Yuhang Yao

Phone: (+1) 412-209-5691
Email: yuhangya@andrew.cmu.edu
Homepage: http://yh-yao.github.io
Carnegie Mellon University

EDUCATION

Carnegie Mellon University   *Aug. 2019 - Present*
Ph.D. in LIONS Lab, Dept. of Electrical and Computer Engineering
- Overall GPA: 4.0/4.0. Advisor: Prof. Carlee Joe-Wong

Shanghai Jiao Tong University   *Sept. 2015 - Jun. 2019*
B.S. in Computer Science, IEEE Honor Class

PUBLICATIONS

- **Yuhang Yao**, Xiao Zeng, Tianyue Cao, Luoyi Fu and Xinbing Wang, "APRP: An Anonymous Propagation Method in Bitcoin Network", in *AAAI 2019 poster*.
- **Yuhang Yao**, Junjie Ou, Yang Li, Luoyi Fu, Xinbing Wang and Guihai Chen, "Turing Index: Cross-domain and Cross-generation Metric of Unraveling Scholars' Impact in Academic Big Data", in *Big Data Research* 2019.
- Jiaqi Liu, **Yuhang Yao**, Xinzhe Fu, Luoyi Fu and Xinbing Wang, "Evolving K-Graph: Modeling Hybrid Interactions in Networks", in *ACM MobiHoc 2017 poster*.
- Jiayu Pan, **Yuhang Yao**, Luoyi Fu and Xinbing Wang, "Core Percolation in Coupled Networks", in *ACM MobiHoc 2017 poster*.
- Xudong Wu, Luoyi Fu, **Yuhang Yao**, Xinzhe Fu, Xinbing Wang and Guihai Chen, "GLP: A novel Framework for Group-level Location Promotion", in *IEEE/ACM Trans. on Networking* 2018.
- Jiaqi Liu, Luoyi Fu, **Yuhang Yao**, Xinzhe Fu, Xinbing Wang and Guihai Chen, "Modeling, Analysis and Validation of Evolving Networks with Hybrid Interactions", in *IEEE/ACM Trans. on Networking* 2018.
- Jiapeng Zhang, Luoyi Fu, Shuhao Li, **Yuhang Yao**, Xinbing Wang, "Core Percolation in Independent Networks", in *IEEE Trans. on Network Science and Engineering* 2018.

RESEARCH EXPERIENCES

Link Prediction and Node Classification with Recurrent Graph Convolutional Network
Nov. 2019 - Present
Advisor: Prof. Carlee Joe-Wong   *LIONS Lab*
- Derived a general model to learn the representation of dynamic evolving graphs which contain complex temporal information in reality, and predict the links and labels in future timestamp;
- Proposed Recurrent Graph Convolutional Network(RGCN), a novel model that uses RNNs to capture temporal evolution of graphs and GCNs to generate node embeddings for prediction;
- Analysed the interpretability of Recurrent Neural Network in evolution of temporal graphs and features by both theory and experiments on Stochastic Block Model and real world networks;
- Experiments on social networks and face-to-face networks demonstrate the performance of our model on mean average precision and mean reciprocal rank over state-of-the-art methods.

Smart Data Pricing for Edge Computing   *Aug. 2019 - Present*
Advisor: Prof. Carlee Joe-Wong   *LIONS Lab*
- Derived a model to describe pricing, resource allocation and relationship migration for heterogeneous edge servers, providers and users;
- Established strategy evaluation on pricing and service quality, especially on lower latency to maximum the expected market size with competition among server providers;
- Analysed the optimal strategy for late comers to catch up with early comers and gain more market share under budget constraint and user dependence;
- Built experiment on both simulated and real edge dataset to demonstrate the validation and accuracy of the model.

Turing Index: A Cross-domain & Cross-generation Metric of Unraveling Scholars' Impact
Mar. 2018 - June. 2019
Advisor: Prof. Xinbing Wang   *Research Center of Intelligent Internet of Things*

- Derived a novel metric named Turing index that captures scholars' impact across different domains and generations;
- Investigated data collected from 14,223,183 scholars, 126,438,664 papers, 533,556,856 references and 310 domains spanning from 1865 to 2016, and applied the Power Law to citation distributions of scholars across different domains and generations;
- Normalized multifarious citation distributions using the inflation adjustment factor that quantifies the absolute status of scholars to eliminate the effect of rapid inflation;
- Established a system containing 12 billion scholars' Turing index data to compare cross-domain and cross-generation scholars.

APRP: An Anonymous Method in Blockchain   *Feb. 2018 - Sep. 2018*
Advisor: Prof. Xinbing Wang   *Research Center of Intelligent Internet of Things*

- Presented a general adversarial network model for formulizing deanonymization attacks;
- Designed a novel propagation method called APRP(Adaptive PageRank Propagation) that adopts PageRank as a propagation delay factor and constantly adjusts the PR-value of nodes to adapt to network dynamics;
- Experimented on both simulated and real Bitcoin networks, and confirmed the superiority of APRP which increased performance by 20% to 50% under various deanonymization attacks.

Core Percolation and Influence Maximization in Social Networks   *Aug. 2016 - Feb. 2018*
Advisor: Prof. Xinbing Wang   *Research Center of Intelligent Internet of Things*

- Studied the core percolation (a fundamental structural transition resulting from preserving core nodes in the network) and influence maximization in multiple interdependent networks;
- Proposed GLR, an alternating greedy leaf reduction method, which recursively switches among networks for node removal, and GLP, a new and novel framework of group-level location promotion in Geo-Social Networks;
- Found that the presence of cores exhibits a jump at the critical point as a first order transition in multi-layer networks, while it undergoes a continuous second order transition in a single network;
- Extensive experiments on real datasets demonstrate that the GLP outperforms four baselines, with notably up to 10 times larger influence spread and 100 times faster seed selection over two individual-level cases.

COURSES AT CMU

- 10-725 Convex Optimization
- 11-785 Introduction to Deep Learning
- 15-719 Advanced Cloud Computing
- 18-661 Intro to Machine Learning for Engineers

SELECTED PROJECTS

Distributed Machine Learning by Spark   *Jan. 2020 - Present*
Advisor: Prof. Greg Ganger

- Developed a distributed machine learning system for large scale datasets based on pyspark;
- Perform super efficient distributed gradient descent on a KDD Cup 2012 dataset with 54,686,452 features and a binary click-through prediction dataset released by Criteo with 882,774,562 features;
- Achieve 60x speedup per iteration over a single machine.

Expert Portrayal in Academic Networks   *Sept. 2018 - June. 2019*
Group Leader

- Created accurate expert portraits and visualizations of scholars in NSF and NSFC;
- Used the Turing index to describe the inflation of domain and times in funding evaluations;
- Studied the career trajectory of scholars from bachelor degrees, PhDs, to faculty terms;
- Site Url: http://acemap.info/.