


# SHIFU HOU

257 Fitzpatrick Hall, University of Notre Dame, Notre Dame, IN 46556-5637

☎ (304) 216-0051 ✉ [sxh1055@case.edu](mailto:sxh1055@case.edu) **in** [Linkedin](#)  [Google Scholar](#)

## RESEARCH INTERESTS

---

- General: machine learning, deep learning, data mining
- Emphases: graph neural networks, natural language processing, heterogeneous information networks
- Applications: cybersecurity, malware detection, social network processing, darknet markets analysis

## EDUCATION

---

**Case Western Reserve University**

June 2019 - July 2022

*Ph.D. in Computer Science*

*Cleveland, OH*

**West Virginia University**

August 2014 - June 2019

*M.S. in Computer Science*

*Morgantown, WV*

**Xiangtan University**

September 2008 - June 2012

*B.E. in Communication Engineering*

*Hunan, China*

## RESEARCH EXPERIENCE

---

**University of Notre Dame**

August 2022 - Present

Postdoc

*Notre Dame, IN*

- Exploring the combination of GNN with new techniques in NLP or CV
- Applying AIGC methods such as stable diffusion on code summarization
- Applying large language model such as Llama and T5 on code security

**Case Western Reserve University**

June 2019 - July 2022

Research Assistant

*Cleveland, OH*

- Developed an AI-driven website to facilitate hierarchical risk assessment of Covid-19
- Designed heterogeneous temporal graph transformer for evolving malware detection
- Applied disentangled representation learning on emotion analysis for social media

**West Virginia University**

August 2014 - June 2019

Research Assistant

*Morgantown, WV*

- Applied transformer on social coding security
- Applied heterogeneous information networks on malware detection
- Explored darknet markets transactions based on graph neural network

**Tencent Security Lab**

July 2018 - September 2018

Research Intern

*Shenzhen, China*

- Designed out-of-sample node embedding model on heterogeneous graphs for Android malware detection
- Applied disentangled representation learning on heterogeneous graphs for Android malware detection

**Cheetah Mobile**

July 2017 - September 2017

Research Intern

*Beijing, China*

- Designed a news recommendation system based on LSTM
- Designed a news classification system based on word2vec

## HONORS

---

MetroLab Innovation of the Month Award	2020
SIGKDD 2017 Best Paper Award	2017
SIGKDD 2017 Best Student Paper Award	2017
National Collegiate Information Safety Contest Third Prize	2011
The ACM Asia Programming Contest Chengdu Site Bronze Medal	2010
The ACM Asia Programming Contest Tianjin Site Bronze Medal	2010
The Sixth Hunan Province Collegiate Programming Contest First Prize	2010
The ACM Asia Programming Contest Wuhan Site Bronze Meda	2009

## PUBLICATIONS

---

**Research papers have been published in top-tier Machine Learning and Data Mining conferences, amassing over 1500 citations on Google Scholar, including KDD (3 papers), AAAI (2 papers), IJCAI (3 papers), WWW (2 papers), CIKM (2 papers), and ACSAC (2 papers).** Full list of papers is as follows:

