

## Peipei Wang

---

CONTACT INFORMATION	3228 EBII, 890 Oval Drive Raleigh, NC 27695	Phone E-mail	(919)-592-2485 pwang7@ncsu.edu
---------------------	--	-----------------	-----------------------------------

---

RESEARCH INTERESTS	<ul style="list-style-type: none"><li>• Regular expression analysis and testing</li><li>• Software testing, program analysis and program repair</li><li>• Cloud System failure detection and diagnosis</li></ul>
--------------------	--

---

EDUCATION	<b>North Carolina State University</b> Ph.D. student in Computer Science	Raleigh, NC May 2013-present
	<b>Xi'an Jiaotong University</b> Master in Computer Software and Theory	Xi'an, China Sep 2010-May 2013
	Bachelor in Software Engineering	Sep 2006-July 2010

---

PUBLICATION	<p><b>Peipei Wang</b>, Rui Bai, Kathryn T. Stolee, "Exploring Regular Expression Evolution", IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), 2019. (to appear)</p> <p><b>Peipei Wang</b>, Kathryn T. Stolee, "How well are regular expressions tested in the wild?", Symposium on the Foundations of Software Engineering (FSE), 2018.</p> <p>Ting Dai, Jingzhu He, Xiaohui Gu, Shan Lu, and <b>Peipei Wang</b>, "DScope: Detecting Real-World Data Corruption Hang Bugs in Cloud Server Systems", ACM Symposium on Cloud Computing (SOCC), 2018.</p> <p>Ting Dai, Daniel Dean, <b>Peipei Wang</b>, Xiaohui Gu, Shan Lu, "Hytrace: A Hybrid Approach to Performance Bug Diagnosis in Production Cloud Infrastructures", IEEE Transactions on Parallel and Distributed Systems (TPDS), 2018</p> <p>Carl Chapman, <b>Peipei Wang</b>, and Kathryn T. Stolee, "Exploring Regular Expression Comprehension". International Conference on Automated Software Engineering (ASE), 2017.</p> <p><b>Peipei Wang</b>, Hiep Nguyen, Xiaohui Gu, Shan Lu, "RDE: Replay DEbugging for Diagnosing Production Site Failures", Proc of IEEE International Symposium on Reliable Distributed Systems (SRDS), 2016.</p> <p>Rui Shu, <b>Peipei Wang</b>, Sigmund A. Gorski III, Benjamin Andow, Adwait Nadkarni, Luke Deshotels, Jason Gionta, William Enck and Xiaohui Gu, "A Study of Security Isolation", ACM Computing Surveys (CSUR), 2016.</p> <p>Daniel Dean, Hiep Nguyen, <b>Peipei Wang</b>, Xiaohui Gu, Anca Sailer, Andrzej Kochut, "PerfCompass: Online Performance Anomaly Fault Localization and Inference in Infrastructure-as-a-Service Clouds", IEEE Transactions on Parallel and Distributed Systems (TPDS), 2015.</p> <p>Daniel Dean, <b>Peipei Wang</b>, Xiaohui Gu, William Enck, Guoliang Jin, "Automatic Server Hang Bug Diagnosis: Feasible Reality or Pipe Dream?", ICAC 2015.</p> <p><b>Peipei Wang</b>, Daniel Dean, Xiaohui Gu, "Understanding Real World Data corruption Bugs in Cloud Systems", Proc. of IEEE International Conference on Cloud Engineering (IC2E), 2015.</p> <p>Daniel Dean, Hiep Nguyen, <b>Peipei Wang</b>, Xiaohui Gu, "PerfCompass: Toward Runtime Performance Anomaly Fault Localization for Infrastructure-as-a-Service Clouds", Proc. of USENIX Workshop on Hot Topics in Cloud Computing (HotCloud), 2014.</p>
-------------	---

---

**PATENT** Yong Qi, **Peipei Wang**, Tao Yang, Yingyao Hao, "An implement of EJB Container for AOP Based on dynamic stack" (China Patent, Publication Number: CN102508668 A, Application Number: CN201110357781.6)

---

**INTERNSHIPS** Unsupervised system metrics segmentation and log differentiation  
IBM Corporation, RTP, NC Jun 2018 - Aug 2018

The goal of this project is to explore the importance of string literals in software development in terms of the percentage of string literals in code commits and bug fixes. We first collect the record of string literals through the commit history of each project. We then compared the difference line by line in the commits to identify the changed string literals. We also classified the types of various string literals, and the string literal changes related to bug fix.

Unsupervised system metrics segmentation and log differentiation  
IBM China Research Laboratory (CRL), Beijing, China Jun 2015 - Aug 2015

The goal of this project is to combine the information from both metrics and logs and to find system actions during significant changes of metrics. We first found the metrics segments by changepoint detection. We then trained the logs into log templates and compared the changes of different templates in each segmentation.

*Visualized library dependencies among projects in JavaScript*

Morgan Stanley, Shanghai, China Jul 2012 - Sep 2012

This project developed a Web UI tool to explore the library dependencies among different projects. Java developed projects often introduce a lot of third-party libraries, and things will become very complex when projects are built upon other projects. Cycles may happen, library conflicts also happen. This tool can plot this relationship into graphs using d3 JavaScript library, look for particular libraries, check the cycles, and find the dependency path from one project to the project which imports the specified library.

---

**PRESENTATION** Peipei Wang, "How well are regular expressions tested in the wild?", Symposium on the Foundations of Software Engineering (FSE), 2018.

Peipei Wang, "Exploring Regular Expression Comprehension", International Conference on Automated Software Engineering (ASE), November 2017.

Peipei Wang, "Understanding Real World Data corruption Bugs in Cloud Systems", International Conference on Cloud Engineering (IC2E), March 2015.

---

**TECHNICAL SKILLS**

- Language: proficient in Java, Python, R scripts, and C++, familiar with C, Shell script, JavaScript, and JSP.
- OS: proficient in Linux
- System: familiar with Docker, Spark, Hadoop, Cassandra, Squid, MySQL, Apache HTTP server, Memcached, etc

---

**LANGUAGES** Fluent in English and Chinese