

# Chenguang Zhu

---

Contact Information	DCS, University of Toronto 3302-40 St. George Street Toronto, ON, Canada M5S 3G4	+1 (416) 876-9265 czhu@cs.toronto.edu www.cs.toronto.edu/~polaris
Research Interests	My interests mainly focus on program analysis and verification techniques with applications in but not limit to various aspects of software engineering research, including software version history analysis, bug finding, and feature location.	
Education	<b>University of Toronto</b> , Toronto, ON, Canada	
	M.Sc., Computer Science	Sep 2015 – Present
	- Supervisor: Marsha Chechik - Area of Study: Software Engineering - GPA 4.0/4.0	
	<b>Harbin Institute of Technology</b> , Harbin, China	
	B.E., Software Engineering	Aug 2011 – Jul 2015
	- Graduated with Highest Honours - Ranked 1st out of 123 students - Major GPA 93.2%	
Publications	<ol style="list-style-type: none"><li>1. <b>Chenguang Zhu</b>, Yi Li, Julia Rubin, and Marsha Chechik. A Dataset for Dynamic Discovery of Semantic Changes in Version Controlled Software Histories. <i>The 14th International Conference on Mining Software Repositories, (MSR 2017)</i>, To appear, 2017.</li><li>2. Yi Li, <b>Chenguang Zhu</b>, Julia Rubin, and Marsha Chechik. Semantic Slicing of Software Version Histories. <i>IEEE Transactions on Software Engineering</i>, To appear, 2017.</li><li>3. Yi Li, <b>Chenguang Zhu</b>, Julia Rubin, and Marsha Chechik. Precise Semantic History Slicing Through Dynamic Delta Refinement. In <i>Proceedings of the 31st IEEE/ACM International Conference on Automated Software Engineering (ASE 2016)</i>, pages 495–506, Singapore, 2016.</li><li>4. Kechao Wang, <b>Chenguang Zhu</b>, Tiantian Wang, and Xiaohong Su. Research on Code Clone Analysis Approach. <i>Application Research of Computers</i> (in Chinese), 33(3), 2016.</li></ol>	
Research Experience	<b>Research Assistant</b>	Sep 2015 – Present
	University of Toronto Supervisor: Marsha Chechik	
	Working in the Software Engineering Group. Conducted research on the following topics:	
	<ul style="list-style-type: none"><li>• Semantic Slicing of Software Version Histories</li><li>- Designed and built semantic history slicing techniques that precisely identifies a subset of change sets (commits) that implement a software functionality (e.g., feature, enhancement, or bug fix) from software version control systems.</li><li>- Provided automated support for software developers to transfer functionalities between branches of a software repository.</li><li>- Our approach was able to identify minimum-length fully-functional history slices, with an average reduction rate of 90.5%, for all 20 randomly selected functionalities from 6 actively developed open-source Java projects.</li></ul>	

- **FHISTORIAN: Locating Features in Version Histories**
  - Developed a dynamic feature location technique which leverages the information embedded in software version histories to discover relationships between features. Software product variants generated with respect to the feature relationships are fully functional and guaranteed to be executable and preserve desired behaviors.
- **A Dataset for Dynamic Discovery of Semantic Changes**
  - Created a benchmark for supporting multiple software analysis research, including change impact analysis, dynamic feature location, and semantic history slicing.
  - The dataset contains more than 180 data items extracted from 15 open-source projects managed in Git. Each data item represents a set of semantic changes related to a particular functionality.

### **Research Intern**

Jun 2016 – Sep 2016

Carnegie Mellon University & NASA Ames Research Center  
Supervisors: Arie Gurfinkel, Temesghen Kahsai

- **Machine Learning-based Invariant Generation in SeaHorn Verification Framework**
  - Worked three months in the SeaHorn team. Conducted research on synthesizing invariants for verification. Developed the ICE decision tree learning model, a machine-learning based invariant generation technique in SeaHorn, a fully automated LLVM-based verification framework.
  - Evaluated the algorithm on SV-COMP benchmarks. The results indicated that our technique can generate invariants effectively and efficiently.

