

## Baishakhi Ray

---

Postdoctoral Researcher  
Department of Computer Science  
University of California Davis  
Davis, CA - 95616

Voice: (303) 748-2958  
Email: bairay@ucdavis.edu  
<http://rayb.info>

### Research Interest

I am primarily interested in *Software Engineering* with a focus on *empirical studies*, *program analysis*, and *software evolution*. Using techniques from diverse domains including machine learning, natural language processing, and rigorous statistical modeling, I analyze large scale software repositories to understand on-going software engineering practices. This data-driven knowledge helps me to build novel program analysis techniques and development tools to improve software quality and programmer productivity.

### Education

**Ph. D., Electrical & Computer Engineering, University of Texas, Austin.**

August 2013 (GPA: 3.97)

Thesis: *Analysis of Cross-System Porting and Porting Errors in Software Projects*

Advisor: Miryung Kim

**MS, Computer Science, University of Colorado, Boulder.**

August 2009 (GPA: 4.0)

Thesis: *SecureWear: Securing Wearable Mobile Social Networks*

Advisor: Richard Han

**B.Tech., Computer Science, Calcutta University, India.**

**B.Sc., Physics, Presidency College, Kolkata, India.**

July 2004 (GPA: 3.84, Top 1% in University)

### Honors

- Best Practical Paper Award, IEEE Symposium on Security and Privacy (S&P Oakland), 2014
- Nominated for Distinguished Paper Award, IEEE/ACM 28th International Conference on Automated Software Engineering (ASE), 2013.
- Google Summer of Code, 2012
- SIGSOFT FSE 2012 CAPS travel award
- Ranked 6<sup>th</sup> out of 15,000 students in B.Sc. Physics (Hons.) examination.
- Selected in Indian Institute of Technology (IIT, Kharagpur) for graduate study.
- Jawaharlal Nehru Summer Scholarship for Advanced Research, India 2001. (100 students are selected nationwide).

