

# JIANFENG CHEN

(919)-457-2034 ◊ <http://jianfeng.us> ◊ [jchen37@ncsu.edu](mailto:jchen37@ncsu.edu)

## EDUCATION

---

**PhD in Computer Science** Aug 2014 - May 2019 (expected)

North Carolina State University, GPA: 4.0/4.0

Coursework: Data-to-Knowledge | DevOps | Advanced AI | Algorithm Analysis | Data Mining | Automated SE

**BS in Computer Science** Sep 2010 - May 2014

Shandong University, China, GPA: 91.1/100

Coursework: Data Structure | OS | Networking | Database System | Numerical Analysis | Image Processing

## SKILLS AND INTERESTS

---

Proficient in **language**: Python, Java; familiar with JavaScript, C++, Matlab and SQL;

Proficient in **data analysis** tools: Scikit-learn, SciPy, Pandas, jMetal, Gephi;

Familiar with **DevOps** tools: Jenkins, Ansible, Travis-CI, AWS Elasticsearch, S3, Docker, Redis.

## SELECTED PROJECTS

---

**Automated Configurations for Cloud-based Workflows** May 2017 - Aug 2017

North Carolina State University

- Presented a novel stochastic method for rapidly configuration cloud-based workflows
- Automatically deployed the workflow with more than 500 sub-tasks to AWS platform. Save more than 30% economy cost within specific deadline requirement, compared to default greedy deployment policy in AWS.

**LACE Data Privatization Tools and its Application** Aug 2016 - Nov 2016

NSA funded project in RAISE Lab

- Distributed a data anonymization package in Python (see <http://tiny.cc/pydp>); tested package via Travis-CI.
- Applied my package to education and medical data sets. Evaluate data set utility through supervised learning.

**Fast Principal-component-analysis (F-PCA) Method for Flight Status Log** May 2016 - Aug 2016

Google Summer of Code program 2016

- Accepted by Google GSoc2016 program among 18,981 applicants (**accept rate: 6%**).
- Hierarchical clustering a dataset(flight status log) with more than 20M entries top-down and bottom-up. Create a PCA-like dimension reduction algorithm and speed it up by spark. Compared my own algorithm with PCA.

**Building Movie Recommendation System** Aug 2015 - Dec 2015

"Netflix Prize" completion extension

- Build a movies recommendation system by training from 100 million Netflix ratings by Factorization Machine, SVM and ANN. Accelerated the learning process through high performance computing (HPC) server.
- Crawled cast, critic reviews from rotten tomatoes and classified the movies basing on Jaro-Winkler Distance. Reduced the RMSE by up to 9% with the help of external information.

**Continuous Integration/Delivery Pipeline** Aug 2015 - Dec 2015

DevOps practice

- Basing on abstract syntax tree, created a test-suite generator to *only* test the diff between two commits.
- Integrated Ansible scripts, Docker and Jenkins to build, test and deploy our "sunrise-calculator" application.

## PUBLICATIONS

---

[1] [Chen, J.](#), and [Menzies, T.](#) "RIOT: a Novel Stochastic Method for Rapidly Configuring Cloud-Based Workflows." *IEEE International Conference on Cloud Computing 2018* (Accept rate: 15%).

[2] [Chen, J.](#), [Nair, V.](#), [Krishna, R.](#), and [Menzies, T.](#) "Sampling as a Baseline Optimizer for Search-based Software Engineering" *To appear. IEEE Transactions on Software Engineering (2018)*.

[3] [Chen, J.](#), [Nair, V.](#), and [Menzies, T.](#) Beyond Evolutionary Algorithms for Search-based Software Engineering. *Information and Software technology Volume 95, Mar 2018, p.281-294*.