### **Research Assistant**

May 2014 – Apr 2015

Harbin Institute of Technology  
Supervisors: Tiantian Wang, Peijun Ma

- **Automatic Code Clone Detection and Clone Pairs Harmfulness Ranking System**
  - Worked one year in the Space Software Engineering Research Center. Developed a token-based code clone detection tool that relies on static analysis to identify copy-pasted code in software systems.
  - Designed and implemented a clone pairs harmfulness ranking algorithm that ranks the harmfulness level of clone pairs based on their locations.
  - The evaluation of our techniques on a set of open-source projects indicated that it can recognize code clones and identified harmful clone pairs effectively.

### **Research Developer**

Mar 2014 – Jul 2014

Harbin Institute of Technology  
Supervisor: Bingrong Hong

- **Remotely Controlled Humanoid Robots for National Robot-soccer Championships of China**
  - Worked four months in Multi-Agent Robotics Research Center. Developed humanoid robots that can perform various actions in soccer games, such as shooting, tackling, blocking, etc, via infrared or Bluetooth remote control.
  - Developed embedded modules on robots that receives commands from an Android application on a cellphone through Bluetooth. Programmed 16 joints to closely imitate complex human motions, to control the mechanical movements of the robots.
  - This design won the gold medal in the 16th National Robot-soccer Championships of China.

Course Projects **Road Estimation for Auto-driving Based on Machine Learning Techniques** Nov 2015 – Dec 2015

University of Toronto

- Designed and implemented a set of methods to detect the road area ahead of a vehicle by identifying the road from photos.
- Preprocessed data by aggregating pixels to generate super-pixels. Applied multiple machine learning models including SVM, KNN and Random Forest. Improved the prediction results using Conditional Random Field.
- Evaluated our techniques on KITTI, a well-known vision benchmark suite. We obtained a best precision of 90.8% and recall of 93.47%.

**A Network Traffic Adjusting System Based on Software Defined Networks**

University of Toronto Oct 2015 – Dec 2015

- Developed a system that dynamically adjusts network traffic based on identified application types of packets. The system operates on software defined networks.
- Applied machine learning algorithms to identify traffic flows of different applications. The prediction pipeline can effectively recognize Skype, BitTorrent, SSL and HTTP flows.

Industrial Experience **Software Engineering Intern** Nov 2014 – Feb 2015

Tencent

- Worked three months in the QQMusic iOS team in Tencent, the fifth largest IT company in the world. Extended QQMusic, the second most popular music app on iOS platform in China.
- Developed an app usage information collection and reporting module by creating protocols for back-end message exchange. The feature is used by more than 10 million users.

**Software Engineering Intern** Jun 2013 – Aug 2013

Neusoft

- Worked three months on a DLNA multimedia player for embedded Linux on vehicles as a member of the Embedded Multimedia team in Neusoft, one of the largest IT solutions and services provider in China.
- Developed a module that enables digital media sharing among multimedia devices, by parsing the DLNA protocol and handling the interaction between media servers and clients.

Teaching Experience **Teaching Assistant**

Department of Computer Science, University of Toronto

- 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 process, and version control.
- CSC108: Introduction to Programming Sep 2015 – Dec 2015  
Graded student work and exams. Topics include Python programming, basic programming principles, and coding style.

- Awards
- University of Toronto Fellowship** 2015 – 2016
    - Total value: CAD \$27,470.
  - 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.
  - Gold Medal in the 16th National Robot-soccer Championships** Jul 2014
    - Awarded by the Chinese Association for Artificial Intelligence, Robot Soccer Committee.
  - 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.

Technical Skills

- Programming:** C, C++, Java, Python, C#, Objective-C
- Modeling:** Matlab, UML
- Productivity:**  $\LaTeX$ , Git, SVN, Eclipse, Visual Studio, Vi, Microsoft Office
- Operating Systems:** Linux, AIX, Mac OS, Microsoft Windows