

# KEWEN WANG

Email: wangkewen001@gmail.com    Website: <http://wangkewen.github.io>

Address: Storrs Mansfield, CT 06269

## EDUCATION

---

<b>University of Connecticut</b> Ph.D. in Computer Science, GPA: 4.0	<i>2014 -2019</i>
<b>Georgia State University</b> Ph.D. in Computer Science, GPA: 3.8	<i>2013 -2014</i>
<b>Beihang University</b> M.S. in Computer Science, GPA: 3.3	<i>2010 -2013</i>
<b>Beijing Information Science and Technology University</b> B.S. in Computer Science, GPA: 3.5	<i>2005 -2009</i>

## TECHNICAL SKILLS

---

<b>Computer Languages</b>	Java, Linux Shell, C, Python
<b>Open Source</b>	Apache Spark, Apache Hadoop, Xen, Apache Tomcat, BTrace, Ganglia
<b>Web Development</b>	JSP, Ajax, CSS, JavaScript, jQuery

## CODING COMPETITION

---

**Google Code Jam 2017** Qualification Round Rank#1483/25k, Round 1C Rank#1664/3775

## ACADEMIC SERVICE

---

Reviewer of IEEE Transactions on Parallel and Distributed Systems (TPDS).  
Reviewer of ACM Transactions on Architecture and Code Optimization (TACO).

## RESEARCH PROJECTS

---

<b>Interference Modeling of Apache Spark Jobs</b> <i>Research Assistant</i>	Aug 2015 - Nov 2016 <i>University of Connecticut</i>
--	---

- Integrated resources consumption and tasks events profiles for Spark Jobs run in Xen virtual machines.
- Built an inference model to predict the execution time of multiple Spark jobs executed in parallel.
- Implemented Spark jobs scheduler to optimize the total execution time.

<b>Performance Prediction for Apache Spark Jobs</b> <i>Research Assistant</i>	Oct 2014 - May 2015 <i>University of Connecticut</i>
--	---

- Parsed JSON format events logs of Apache Spark jobs, and analyzed tasks execution pattern.
- Established an analytical performance model to predict time, I/O overhead and memory consumption.

<b>Learning environment for Smart Grid security</b> <i>Research Assistant</i>	Aug 2013 - Feb 2014 <i>Georgia State University</i>
--	--

- Implemented an online tool using JSP and jQuery to schedule Smart Grid emulator for courses design.

<b>Optimizing Hadoop MapReduce</b> <i>Research Assistant</i>	Nov 2011 - Dec 2012 <i>Beihang University</i>
---	--

- Applied BTrace to trace MapReduce job functions, monitored resources consumption using Ganglia.
- Constructed Hadoop performance model for execution time prediction.
- Designed heuristic search algorithm to find optimal configuration for MapReduce jobs.

## WORK EXPERIENCE

---

### Full Stack Developer

Nov 2011 - Jan 2012

#### Science and Technology Research Institute of Beihang University

- Built website on the framework Struts+Spring+Hibernate, and loaded project archives into MySQL.
- Implemented information retrieval and display using JSP, JavaScript and Ajax.

### Software Engineer Intern

Mar 2010 - May 2010

#### NDtech Inc. Beijing, China

- Analyzed ANTLR (an open source parser generator) to learn C# parser and Script#.
- Applied Script# to write JavaScript by compiling C#.

### Test Engineer Intern

Oct 2008 - Nov 2008

#### National Computer Products Quality Supervising Test Center, Beijing, China

- Applied black-box testing to test an information management software.

## AWARDS

---

### Predoctoral Fellowship

2017

Computer Science and Engineering department at University of Connecticut

### Third Class Scholarship

2011

Beihang University

### Academic Scholarship

2008

Beijing Information Science and Technology University

### Municipal 2nd Prize of 21st National Middle School Students Physics Competition

2004

City of Xianning, China

## PUBLICATIONS

---

1. Modeling Interference for Apache Spark Jobs. Wang, Kewen, Mohammad Maifi Hasan Khan, Nhan Nguyen, and Swapna Gokhale. IEEE 9th International Conference on Cloud Computing (CLOUD), 2016.
2. CSMiner: An Automated Tool for Analyzing Changes in Configuration Settings across Multiple Versions of Large Scale Cloud Software. Nguyen, Nhan, Mohammad Maifi Hasan Khan, and Kewen Wang. IEEE 9th International Conference on Cloud Computing (CLOUD), 2016.
3. Performance Prediction for Apache spark Platform. Wang, Kewen, Mohammad Maifi Hasan Khan. IEEE 17th International Conference on High Performance and Communications (HPCC), 2015.
4. Integrated Learning Environment for Smart Grid Security. Wang, Kewen, Yi Pan, Wen-Zhan Song, Weichao Wang, and Le Xie. The Fourth International Conference on Advanced Communications and Computation (INFOCOMP), 2014.
5. Predator - An experience guided configuration optimizer for Hadoop MapReduce. Wang, Kewen, Xuelian Lin, and Wenzhong Tang. IEEE 4th International Conference on Cloud Computing Technology and Science (CloudCom), 2012.