

# Kaituo Li

Software Engineer

Amazon

email: kaituo@amazon.com

## Research Interests

---

Software Engineering

- Program analysis and testing (invariant inference, symbolic execution, automated test generation, testing and verification of distributed and concurrent systems).

Image Processing

- Digital watermarking and information embedding

## Education

---

- **University of Massachusetts, Amherst** Amherst, MA  
*Ph.D. in Computer Science* *Dec 2012 – Sep 2016*  
*Advisor: Yannis Smaragdakis*
- **University of Massachusetts, Amherst** Amherst, MA  
*Master of Science in Computer Science—GPA: 3.67* *Sep 2009 – Feb 2013*  
*Advisor: Yannis Smaragdakis*
- **Zhejiang University** Hangzhou, China  
*Master of Engineering in Software Engineering—GPA: 3.88* *Sep 2007 – Jul 2009*  
*Entrance exam waived based on outstanding undergraduate record*  
*Department rank: 1 (Department size: 139 graduates)*
- **Jilin University** Changchun, China  
*Bachelor of Engineering in Software Engineering—GPA: 3.83* *Sep 2003 – Jul 2007*  
*Department rank: 1 (Department size: 231 graduates)*

## Publications List

---

### ◆ Journal Papers

1. **Kaituo Li**, Christoph Reichenbach, Christoph Csallner, and Yannis Smaragdakis, “Residual Investigation: Predictive and Precise Bug Detection”, [ACM Transactions on Software Engineering and Methodologies](#), Vol. 24, No. 2, 2014.
2. **Kaituo Li**, Yu Hu and Jiachen Zhang, “Design and implementation of software firewall supporting dynamic evolution,” [Journal of Computer Applications](#), Vol. 28, No. S1, pp. 302-304, 2008 (in Chinese).

### ◆ Conferences Papers [with acceptance rate]

1. **Kaituo Li**, Pallavi Joshi, Aarti Gupta, and Malay K. Ganai. “ReproLite: A Lightweight Tool to Quickly Reproduce Hard System Bugs”, in: The ACM Symposium on Cloud

- Computing 2014 ([SoCC 2014](#)), 2014. [24%]
2. **Kaituo Li**, Christoph Reichenbach, Yannis Smaragdakis, and Michal Young, “**Second-Order Constraints in Dynamic Invariant Inference**”, in: The 2013 9th Joint Meeting on Foundations of Software Engineering ([ESEC/FSE 2013](#)), 2013. [20%]
  3. **Kaituo Li**, Christoph Reichenbach, Yannis Smaragdakis, Yanlei Diao, Christoph Csallner, “**SEDGE: Symbolic Example Data Generation for Dataflow Programs**”, in: The 28th IEEE/ACM International Conference on Automated Software Engineering ([ASE 2013](#)), 2013. [17%]
  4. **Kaituo Li**, Christoph Reichenbach, Christoph Csallner, and Yannis Smaragdakis, “**Residual Investigation: Predictive and Precise Bug Detection**”, in: The 2012 International Symposium on Software Testing and Analysis ([ISSTA 2012](#)), 2012. **ACM SIGSOFT Distinguished Paper Award.** [28.7%]

◆ **Refereed Workshop**

1. **Kaituo Li**, Dan Zhang and Dr Chen, “**Exposure time change attack on image watermarking systems**”, in: The 7<sup>th</sup> International Workshop on Digital Watermarking ([IWDW 2008](#)), pp. 170-183, 2008.
2. **Kaituo Li**, Dan Zhang, “**Attack on digital watermarking based on exposure**,” in: The 14th National Conference on Image and Graphics ([NCIG2008](#)), pp. 162-165, 2008 (in Chinese).

## Research Experience

---

Graduate Research Assistant Sep 2009 to Jan 2012  
 Yannis Smaragdakis University of Massachusetts Amherst

- Explored techniques for expressing, discovering, and employing statically known constraints for improving dynamically discovered invariants.
- Introduced the concept of “residual investigation” for program analysis and enhanced the static bug analyzer FindBugs with several residual investigations.
- Adapted conventional dynamic-symbolic execution in a conventional programming language to the unique features of the dataflow domain.

Graduate Research Assistant Sep 2007 to Jun 2009  
 Deren Chen & Dan Zhang Zhejiang University, China

- Created a generic digital watermarking attack method for robustness test of image watermarking algorithms.