### Research & Work Experience

**University of California Davis, CA, USA**

Postdoctoral Research Fellow

Advisor: Prem Devanbu

October 2013 - present

**Microsoft Research, Redmond, USA**

Research Intern

Mentors: Christian Bird, Nachiappan Nagappan, Thomas Zimmermann

May 2013–August 2013

**Google Summer of Code, Google Inc.**

Research Intern

May 2012 – August 2012

Mentors: Suzette Person, Neha Rungta, NASA

**The University of Texas at Austin, TX, USA**

Graduate Research Assistant

January 2011–May 2013

Adviser: Miryung Kim

**Avaya Research Lab, Westminister, CO, USA**

Research Intern

May 2008 – Aug 2008

**Ericsson Pvt. Ltd. Boulder, CO, USA**

Software Engineer

February 2009–June 2010

**Ixia, Sasken, and Texas Instruments, India**

Software Engineer

August 2004–July 2007

**Publications**

1. *A Large Scale Study of Programming Languages and Code Quality in Github*. **B. Ray**, D. Posnett, V. Filkov, P. T. Devanbu. In ACM SIGSOFT, 22<sup>nd</sup> International Symposium on the Foundations of Software Engineering (FSE'14), pages: 155-165, acceptance rate: 22%.
2. *Using Frankencerts for Automated Adversarial Testing of Certificate Validation in SSL/TLS Implementations*. C. Brubaker, S. Jana, **B. Ray**, S. Khurshid, and V. Shmatikov. In 35<sup>th</sup> IEEE Symposium on Security and Privacy, 2014 (S&P Oakland'14), pages: 114-129, acceptance rate: 13%, **Best Practical Paper Award**.
3. *Detecting and Characterizing Semantic Inconsistencies in Ported Code*. **B. Ray**, M. Kim, S. Person, N. Rungta. In 28<sup>th</sup> IEEE/ACM International Conference on Automated Software Engineering, 2013 (ASE'13), pages: 367-377, acceptance rate: 23%, **Nominated for Distinguished Paper Award, invited for ASE journal special issue**.
4. *An Empirical Study of API Stability and Adoption in the Android Ecosystem*. T. McDonnell, **B. Ray**, M. Kim. In 29<sup>th</sup> IEEE International Conference on Software Maintenance, 2013 (ICSM'13), pages: 70-79, acceptance rate: 22%.
5. *A Case Study of Cross-System Porting in Forked Projects*. **B. Ray**, M. Kim. In ACM SIGSOFT, the 20<sup>th</sup> International Symposium on the Foundations of Software Engineering (FSE'12), pages: 53:1-53:11, acceptance rate: 17%.
6. *Repertoire: A Cross-System Porting Analysis Tool for Forked Software Projects*. **B. Ray**, C. Wiley, M. Kim. In ACM SIGSOFT the 20<sup>th</sup> International Symposium on the Foundations of Software Engineering, Formal Research Tool Demonstration (FSE'12), pages: 8:1-8:4.
7. *An Empirical Study of Supplementary Bug Fixes*. J. Park, M. Kim, **B. Ray**, D. Bae. In The 9<sup>th</sup> IEEE Working Conference on Mining Software Repositories (MSR'12), pages: 40-49, acceptance rate: 28% **Invited to the Special Issue of Journal of Empirical Software Engineering (EMSE)**.
8. *PTask: Operating System Abstractions To Manage GPUs as Compute Devices*. C.J. Rossbach, J. Currey, M. Silberstein, **B. Ray**, E. Witchel. In Proceedings of the 23<sup>rd</sup> ACM Symposium on Operating System Principles (SOSP'11), pages: 233-248, acceptance rate: 17%.
9. *Touch Me wE@r: Getting Physical with Social Networks*, A. Beach, **B. Ray**, L. Buechley. In 2009 Workshop on Sensor-based Models and Feedback Systems for Social Computing. Associated with SocialCom 2009, pages: 960-965.
10. *A Protocol for Building Secure and Reliable Covert Channel*. **B. Ray** and S. Mishra. In 6<sup>th</sup> Annual Conference on Privacy, Security and Trust, 2008. (PST'08), pages: 246-253
11. *WhozThat?: Evolving an Ecosystem for Context-Aware Mobile Social Networks*. A. Beach, **B. Ray**, et al., In IEEE Network Magazine Special Issue on Composable context aware services, 2008, pages: 50-55.

## Book Chapter

12. *SecureWear: A Framework for Securing Mobile Social Networks*. **B. Ray**, R. Han. Advances in Computer Science and Information Technology. Computer Science and Engineering, Vol. 85, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Springer Berlin Heidelberg, 2012, pages: 515-524.

## Under Submission

13. *Gender and Tenure Diversity in GitHub Teams*. B. Vasilescu, D. Posnett, **B. Ray**, M. Brand, A. Serebrenik, P. Devanbu, V. Filkov, In Proceedings of the International Conference on Human Factors in Computing Systems (CHI'15), 10 pages, acceptance rate: 23%, (accepted for publication).
14. *The Uniqueness of Changes: Characteristics and Applications*. **B. Ray**, M. Nagappan, C. Bird, N. Nagappan, T. Zimmermann. Microsoft Research Technical Report, 2014, 10 pages.
15. *Assert Use in GitHub Projects*. C. Casalnuovo, P. Devanbu, A. Oliveira, V. Filkov, **B. Ray**, 10 pages.
16. *Detecting and Characterizing Semantic Porting Inconsistencies in Copy and Paste Code*. **B. Ray**, M. Kim, S. Person, N. Rungta., 27 pages.

## Software Releases

- Frankencerts, <https://github.com/sumanj/frankencert>.  
Framework to test certificate validation code in SSL/TLS implementations.  
*Has received 51 github 'stars' as of October 2014.*
- Repertoire, <https://github.com/baishakhir/RepertoireTool>.  
A cross-system porting analysis tool for forked software projects.  
*Published in FSE'12 tool demo track.*

