

Zhouhan Chen

60 5th Avenue, 7th Floor
New York, NY 10011
+1 281-777-5808
zc1245@nyu.edu
zhouhanc.com

RESEARCH FOCUS

Cyber Security, Social Network Analysis, Misinformation Detection, Data Mining, Web crawling

EDUCATION

Ph.D. Student, Data Science

Center for Data Science, New York University, New York, NY August 2018 - May 2022 (expected)

Master of Science, Computer Science

Rice University, Houston, TX

August 2017 - May 2018

Bachelor of Science, Computer Science

Rice University, Houston, TX

August 2013 - May 2017

RELATED COURSEWORK

Natural Language Processing with Representation Learning, Machine Learning, Big Data, Optimization and Computational Linear Algebra, Computer Networks, Computer Security, Operation System, Distributed Program Construction, Artificial Intelligence

INDUSTRY EXPERIENCE

Google

June 2020 - August 2020 (expected)

Incoming Google Research Intern, Sunnyvale, CA

- Google Research Anti-abuse Team. Research will focus on virus detection and system security.

Twitter

June 2019 - August 2019

Health Machine Learning Intern, San Francisco, CA

- Designed and deployed an automated system to detect coordinated manipulation on the platform using unsupervised graph clustering. The system detects +5k bad actors every day with +95% accuracy.
- Investigated behaviors of suspicious accounts, and discovered new features that can be used in next-generation spam detection algorithm.
- Collaborated closely with other teams including Trust & Safety, Site Integrity and Investigation to share URL threat intelligence and to define a clearer anti-abuse policy.

Webroot

June 2018 - August 2018

Machine Learning Intern, San Diego, CA

- Designed a phishing domain classifier using LSTM and deep neural networks in Pytorch.

- Deployed a real-time, online-learning phishing domain detector on Amazon Web Service, scanning millions of URLs per hour.
- Performed extensive neural network performance tuning, hyperparameter search, and increased test accuracy from baseline 70% to 89%.

Quantlab Financial

June 2017 - August 2017

Data Engineer Intern, Houston, TX

- Developed API endpoints in Elixir to expose real-time trading data to internal teams.
- Created an Elixir pipeline that pushes Gigabytes of market data to Elasticsearch every second.
- Created a web application to display trading data in real-time using Chart.js.
- Implemented automatic web testing framework using Docker, Mocha and NodeJS.

App Annie

May 2016 - August 2016

Software Development Engineer Intern, Beijing, China

- Customized Hue (a web application for analyzing big data) to support database connections to Hadoop Hive and Presto to facilitate data accessibility for internal customers.
- Developed Python ETL jobs to migrate data from PostgreSQL to Amazon S3.

PUBLICATIONS

1. Z. Chen, et al., **Proactive Discovery of Fake News Domains from Real-Time Social Media Feeds**, WWW' 20, the Web Conference, Proceedings, 2020
2. Z. Chen, et al., **An Unsupervised Approach to Detect Spam Campaigns that Use Botnets on Twitter**, arXiv:1804.05232, 2018
3. Z. Chen, et al., **Hunting bots on Twitter: an unsupervised approach**, Social Informatics: 9th International Conference, Proceedings, 2017
4. R. Tanash, **Z. Chen**, et al., **The Decline of Social Media Censorship and the Rise of Self-Censorship after the 2016 Failed Turkish Coup**, 7th USENIX Workshop on Free and Open Communications on the Internet, Proceedings, 2017
5. R. Tanash, **Z. Chen**, et al., **The Dynamics of Social Media Censorship in Transitioning Democracies**, Annual Meeting of the American Political Science Association, Proceedings, 2016
6. R. Tanash, **Z. Chen**, et al., **Known Unknowns: An Analysis of Twitter Censorship in Turkey**, 14th ACM Workshop on Privacy in the Electronic Society, Proceedings, 2015

PRESENTATION

- **Hunting Malicious Bots on Twitter: An Unsupervised Approach**, Presented at *Twitter Data Science Tech Talk*, December 2017

PROGRAM COMMITTEE

- ACL SRW (Association for Computational Linguistics Student Research Workshop), 2020.

PROGRAMMING SKILLS

- Languages: Python, Java, JavaScript, C, Scala, Bash, Git, SQL, MongoDB
- Web Framework and Cloud Computing: Django, Flask, NodeJS, Docker, AWS (EC2, EMR, S3), Hadoop (Hive, Hue, HDFS)
- Deep learning and Data analytics: Pytorch, Tensorflow, Pandas, Scikit-learn, d3js

HONORS

- Rice University Honor: Distinction in Research and Creative Works May 2017
- A university honor granted annually at commencement. Three CS students received this distinction in 2017.
 - Recommended by Computer Science Department for creativity, self-initiation, independence, perseverance, and dedication.
- 2nd Place in the School of Engineering at Rice Undergraduate Research Symposium April 2017
- President's Honor Roll Fall 2014, Fall 2015

ACTIVITIES

- President of Rice Road Runners Club July 2014 - September 2016
- Organized weekly group run of 10 students
 - Collaborated with other clubs to promote students' health and well-being
- Certified Personal Trainer, National Council on Strength and Fitness (NCSF) 2017-
- Certified Sport Nutrition Specialist, National Council on Strength and Fitness (NCSF) 2017-
- Ultramarathon Runner The North Face Endurance Challenge (New York) 50km Finisher 2019-