Undergraduate Research Assistant Dec 2006 to Jun 2007  
 Jiachen Zhang Jilin University, China

- Created a software firewall that can be installed, started, stopped, updated, and uninstalled without bringing down the firewall by using netfilter and OSGi framework.

## Teaching Experience

---

Teaching Assistant Spring 2014  
CMPSCI 326 – Web Programming University of Massachusetts Amherst

- Acted as a venture capitalist. Responsible for grading written assignments, holding office hours, and answering questions in the online discussion forums.

Instructor/Teaching Assistant Spring 2016, Fall 2015, Spring 2015, Fall 2014, Fall 2013  
CMPSCI 230 – Computer Systems Principles University of Massachusetts Amherst

- Prepared Lectures, homework, laboratory and exams; lecturing to the students, leading discussions, setting up laboratory apparatus, grading exams and term projects

Teaching Assistant Spring 2013, Fall 2012  
CMPSCI 220 – Programming Methodology University of Massachusetts Amherst

- Led a weekly discussion section. Responsible for holding office hours, grading exams.

Teaching Assistant Spring 2012  
CMPSCI 105 – Computer Literacy University of Massachusetts Amherst

- Taught a semi-weekly lab session. Responsible for holding office hours, grading problem sets and exams.

## Industry Experience

---

Amazon, Inc., Seattle, WA Sep 2016 - present  
Software Engineer

- At AWS RedShift data warehousing team, my daily tasks span across the stack: distributed systems, networking, databases, operating systems, parallel computing, and architecture. I built tools in production for correctness verification and performance analysis of cloud database storage layer. My tools are responsible for extracting relevant information from raw disks, logs, and distributed file services that are scattered across hundreds of commodity machines. Then my tools aggregate and consolidate the extracted information, aiming to diagnose and detect performance and correctness issues. My tools are integrated as part of Redshift replace and backup workflow.

Lattice Engines, Inc., Boston, MA Summer 2014  
Engineering Intern

- Created a data mining system that can forecast the time required to complete a job in the Lattice SaaS platform, which can vary for a number of reasons: variance in input data set size, hardware resources, software modifications, network availability, network traffic, etc. This work includes statistical modeling, and building infrastructure to efficiently store, analyze and extract data.

NEC Laboratories America Inc, Princeton, NJ Summer 2013  
Research Intern

- Designed a domain-specific language that allows testers to specify all aspects of a potential scenario that causes a given bug in distributed systems.
- Developed a scheduler that takes control over the distributed system execution to enforce certain schedule of system events.
- Implemented a log extractor that automatically generate potential buggy scenario from a sequence of log messages that a tester believes indicates the cause of the bug.

OriginLab Corporation, Northampton, MA

Software Engineer in Test

Summer 2012

- Implemented and maintained automated tests in C# and Ranorex.
- Worked with development and test engineers to identify software defects.
- Assisted in the research and development of new features.

## Professional Service

---

- Reviewer for ICSE 2010, POPL 2013, IET Software
- Student Volunteer for OOPSLA 2010
- Steward for Graduate Employee Organization, UMass Computer Science Department, Spring 2009
- Graduate Representative, UMass Computer Science Department, 2013

## Computer Skills

---

Software Tools:	Matlab, Ant, Junit, Z3, Origin, Ranorex, scikit-learn, Pandas, JAMS (proficient); Mathematica, LINDO, LINGO, GNU Make, Maven (prior experience)
Programming:	Java, AspectJ (expert); C, C++, SQL, Python, C# (proficient); Bash, Perl, Ruby (prior experience)
Distributed Computing:	Cassandra, Pig, Hadoop, Apache Accumulo (proficient)
Networking:	XML (expert); JavaScript, HTML, CSS, ASP, JSP, Macromedia Dreamweaver (prior experience)
Database:	Microsoft SQL Server, Oracle, MySQL, ODBC, JDBC, DB2, postgresSQL (prior experience)
IDEs:	Eclipse, Visual Studio, Vi (proficient)
Platform:	Windows, Linux/Unix (proficient)

## Software

---

**SEDGE:** <https://github.com/kaituo/sedge/>

Designed and implemented SEDGE, a dynamic symbolic execution system that generates test cases to exercise the key behavior of the operators that comprise a Pig Latin program.

**usemetainv4daikon:** <http://code.google.com/p/usemetainv4daikon/>

Extended Daikon with an annotation mechanism for "meta-invariants" (a.k.a. second-order

constraints). Users can pass a separate configuration file with meta-invariants to Daikon. For instance, they can declare that the precondition of method foo is stronger than that of method bar for unrelated methods foo and bar (even though they don't know what the preconditions are). This changes the Daikon processing of the low-level observations and allows eliminating spurious invariants and adding correct invariants without having to re-run any test suites.

