bit.ly/3xWpPFs>

Professional Service
 ICST 2022 Program Committee Member – Industry Track
 ASE 2021 External Reviewer
 FSE 2021 External Reviewer
 FASE 2021 External Reviewer
 ICST 2021 Program Committee Member – Industry Track
 ASE 2020 External Reviewer
 ISSRE 2020 External Reviewer
 ICST 2020 Program Committee Member – Industry Track
 ASE 2019 External Reviewer
 ISSTA 2019 External Reviewer
 ICST 2019 External Reviewer
 ICSE 2019 External Reviewer
 ASE 2018 External Reviewer

Invited Talks	<p>Handling Nondeterminism in Data Science Applications Sep 2021 Guest lecture for the course EE382V: Software Testing in the Era of Nondeterminism, The University of Texas at Austin, Austin, TX, USA</p> <p>A Framework for Checking Regression Test Selection Tools Nov 2020 The University of Texas at Austin, Austin, TX, USA</p> <p>GenSlice: Generalized Semantic History Slicing Sep 2020 IEEE International Conference on Software Maintenance and Evolution (ICSME), Adelaide, Australia</p> <p>Automated Analysis and Tooling for Supporting Software Evolution Jun 2018 Harbin Institute of Technology, Harbin, China</p>
Teaching Experience	<p>Teaching Assistant</p> <p>Department of Electrical and Computer Engineering, University of Texas at Austin</p> <ul style="list-style-type: none"> - EE 382V: Algorithmic Foundations for Software Systems Jun 2021 – Aug 2021 Held office hours, prepared exam materials, and graded student work. Topics include discrete math basics, theoretical foundations and implementation aspects of algorithms. - EE 360C: Algorithms Jan 2021 – May 2021 Held office hours, prepared assignment materials, and graded student work. Topics include proof-based reasoning about algorithms, asymptotic complexity analysis, and algorithm design. - EE 461L: Software Engineering and Design Lab Aug 2020 – Dec 2020 Taught weekly tutorials, held office hours, and graded student work. Topics include advanced software engineering tools, design patterns, and advanced Python programming. - EE 360T/EE382C-16: Software Testing Jan 2020 – May 2020 Held office hours and graded student work. Topics include software testing, symbolic execution, and model checking. <p>Department of Computer Science, University of Toronto</p> <ul style="list-style-type: none"> - CSC410: Software Testing and Verification Sep 2016 – Dec 2016 Taught weekly tutorials, held office hours, and graded student work. Topics include software testing, theorem proving, model checking, and static analysis. - CSC207: Software Design Jan 2016 – Apr 2016 Taught lab section, held office hours, and graded student work. Topics include Android application development, Java programming, the software design basics, and version control. - CSC108: Introduction to Programming Sep 2015 – Dec 2015 Graded student work and exams. Topics include Python programming and coding styles.
Awards	<p>Best Artifact Award Aug 2021 - The 29th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (FSE 2021)</p> <p>Best Artifact Award Sep 2020 - The 36th IEEE International Conference on Software Maintenance and Evolution (ICSME 2020)</p> <p>UT Austin Graduate School Fellowship 2017 – 2020 - Total value: USD \$36,000.</p> <p>National Scholarship 2013 – 2014 - Awarded by the Ministry of Education of China, to the student ranked first in the department. - Received the award in two consecutive academic years.</p> <p>Gold Medal in the 16th National Robot-soccer Championships Jul 2014 - Awarded by the Chinese Association for Artificial Intelligence, Robot Soccer Committee.</p> <p>First-class People’s Scholarship 2012 – 2014 - Awarded to the top 1% of the students in the department. - Received the award in three consecutive academic years.</p>

Technical Skills **Programming:** Java, Python, Shell, C, C++
Productivity: Emacs, \LaTeX , Git, SVN, IntelliJ, PyCharm
Operating Systems: Linux, Mac OS, Microsoft Windows