YUE YU

E-mail: yueyu@gatech.edu | Tel: (+1) 217-721-1269 | Webpage: yueyu1030.github.io

Address: CODA Building 1312, 756 W Peachtree St NW, Atlanta, GA 30308, USA

EDUCATION

School of Computational Science and Engineering, Georgia Institute of Technology **Atlanta, GA, USA**

Ph.D. in Computational Science and Engineering *Aug. 2019 - Present*

- Ph.D. Advisor: Dr. Chao Zhang;
- GPA: 4.0/4.0;
- Research Interest: Text Mining, Applied Machine Learning, Information Extraction.

Department of Electronic Engineering, Tsinghua University **Beijing, China**

B.Eng. in Electronic Engineering *Aug. 2015 - July 2019*

- GPA: 3.87/4.00 (Rank: 5/246) ;
- Research Assistant in the Future Internet & Communication Lab advised by Dr. Yong Li;

PUBLICATIONS & MANUSCRIPTS

Publications (* stands for equal contribution, same as belows):

1. **Yue Yu**, Yinghao Li, Jiaming Shen, Hao Feng, Jimeng Sun and Chao Zhang, "STEAM: Self-Supervised Taxonomy Expansion via Path-Based Multi-View Co-Training", In *Proceedings of the 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'20)*, 2020.
2. Chen Liang*, **Yue Yu***, Haoming Jiang*, Siawpeng Er, Ruijia Wang, Tuo Zhao and Chao Zhang, "BOND: Bert-Assisted Open-Domain Named Entity Recognition with Distant Supervision", In *Proceedings of the 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'20)*, 2020.
3. **Yue Yu***, Tong Xia*, Huandong Wang, Jie Feng and Yong Li, "Semantic-aware Spatio-temporal App Usage Representation via Graph Convolutional Network", In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT/UbiComp'20)*, 2020.
4. Mingyang Zhang, Tong Li, **Yue Yu**, Yong Li, Pan Hui and Yu Zheng, "Urban Anomaly Analytics: Description, Detection and Prediction", In *IEEE Transactions on Big Data*, 2020.
5. Tong Xia, Yong Li, **Yue Yu**, Fengli Xu, Qingmin Liao and Depeng Jin, "Understanding Urban Dynamics via State-sharing Hidden Markov Model" (extended version of previous WWW paper), In *IEEE Transactions on Knowledge and Data Engineering*, 2020.
6. Tong Xia*, **Yue Yu***, Fengli Xu, Funing Sun, Diansheng Guo, Depeng Jin and Yong Li, "Understanding Urban Dynamics via State-sharing Hidden Markov Model", in *Proceedings of the Web Conference (WWW'19)*, 2019.
7. Chen Gao, Chao Huang, **Yue Yu**, Huandong Wang, Yong Li and Depeng Jin, "Privacy-preserving Cross-domain Location Recommendation', in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT / UbiComp'19)*, 2019.

Manusripts:

1. **Yue Yu***, Simiao Zuo*, Haoming Jiang, Wendi Ren, Tuo Zhao and Chao Zhang, "Anonymous Submission", submitted to *The 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP'20)*, 2020.
2. Rongzhi Zhang, **Yue Yu** and Chao Zhang, "Anonymous Submission", submitted to *The 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP'20)*, 2020.
3. Xiaojun Xu, **Yue Yu**, Le Song, Chengfeng Liu, Havya Kaikhura, Carl A. Gunter and Bo Li, "Adversarial Edge Detection For Graph Neural Networks", submitted to *the ACM Conference on Computer and Communications Security*, 2020.

RESEARCH EXPERIENCE

Data Mining and Machine Learning Group, Georgia Tech **Atlanta, GA, USA**

Advisor: Dr. Chao Zhang

- **Self-Supervised Taxonomy Expansion via Path-Based Multi-View Co-Training** *Aug. 2019 - Feb. 2020*
  - Proposed a self-supervised framework that performs taxonomy expansion with natural supervision signals from existing taxonomies and text corpora;
  - Designed a mini-path-based anchor format to generate training data that better captures structural information in taxonomies for expansion;
  - Proposed a multi-view co-training procedure that integrates multiple sources of information in an end-to-end model.
- **Bert-Assisted Open-Domain Named Entity Recognition with Distant Supervision** *Aug. 2019 - Feb. 2020*
  - Demonstrated that the pre-trained LM provides additional semantic information during the training process and reduce the label noise for distantly-supervised NER.
  - Designed a two-stage framework to fully exploit the power of language models in our task. Specifically, we refine the distant label iteratively with the LM in the first stage and improve the model fitting under the teacher-student framework in the second stage, which is able to address the challenge of noisy and incomplete annotation.

Secure Learning Lab, University of Illinois at Urbana-Champaign **Urbana, IL, USA (Remote)**

Advisor: Dr. Bo Li

- **Robust Graph Neural Networks against Adversarial Edges** *July 2018 - Feb 2019*
  - Designed a novel graph generation method together with link prediction method to detect malicious edges.
  - Leveraged novel features to perform outlier detection for detection.
  - Conducted extensive experiments and showed that the proposed detection mechanism can achieve AUC above 90% against attack strategies on both real-world network datasets and synthetic networks.

Future Internet & Communication Lab, Tsinghua University **Beijing, China**

Advisor: Dr. Yong Li

- **Spatio-temporal Data Mining and Recommender Systems** *Dec. 2017 - July 2019*
  - Designed a novel urban dynamic revealing system based on state-sharing HMM to identify the typical dynamic patterns on various regions of the city with different urban functions.
  - Presented a new framework for privacy-preserving cross-domain location recommendation, in which data from the transferred domain is protected with the criteria of differential privacy. Designed confidence-enhanced collective matrix factorization (CCMF) to balance two domains' influence for better recommendation.
  - Built a heterogeneous App usage graph by regarding App, time, and location units as nodes and their co-occurrence relations as edges. Developed a Graph Convolutional Network with meta path-based objective function to learn the semantic-aware representations of units.

HONORS AND AWARDS

- ACM SIGKDD Student Registration Award *Aug. 2020*
- Excellent Graduate, Tsinghua University & Beijing City (Top 2% over 3292 graduate students) *July 2019*
- Comprehensive Scholarship, Tsinghua Univiersity (Top 1%) *Oct. 2018*
- Award from Tsinghua University Initiative Scientific Research Program (5000 USD) *May 2018*
- Comprehensive Scholarship, Tsinghua Univiersity (Top 5%) *Oct. 2017*
- Comprehensive Scholarship, Tsinghua Univiersity (Top 5%) *Oct. 2016*

PROFESSIONAL SKILLS

- Programming language: C++, Python, MATLAB, Latex.

- Deep learning frameworks: Keras, Pytorch.

MISC

- **Teaching Experience**: Teaching Assistant for CX4240: Introduction to Computational Data Analysis. *Jan. 2020 - Now*
- **Reviewing Experience**: Reviewer for AAAI 2020, CIKM 2019.