## Teaching Experience

### Mentoring

- Casey Casalnuovo, Graduate Student, University of California, Davis. Co-advised to conduct an empirical study on the use of assertions in source code [14].
- Vincent Hellendoorn, Visiting Graduate Student, University of California, Davis. Helped in multiple studies related to software quality and source code structure.
- Saheel Godhane, Graduate Student, University of California, Davis. Mentored to implement control and dataflow analysis using LLVM/Clang infrastructure.
- Connie Nguyen, Undergraduate Student, University of California, Davis. Taught how to collect, process, and clean-up data to conduct multiple empirical studies.
- Abilio Oliveira, Undergraduate Student, University of California, Davis. Co-advised to study assertions usage in source code [14].
- Lisa Hua, Graduate Student, The University of Texas at Austin. Mentored Masters Thesis on *A case study of cross-branch porting in Linux Kernel*.
- Tyler McDonnell, Undergraduate Student, The University of Texas at Austin. Co-advised undergraduate thesis to conduct an empirical study on API Stability and Adoption in the Android Ecosystem [4].
- Taariq Chasmawala, High School Student, Austin, Texas. Mentored science fair projects on how to detect surrounding movements using camera and GPU (won awards from Intel and the U.S. Army).

### Guest Lecturer

- Introduction to Programming and Problem Solving, Fall 2014, University of California, Davis
- Software Engineering, Fall 2013, University of California, Davis
- Software Engineering and Design Laboratory, Fall 2013, The University of Texas at Austin

### **Teaching Assistant**

- Mobile Computing, Fall 2010, The University of Texas at Austin

### **Grader**

- Operating Systems, Spring 2008, University of Colorado Boulder.
- Networking, Fall 2008, University of Colorado Boulder.

## **Academic Services**

### **Program Committee**

- Mining Software Repositories, 2015 (MSR)
- Mining Software Repositories mining challenge, 2015 (MSR challenge)
- India Software Engineering Conference, 2015 (ISEC)
- Foundation of Software Engineering Artifact, 2014 (FSE Artifact)
- OOPSLA Artifact, 2013

### **Journal Reviewer**

- Transactions on Software Engineering (TSE)
- CSI Journal Computer Standards & Interfaces
- Information and Software Technology

### **External Reviewer**

- Automated Software Engineering, 2014 (ASE)
- Mining Software Repository, 2012 (MSR)
- OOPSLA 2012

### **Other**

- Vice President, Graduate Women in Engineering, ECE Department, The University of Texas at Austin

## **Talks**

### **A Case Study of Cross-Systems Porting in Forked Projects**

- Mining Summer School MSR Vision 2020, August 2012, Kingston, ON, Canada.
- 20th International Symposium on the Foundation of Software Engineering (FSE), November 2012, Cary, North Carolina, USA.
- IBM Research, March 2013, Delhi, India.

### **Detecting and Characterizing Semantic Inconsistencies in Ported Code**

- NASA Ames Research Center, Mountain View, CA, November 2013.
- 28th IEEE/ACM International Conference on Automated Software Engineering (ASE), Silicon Valley, CA, November 2013.

### **Analysis of Cross-System Porting and Porting Errors in Software Projects**

- Fujitsu Laboratories America, Sunnyvale, CA, January 2014.

## **Media Coverage**

- Language Study [1]: [SlashDot](#), [The Register](#), [Reddit](#), [InfoWorld](#), [Hacker News](#).
- Frankencerts [2]: [Reddit](#), [Golem](#), [Heise](#).

## References

### **Premkumar T. Devanbu**

Professor  
Department of Computer Science  
University of California, Davis  
Kemper Hall, 1 Shields Avenue  
Davis, CA 95616, USA  
devanbu@cs.ucdavis.edu

### **Sarfraz Khurshid**

Associate Professor  
Electrical and Computer Engineering Dept.  
The University of Texas at Austin  
1 University Station C5000  
Austin, TX 78712, USA  
khurshid@ece.utexas.edu

### **Thomas Zimmermann**

Senior Researcher  
Microsoft Research  
1 Microsoft Way  
Redmond, WA 98052, USA  
tzimmer@microsoft.com

### **Miryung Kim**

Associate Professor  
Department of Computer Science  
University of California, Los Angeles  
Boelter Hall, 420 Westwood Plaza  
Los Angeles, CA 90095, USA  
miryung@cs.ucla.edu

### **Vitaly Shmatikov**

Associate Professor  
Department of Computer Science  
The University of Texas at Austin  
2317 Speedway  
Austin, TX 78712, USA  
shmat@cs.utexas.edu