**getmetainv:** <http://code.google.com/p/getmetainv/>

Designed and implemented the tool getmetainv that can automatically generate "meta-invariants" (a.k.a. second-order constraints) that are constraints about invariants. For instance, our tool can find the precondition of method foo is stronger than that of method bar with a high degree of precision.

**Dsc:** <http://ranger.uta.edu/~csallner/dsc/index.html>

Contributed to the design and implementation of the Dsc dynamic symbolic execution engine. Designed and implemented support for the dynamic symbolic execution of instance methods. In particular, implemented support for the reasoning of constraints on the fields of the receiver object, which enables Dsc to analyze parameter-less instance methods. In addition, implemented support for inferring the variable names from the Java byte code. Introduced a smart expression simplification system using Z3 constraint solver and Java. Improved the implementation of the data structure representing the model (test case) by Z3 constraint solver and the data structure representing the symbolic heap inside the Java virtual machine.

## Honors and Awards

---

- ACM SIGOPS Travel Grants Oct 2014
- Travel Grant, School of Computer Science, Sep 2013, Sep 2014  
University of Massachusetts Amherst
- ACM SIGSOFT Distinguished Paper Award Jul 2012
- ACM SIGSOFT CAPS Travel Award Jun 2012, Sep 2013
- He Zhijun Scholarship, China Jun 2009
- Graduate with Distinction, Zhejiang Province, China May 2009
- Graduate with Distinction, Zhejiang University, China May 2009
- Tencent Scholarship, the Highest Band, China Oct 2008
- First Class Honor Award, Zhejiang University, China Oct 2008
- Insigma Hengtian Scholarship, China Dec 2007
- Outstanding Academic Research Ability Scholarship, Jun 2007  
Jilin University, China
- First Class Stipend, College of Software Technology, May 2007, Sep 2008  
Zhejiang University, China
- Graduate with Distinction, Jilin University, China Nov 2006
- Excellent League Member, Jilin University, China Sep 2005
- First Class Scholarship, Jilin University, China Sep 2004, Sep 2005, Sep 2006,

- Excellent Student, Jilin University, China Jun 2007  
Sep 2004, Sep 2005, Sep 2006,  
Jun 2007
- Outstanding English Ability Scholarship, Jilin University, China Sep 2004

## Professional Membership

---

Java Community Process (JCP)

## Relevant Coursework

---

Discrete Math, Artificial Intelligence, Machine Learning, Theory of Computation, Database Design and Implementation, Advanced Software Engineering, Modern Computer Architecture, Object Oriented Languages, Software Quality Assurance, Case Study of Project Management, System Analysis & Design, Software Requirement Engineering, Webservice Technology, Project Management, Algorithm Analysis and Design, Assembly Language, Computer Networks, Operating System, Fundamentals of Compiler and Implementation, Design Patterns, Computer Security, Probability & Statistics, Unix Tools and Scripting

## Reference

---

Prof. Yannis Smaragdakis  
 email: [yannis@smaragd.org](mailto:yannis@smaragd.org)  
 Dept. of Informatics  
 University of Athens  
 Ilisia , 15784  
 Greece

Dr. Christoph Reichenbach  
 email: [creichen@gmail.com](mailto:creichen@gmail.com)  
 R<sup>2</sup> Software & Systeme UG  
 Schwalbenweg 19  
 63500 Seligenstadt  
 Germany

Prof. Christoph Csallner  
 email: [csallner@uta.edu](mailto:csallner@uta.edu)  
 Computer Science and Engineering  
 Department  
 University of Texas at Arlington  
 Box 19015  
 Arlington, TX 76019-0015

Prof. Michal Young  
 email: [michal@cs.uoregon.edu](mailto:michal@cs.uoregon.edu)  
 1202 Dept. of Computer & Information  
 Science  
 Deschutes Hall  
 University of Oregon  
 Eugene, OR 97403-1202

Dr. Pallavi Joshi  
 email: [pallavijoshi84@gmail.com](mailto:pallavijoshi84@gmail.com)  
 Google  
 1600 Amphitheatre Parkway  
 Mountain View, CA 94043

Prof. Aarti Gupta  
 email: [aartig@cs.princeton.edu](mailto:aartig@cs.princeton.edu)  
 Department of Computer Science  
 Princeton University  
 35 Olden Street  
 Princeton, NJ 08540

