

# Xiao (Lester) Yu

☎ +1 (919) 518 4431 • ✉ xyu10@ncsu.edu • 🌐 lesteryu.com

## Education

---

### Doctor of Philosophy in Computer Science

*North Carolina State University*

Advisor: Guoliang Jin

**May 2018 (Expected)**

*Raleigh, NC, USA*

### Master of Science in Computer Software & Theory

*East China Normal University*

Advisor: Geguang Pu

**June 2011**

*Shanghai, China*

### Bachelor of Engineering in Software Engineering

*East China Normal University*

**June 2008**

*Shanghai, China*

## Research Interests

---

Understanding and addressing reliability problems in complex software systems through data analytics and program analysis, with a focus on log/trace-related techniques.

## Research Experience

---

### Understanding and Addressing Inefficiencies in Web-based Applications

*North Carolina State University*

*2015–Present*

- Realizing the prevalence of web-based applications, investigated their performance inefficiencies related to two common aspects: the request-based execution model and frequent database accesses.
- On inefficiencies related to the request-based execution model,
  - Proposed a data-driven analysis approach combining program analysis and tracing to infer data dependencies between request-handler methods.
  - Investigated how the inferred data dependencies could enable inter-request analysis, which works across request-handler methods to identify potential inefficiencies, such as repetitive computations.
- On inefficiencies related to database accesses,
  - Conducted an empirical study on performance bugs related to database accesses with the goal of finding future research directions to address inefficiencies in database accesses.
  - Identified root causes and related factors on both the application side and the database side, summarized and explained how such bugs could be introduced, triggered, and fixed.

### Log-based Monitoring for Cloud Infrastructures

*NEC Laboratories America and North Carolina State University*

*2014–2015*

- Proposed a lightweight workflow monitoring approach on distributed and interleaved logs from complex task executions in order to address the management complexity in multi-tenant cloud infrastructures.
- Implemented and applied the approach on OpenStack to monitor and detect failures and performance problems in task executions.

### Analyzing Execution Traces for Bottlenecks in Windows Kernel Drivers

*Microsoft Research and North Carolina State University*

*2012–2013*

- o From large-scale real-world execution traces, identified that system performance could be compromised by *cost propagation*, an adverse effect caused by synchronizations and component dependencies in kernel drivers and the Windows kernel.
- o Proposed and developed a practical two-step approach with effective data and pattern abstractions to (1) measure performance impacts manifested through cost propagation, and (2) discover runtime behavioral patterns closely related to performance problems.

### **Dynamic Test Generation for C Programs**

*East China Normal University*

2007–2011

- o Built an automated test generation tool for C programs with the goal of improving the efficiency and usability of dynamic symbolic execution.
- o Designed practical algorithms as supplements to the linear constraint solver used in the test generation tool to improve the handling of complex data structures and pointers.
- o Adopted the side-effect analysis that eliminates irrelevant path conditions, and a partial execution technique that reduces paths to explore in loop iterations, in order to relieve the path explosion problem.
- o Designed parallel algorithms for dynamic symbolic execution based on partial orders in program paths.

### **Validation of Design Patterns in Object-Oriented Programs**

*East China Normal University*

2009–2010

- o Designed a relational calculus and an object model to describe object-oriented design patterns.
- o Implemented a tool to detect and validate the use of design patterns in Java program with the relational calculus and the object model.

### **Analysis and Verification of rCOS**

*East China Normal University*

2007–2008

- o Extended rCOS, a component-based modeling system, to support assertions, invariants, and parallelism.
- o Proposed an rCOS-to-SPIN transformation approach to enable model verification.
- o Constructed operational semantics for the extended rCOS as a guide for the implementation of rCOS.

## **Publications**

---

- [1] **Xiao Yu** and Guoliang Jin. Dataflow Tunneling: Mining Inter-request Data Dependencies for Request-based Applications. In *Proceedings of the 40th International Conference on Software Engineering, ICSE '18*, 2018. To appear.
- [2] **Xiao Yu**, Wei Yang, Mengqi Gu, Guoliang Jin, Tao Xie, and Xintao Wu. Characterizing Performance Bugs in Database Applications. Submission Under Review, 2017.
- [3] **Xiao Yu**, Pallavi Joshi, Jianwu Xu, Guoliang Jin, Hui Zhang, and Guofei Jiang. CloudSeer: Workflow Monitoring of Cloud Infrastructures via Interleaved Logs. In *Proceedings of the 21st International Conference on Architectural Support for Programming Languages and Operating Systems, ASPLOS '16*, pages 489–502, New York, NY, USA, 2016. ACM.
- [4] **Xiao Yu**, Shi Han, Dongmei Zhang, and Tao Xie. Comprehending Performance from Real-World Execution Traces: A Device-Driver Case. In *Proceedings of the 19th International Conference on Architectural Support for Programming Languages and Operating Systems, ASPLOS '14*, pages 193–206, New York, NY, USA, 2014. ACM.
- [5] **Xiao Yu**, Shuai Sun, Geguang Pu, Siyuan Jiang, and Zheng Wang. A Parallel Approach to

Concolic Testing with Low-cost Synchronization. *Electronic Notes in Theoretical Computer Science*, 274:83 – 96, 2011.

