

Zhouhan Chen

Cell: 281-777-5808
Email: chenzhouhan0124@gmail.com
Website: zhouhanc.com

EXPERTISE

Cyber Security, Anti-Abuse, Social Network Analysis, Data Engineering, Data Visualization.

CURRENT OCCUPATION

Safe Link Network Corporation July 2022 -
Founder.

1. Project **Information Tracer** - spam and fake news detection on social media.
2. For Information Tracer, I've received \$12,000 funding from NYC Media Lab, and won the 2022 Twitter Chirp Developer Challenge (Content Safety Tools Category, \$10,000 prize).
3. Project **Malware Discoverer** - malicious URL redirection campaign detection. I share threat intelligence reports based on 15,000+ daily deep crawls.
4. I also provide consulting to companies and NGOs such as Amnesty International USA.

EDUCATION

PhD in Data Science
New York University, New York, NY August 2018 - May 2022
Advisors: Professor Richard Bonneau, Professor Juliana Freire, Professor Joshua Tucker
Thesis: Be careful what you click: tracing URLs that spread false or malicious information

Master of Science in Computer Science
Rice University, Houston, TX August 2017 - May 2018
Advisor: Professor Devika Subramanian
Thesis: An Unsupervised Approach to Detect Spam Campaigns that Use Botnets on Twitter

Bachelor of Science in Computer Science
Rice University, Houston, TX August 2013 - May 2017

INDUSTRY EXPERIENCE

Amazon June 2021 - August 2021
Security Engineer Intern, New York City

- I was a member of the at Security Analytics and AI Research Team, building cloud-native threat intelligence systems to protect customers using Amazon Web Services.
- Designed unsupervised clustering algorithms using LSH (locality sensitive hashing), multi-hop algorithms and rule-based logics.
- Built end-to-end clustering system using CloudFormation, Lambad, Spark and EMR.
- Investigated unknown phishing and malvertising campaigns, and collaborated with internal AWS security teams.

Google

June 2020 - August 2020

Software Engineer Intern, Remote

- I was a member of the Security and Anti-abuse Research Team. I contributed to the open-source project Maldoca (Malicious Microsoft Office document analyzer).
- Investigated malicious PDF documents based on Adobe CVE (Common Vulnerabilities and Exposures), Virustotal samples and Yara Rules, and identified high-impact features.
- Implemented C++ functions, protobufs and unit tests for PDF parser and feature extractor.

Twitter

June 2019 - August 2019

Platform 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 and Site Integrity 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. **Zhouhan Chen** et al., What we learned about The Gateway Pundit from its own web traffic data, International Workshop on Cyber Social Threats, 2022
2. **Zhouhan Chen** and Juliana Freire, Discovering and Measuring Malicious URL Redirection Campaigns from Fake News Domains, 2021 IEEE Security and Privacy Workshops (SPW), 2021 (*Best Paper Award*)

3. **Zhouhan Chen** et al., An Automatic Framework to Continuously Monitor Multi-Platform Information Spread, Workshop on Misinformation Integrity in Social Networks, 2021
4. **Zhouhan Chen** and Juliana Freire, Proactive Discovery of Fake News Domains from Real-Time Social Media Feeds, WWW' 20, the Web Conference, Proceedings, 2020 (*Best Paper Award*)
5. **Zhouhan Chen** and Devika Subramanian, An Unsupervised Approach to Detect Spam Campaigns that Use Botnets on Twitter, arXiv:1804.05232, 2018
6. **Zhouhan Chen** and Devika Subramanian, Hunting Bots on Twitter: An Unsupervised Approach, Social Informatics: 9th International Conference, Proceedings, 2017
7. Rima Tanash, **Zhouhan Chen**, Dan Wallach, and Melissa Marschall, 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
8. Rima Tanash, Abdullah Aydogan, **Zhouhan Chen**, et al., Detecting Influential Users and Communities in Censored Tweets Using Data-Flow Graphs, Annual Meeting of the Society for Political Methodology (POLMETH), Proceedings, 2016
9. Rima Tanash, **Zhouhan Chen**, Tanmay Thakur, Dan Wallach, and Devika Subramanian, Known Unknowns: An Analysis of Twitter Censorship in Turkey, 14th ACM Workshop on Privacy in the Electronic Society, Proceedings, 2015

INDUSTRY TALKS

- **Uncovering Multi-platform Misinformation Campaigns with Information Tracer**, Presented to *DEFCON 30*, August 2022
- **Discovering and Tracing Misinformation**, Presented to *Google Trust and Safety Intelligence Investigation Team*, August 2020
- **Unsupervised Discovery of Malware Redirection Campaigns from Fake News Domains**, Presented to *Amazon Web Service Security Team*, August 2020
- **Hunting Malicious Bots on Twitter: An Unsupervised Approach**, Presented at *Twitter Data Science Tech Talk*, December 2017

Media Coverage

1. Announcing funding for three AI & Local News teams, NYC Media Lab, 2022
2. How CDS PhD student Zhouhan Chen hunts malicious URL redirection campaigns, 2021
3. Information Tracer, a proactive framework to fight COVID-19 infodemic, 2021
4. Zhouhan Chen: CS Researcher and Marathon Runner, 2017

PROGRAMMING SKILLS

- Python, Java, JavaScript, Bash, Git, SQL, MongoDB, Spark, Bootstrap, D3js
- Proficient in deploying servers on cloud computing platforms. Frequent user of Flask, Docker, AWS, GCP, Hadoop (Hive, Hue, Presto)