1. **Shifu Hou**, Lingwei Chen, Mingxuan Ju, Yanfang Ye . "Leveraging Comment Retrieval for Code Summarization", The 45th European Conference on Information Retrieval (ECIR), 2022.
2. **Shifu Hou**, Lingwei Chen, Yanfang Ye . Summarizing Source Code from Structure and Context," International Joint Conference on Neural Networks (IJCNN), 2022.
3. Mingxuan Ju, **Shifu Hou**, Yujie Fan, Jianan Zhao, Yanfang Ye , Liang Zhao. "Adaptive Kernel Graph Neural Network", 36th AAAI Conference on Artificial Intelligence (AAAI), 2022.
4. **Shifu Hou**, Yujie Fan, Mingxuan Ju, Yanfang Ye , Wenqiang Wan, Kui Wang, Yinming Mei, Qi Xiong, Fudong Shao. 2021. Disentangled Representation Learning in Heterogeneous Information Network for Large-scale Android Malware Detection in the COVID-19 Era and Beyond. 35th AAAI Conference on Artificial Intelligence (AAAI), 2021.
5. Mingxuan Ju, Wei Song, Shiyu Sun, Yanfang Ye , Yujie Fan, **Shifu Hou**, Kenneth Loparo, Liang Zhao. "Dr.Emotion: Disentangled Representation Learning for Emotion Analysis on Social Media to Improve Community Resilience in the COVID-19 Era and Beyond", The Web Conference (WWW), 2021.
6. Yujie Fan, Mingxuan Ju, **Shifu Hou**, Yanfang Ye , Wenqiang Wan, Kui Wang, Yinming Mei, Qi Xiong. "Heterogeneous Temporal Graph Transformer: An Intelligent System for Evolving Android Malware Detection", International Conference on Knowledge Discovery and Data Mining (ACM SIGKDD), 2021.
7. Yanfang Ye , Yujie Fan, **Shifu Hou**, Yiming Zhang, Yiyue Qian, Shiyu Sun, Qian Peng, Mingxuan Ju, Wei Song, Kenneth Loparo. "a-Satellite: An AI-driven System and Benchmark Datasets for Dynamic COVID-19 Risk Assessment in the United States", IEEE Journal of Biomedical and Health Informatics (J-BHI), 2020.
8. Yanfang Ye , Yujie Fan, **Shifu Hou**, Yiming Zhang, Yiyue Qian, Shiyu Sun, Qian Peng, Mingxuan Ju, Wei Song, Kenneth Loparo. "Community Mitigation: A Data-driven System for COVID-19 Risk Assessment in a Hierarchical Manner", International Conference on Information and Knowledge Management (CIKM), 2020.
9. Yiming Zhang, Yujie Fan, **Shifu Hou**, Yanfang Ye, Xusheng Xiao, Pan Li, Chuan Shi, Liang Zhao, Shouhuai Xu. "Cyber-guided Deep Neural Network for Malicious Repository Detection in GitHub", IEEE International Conference on Knowledge Graph (ICKG), 2020.
10. Yanfang Ye, **Shifu Hou**, Lingwei Chen, Jingwei Lei, Wenqiang Wan, Jiabin Wang, Qi Xiong, Fudong Shao. "Out-of-sample Node Representation Learning for Heterogeneous Graph in Real-time Android Malware Detection", 28th International Joint Conference on Artificial Intelligence (IJCAI), 2019.
11. Yujie Fan, Yiming Zhang, **Shifu Hou**, Lingwei Chen, Yanfang Ye, Chuan Shi, Liang Zhao, Shouhuai Xu. "iDev: Enhancing Social Coding Security by Cross-platform User Identification Between GitHub and Stack Overflow", 28th International Joint Conference on Artificial Intelligence (IJCAI), 2019.