- [6] Kang Miao, Siyuan Jiang, **Xiao Yu**, and Ji Zhao. Run-time Discovery of Java Design Patterns. In *the 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce*, AIMSEC '11, pages 3329–3332, 2011.
- [7] Libo Feng, **Xiao Yu**, Geguang Pu, Siyuan Jiang, Huibiao Zhu, and Bing Gu. Property Checking for Design Patterns. In *Proceedings of the IASTED International Conference on Software Engineering*, SE '10, pages 87–94. ACTA Press, 2010.
- [8] Zheng Wang, **Xiao Yu**, Tao Sun, Geguang Pu, Zuohua Ding, and JueLiang Hu. Test Data Generation for Derived Types in C Program. In *the Third IEEE International Symposium on Theoretical Aspects of Software Engineering*, TASE '09, pages 155–162, 2009.
- [9] Tao Sun, Zheng Wang, Geguang Pu, **Xiao Yu**, Zongyan Qiu, and Bing Gu. Towards Scalable Compositional Test Generation. In *the Ninth International Conference on Quality Software*, QSIC '09, pages 353–358, 2009.
- [10] Zheng Wang, **Xiao Yu**, Geguang Pu, Libo Feng, Huibiao Zhu, and Jifeng He. Execution Semantics for rCOS. In *Proceedings of the 15th Asia-Pacific Software Engineering Conference*, APSEC '08, pages 119–126, 2008.
- [11] **Xiao Yu**, Zheng Wang, Geguang Pu, Dingding Mao, and Jing Liu. The Verification of rCOS Using Spin. *Electronic Notes in Theoretical Computer Science*, 207:49 – 67, 2008.

## Patents

---

### **CloudSeer: Using Logs to Detect Errors in the Cloud Infrastructure**

*Pallavi Joshi, Hui Zhang, Jianwu Xu, Xiao Yu, and Guofei Jiang*

*US9720753, 8/1/2017*

## Internship and Visiting Experience

---

### **NEC Laboratories America**

*Research Intern, Autonomic Management Department*

*Mentor: Pallavi Joshi*

*Princeton, NJ, USA*

*5/2014–12/2014*

### **University of Illinois at Urbana-Champaign**

*Visiting Student, Computer Science*

*Advisor: Tao Xie*

*Urbana, IL, USA*

*9/2013–5/2014*

### **Food and Drug Administration**

*Visiting Student, Center for Devices and Radiological Health*

*Mentor: Yi Zhang*

*Silver Spring, MD, USA*

*5/2013–8/2013*

### **Microsoft Research**

*Research Intern, Software Analytics Group*

*Mentor: Shi Han*

*Beijing, China*

*5/2012–8/2012*

## Teaching Experience

---

### Teaching Assistant

North Carolina State University

- **CSC 501: Operating Systems Principles** Spring 2017, Fall 2017  
Lecturers: Xiaohui (Helen) Gu and Guoliang Jin
- **CSC 568: Enterprise Storage Architecture** Spring 2015  
Lecturer: Vincent W. Freeh
- **CSC 506: Architecture of Parallel Computers** Spring 2012, Spring 2015  
Lecturer: Edward F. Gehringer
- **CSC 510: Software Engineering** Fall 2011  
Lecturer: Annie I. Antón

## Conference Presentations

---

### CloudSeer: Workflow Monitoring of Cloud Infrastructures via Interleaved Logs

ASPLOS 2016, Atlanta, GA, USA

### Comprehending Performance from Real-World Execution Traces: A Device-Driver Case

ASPLOS 2014, Salt Lake City, UT, USA

## Professional Activities

---

**Program Committee Member:** ASPLOS 2018 (Shadow)

**Reviewer:** Empirical Software Engineering (Special Issue on Automatic Software Repair)

**Subreviewer:** ISSTA 2012 & 2013, ICSM 2012, OOPSLA 2013, MSR 2014, ICST 2014, ASE 2014

**Student Volunteer:** FSE 2012

## References

---

### Guoliang Jin

Assistant Professor

Department of Computer Science

North Carolina State University

✉ [guoliang\\_jin@ncsu.edu](mailto:guoliang_jin@ncsu.edu)

☎ (1) 919-515-7880

### Xiaohui (Helen) Gu

Associate Professor

Department of Computer Science

North Carolina State University

✉ [xgu@ncsu.edu](mailto:xgu@ncsu.edu)

☎ (1) 919-515-7045

### Hui Zhang

*Former:* Department Head of Systems Research

NEC Labs America

*Current:* Senior Staff Engineer

Ant Financial Network Technology Inc.

✉ [hui.zhang.princeton@gmail.com](mailto:hui.zhang.princeton@gmail.com)

☎ (86) 153-8402-5406

### Shi Han

Lead Researcher

Software Analytics & Data Intelligence Group

Microsoft Research

✉ [shihan@microsoft.com](mailto:shihan@microsoft.com)