Tianyi Chen

E-mail: ctianyi7@gmail.com mobile: 669-900-2291

EDUCATION

Graduate School of Arts and Sciences, Boston University **2019 - 2023**
• Ph.D., Department of Computer Science
Advisor: Charalampos E. Tsourakakis

Graduate School of Arts and Sciences, Boston University **2017 - 2019**
• M.S., Department of Computer Science

School of Software Engineering, Xi'an Jiaotong University **2013 - 2017**
• B.S., Department of Software Engineering

PUBLICATION

- Tianyi Chen, Charalampos E. Tsourakakis, "Modeling the Emergence of Polarized Communities: Theory and Applications", *under review*.
- Charalampos E. Tsourakakis, Tianyi Chen, Naonori Kakimura, Jakub Pachocki, "Novel Dense Subgraph Discovery Primitives: Risk Aversion and Exclusion Queries", *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases* (ECML PKDD) 2019.
- Ming Fan, Jun Liu, Xiapu Luo, Kai Chen, Tianyi Chen, Zhenzhou Tian, Xiaodong Zhang, Qinghua Zheng, Ting Liu, "Frequent Subgraph based Familial Classification of Android Malware", *IEEE 27th International Symposium on Software Reliability Engineering* (ISSRE) 2016, ***BEST RESEARCH PAPER AWARD.***

EXPERIENCE

Software Engineer Intern, Google **June 2020 - August 2020**
Project: Botnets Anomaly Detection Based on Graph Convolutional Network (GCN)
- Sampled neighborhoods in parallel using C++ and GCN input pipeline from user publisher bipartite graph.
- Implemented distributed GCN auto-encoder with skip-connection in TensorFlow that reconstruct both seed node features and subgraph adjacency.
- Anomaly detection model based on encoder-decoder's reconstruction error showed high correlation to anomaly metrics, outperformed benchmark MLP model, and found new suspicious publishers.

Software Engineer Intern, Google **May 2019 - August 2019**
Project: Android Malware Detection Based on Dynamic Analysis Trace
- Used C++ flume to fetch and preprocess Android APK evaluation data in parallel.
- Implemented LSTM with TensorFlow Estimator to extract features from dynamic analysis sequence.
- Model reached 0.96 recall rate while the bag-of-words baseline only got 0.88.

Research Assistant, Harvard University **May 2018 - May 2019**
Project: Functional Object-Oriented Graph Automation platform
- Transformed Python program into workflow for reproducing computations and reusing data.
- Investigated on the performance of graph database Neo4j and deployed it to improve query efficiency in workflow structures ten times compared with PostgreSQL.
- Designed and built front-end GUI with HTML, CSS and JavaScript, jQuery. Built server on Google Cloud with Python Flask to handle requests.

Research Assistant, Xi'An Jiaotong University **December 2014 - June 2017**
Project: Detection and Classification of Android Malware
- Managed data collection of permissions and sensitive APIs used by 10000 malicious Android APP samples, de-compiled those Apps and analyzed the Manifest files as well as Smali codes.
- Investigated function-call relation graph in samples, and participated in discovery of sensitive subgraph structure.
- Cleaned API data and trained random forest model with both API features and sensitive subgraph features.

SKILLS

Programing language: C++, Java, JavaScript, SQL, Python, Git, HTML, CSS.
Machine learning: TensorFlow, numPy, Pandas, Scikit-Learn, Matlab, WEKA.