12. Yiming Zhang, Yujie Fan, Wei Song, **Shifu Hou**, Yanfang Ye, Xin Li, Liang Zhao, Chuan Shi, Jiabin Wang, Qi Xiong. "Your Style Your Identity: Leveraging Writing and Photography Styles for Drug Trafficker Identification in Darknet Markets over Attributed Heterogeneous Information Network", The Web Conference (WWW), 2019.
13. **Shifu Hou**, Yujie Fan, Yiming Zhang, Yanfang Ye, Jingwei Lei, Wenqiang Wan, Jiabin Wang, Qi Xiong and Fudong Shao. "aCyber: Enhancing Robustness of Android Malware Detection System against Adversarial Attacks on Heterogeneous Graph based Model", International Conference on Information and Knowledge Management (CIKM), 2019.
14. Lingwei Chen, **Shifu Hou**, Yanfang Ye, Thirimachos Bourlai, Shouhuai Xu, Liang Zhao. "iTrustSO: An Intelligent System for Automatic Detection of Insecure Code Snippets in Stack Overflow", Proceedings of International conference on Advances in Social Network Analysis and Mining (ASONAM), 2019.
15. Yujie Fan, **Shifu Hou**, Yiming Zhang, Yanfang Ye, Melih Abdulhayoglu. "Gotcha - Sly Malware! Scorpion: A Metagraph2vec Based Malware Detection System", Proceedings of ACM International Conference on Knowledge Discovery and Data Mining (ACM SIGKDD), 2018.
16. **Shifu Hou**, Yanfang Ye, Yangqiu Song, Melih Abdulhayoglu. "Make Evasion Harder: An Intelligent Android Malware Detection System", 27th International Joint Conference on Artificial Intelligence (IJCAI), 2018.
17. Yanfang Ye, **Shifu Hou**, Lingwei Chen, Xin Li, Liang Zhao, Shouhuai Xu, Jiabin Wang, Qi Xiong. "ICSD: An Automatic System for Insecure Code Snippet Detection in Stack Overflow over Heterogeneous Information Network", Annual Computer Security Applications Conference (ACSAC), 2018.
18. Lingwei Chen, **Shifu Hou**, Yanfang Ye, Shouhuai Xu. "DroidEye: Fortifying Security of Learning-based Classifier against Adversarial Android Malware Attacks", Proceedings of International conference on Advances in Social Network Analysis and Mining (ASONAM), 2018.
19. Yiming Zhang, Yujie Fan, **Shifu Hou**, Jian Liu, Yanfang Ye, Thirimachos Bourlai. "iDetector: Automate Underground Forum Analysis Based on Heterogeneous Information Network", Proceedings of International conference on Advances in Social Network Analysis and Mining (ASONAM), 2018.
20. **Shifu Hou**, Yanfang Ye, Yangqiu Song, Melih Abdulhayoglu. "HinDroid: An Intelligent Android Malware Detection System Based on Structured Heterogeneous Information Network", Proceedings of ACM International Conference on Knowledge Discovery and Data Mining (ACM SIGKDD) , 2017.
21. Lingwei Chen, **Shifu Hou**, Yanfang Ye. "SecureDroid: Enhancing Security of Machine Learning-based Detection against Adversarial Android Malware Attacks", Annual Computer Security Applications Conference (ACSAC), 2017.
22. Yanfang Ye, Lingwei Chen, **Shifu Hou**, William Hardy, Xin Li. "DeepAM: A Heterogeneous Deep Learning Framework for Intelligent Malware Detection", Knowledge and Information Systems (KAIS), Vol. PP (52): 1 21, 2017.
23. **Shifu Hou**, Aaron Saas, Lingwei Chen, Yanfang Ye, Thirimachos Bourlai. "Deep Neural Networks for Automatic Android Malware Detection", International conference on Advances in Social Network Analysis and Mining (ASONAM) , 2017.
24. **Shifu Hou**, Lingwei Chen, Yanfang Ye, Lifei Chen. "Deep Analysis and Utilization of Malware's Social Relation Network for its Detection", Proceedings of International Conference on Web-Age Information Management (WAIM) , 2017.
25. Lingwei Chen, **Shifu Hou**, Yanfang Ye, Lifei Chen. "An Adversarial Machine Learning Model against Android Malware Evasion Attacks", Proceedings of International Conference on Web-Age Information Management (WAIM) , 2017.
26. **Shifu Hou**, Aaron Saas, Yanfang Ye, Lifei Chen. "Deep4MalDroid: A Deep Learning Framework for Android Malware Detection Based on Linux Kernel System Call Graphs", IEEE/WIC/ACM International Conference on Web Intelligence Workshops (WIW), 2016.
27. **Shifu Hou**, Aaron Saas, Yanfang Ye, Lifei Chen. "DroidDeliver: An Android Malware Detection System Using Deep Belief Network Based on API Call Blocks", Proceedings of International Conference on Web-Age Information

Management(WAIM) , 2016.

28. William Hardy, Lingwei Chen, **Shifu Hou**, Yanfang Ye, Xin Li. "DL4MD: A Deep Learning Framework for Intelligent Malware Detection", International Conference on Data Mining (DMIN) , 2016.

29. **Shifu Hou**, Lifei Chen, Egemen Tas, Igor Demihovskiy, Yanfang Ye. "Cluster-Oriented Ensemble Classifiers for Malware Detection", IEEE International Conference on Semantic Computing (IEEE ICSC) , pp. 189 196, 2015.

## PROFESSIONAL SERVICES

---

### Conference Program Committee

- AAI, ECML, CIKM, WEB

### Journal Reviewer

- TDSC, TOPS, TETA, TAI

## SKILLS

---

Python, Pytorch, Huggingface, Deep Graph Library (DGL), C/C++