Ya (Grace) Xiao

Email: grace.xiao@bytedance.com **Web:** https://gracexy11.github.io/ **Tel:** (+1) 540-739-4136

EDUCATION

Virginia Tech, Blacksburg, VA — Aug 2017 – Jul 2022
Ph.D. in Computer Science, GPA: 3.94/4.0
Thesis Title: Neural Network-based Methodologies for Securing Cryptographic Code

Beijing University of Posts and Telecommunications (BUPT), Beijing, China — Sep 2010 - Jun 2017
M.S. in Information Security
B.S. in Accounting, minor in Information Security

EXPERIENCE

ByteDance, San Jose, California — Aug 2022 – Present
Machine Learning Engineer in TikTok Ads Integrity Team

- Build highly scalable machine learning systems for ads/business integrity, to improve TikTok user's experiences with ads across the platform.
- Collaborate with strategy team, product managers, policy team and other key stakeholders to define products and drive initiatives from engineering viewpoint.

Oracle Labs, Brisbane, Australia — Jun 2019 – Aug 2019
Research Internship in Program Analysis team — Director: Dr. Cristina Cifuentes

- **Static Analysis for Cryptographic Vulnerability Detection**
  - Develop a static dataflow analysis in Oracle bug checker Parfait for cryptographic vulnerability detection.
  - Implemented in C++ using LLVM.

Virginia Tech, Blacksburg, VA — Aug 2017 – Jul 2022
Graduate Research Assistant — Supervisor: Dr. Danfeng (Daphne) Yao

- **Program Analysis guided Code Embedding Techniques**
  - Design API embedding approaches with inter-procedural slicing and dataflow graph construction.
  - Compare NLP embedding approaches (word2vec, ELMo and BERT) for programming API embedding.
  - Implement in Python using Tensorflow.
- **Neural Network based API Completion for Securing Java Cryptographic APIs**
  - Design a multi-path based LSTM with an advanced low-frequency enhancing loss function for API completion.
  - Implemented in Python using Tensorflow.
- **Static Analysis for Cryptographic Vulnerability Detection in Java and Python**
  - Develop a high-precision, scalable detector for cryptographic API misuses in massive-sized projects.
  - Implemented in Java with Soot framework.
  - Implemented in Python with Python libraries Bandit, Astroid and RedBaron.
- **Measurement on Code Randomization Countermeasures under JIT-ROP Attacks**
  - Compare five code randomization tools (zipr, selfrando, CCR, Multicompiler, and Shuffler) against JIT-ROP attacks.
  - Evaluate the JIT-ROP gadget availability, quality, and their Turing-complete expressiveness.
- **A Neural Network based Approach for Black-box Cryptanalysis**
  - Design a neural network based approach to evaluate the security of a cipher in the black-box manner.
  - Implemented in Python with Tensorflow.
- **Deep Learning-Based Anomaly Detection in Cyber-Physical Systems**
  - Develop a LSTM based sequence model to identify anomalies of CPS systems.
  - Implemented in Python with using Tensorflow and Keras.
- **High-accuracy Insider Threat Detection and Reasoning with Probabilistic Evidence**
  - Build probabilistic models to discover anomalies from huge high dimensional data.
  - Implemented in Python with Pyro.
- **Data Sampling Techniques for Machine Learning with Imbalanced Dataset**
  - Experiment with several data imbalance solutions, including oversampling, subsampling, and weighted loss function, on MIMIC-III clinic dataset.
  - Implemented in Python with Tensorflow and Keras.

Graduate Teaching Assistant

- CS4264 Principles of Computer Security (Fall 2017) — Instructor: Dr. Matthew Hicks

TECHNICAL SKILLS

- **Applied Machine Learning:** Word Embedding, Natural Language Modeling, Clustering
- **Security:** Secure Coding, Vulnerability Detection, Automatic Repair, Insider Threat Detection, Anomaly Detection
- **Engineering**: Python, Java, C++, Soot, LLVM, Tensorflow, Kera, Pyro, Java Cryptography Architecture, Spring Security

PUBLICATIONS

[TOSEM'23] **Ya Xiao**, Wenjia Song, Salman Ahmed, Xinyang Ge, Bimal Viswanath, Na Meng, and Danfeng (Daphne) Yao. "Measurement of Embedding Choices on Cryptographic API Completion Tasks". ACM Transactions on Software Engineering and Methodology, 2023.

[TSE'23] **Ya Xiao**, Wenjia Song, Jingyuan Qi, Bimal Viswanath, Patrick McDaniel, Danfeng (Daphne) Yao. "Specializing Neural Networks for Cryptographic Code Completion Applications". IEEE Transactions on Software Engineering, 2023.

[FSE'21] **Ya Xiao**. "Multi-location Cryptographic Code Repair with Neural-Network-Based Methodologies", Doctoral Symposium of ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE), August 2021.

[CCS'20] Salman Ahmed, **Ya Xiao**, Kevin Z. Snow, Gang Tan, Fabian Monrose, and Danfeng (Daphne) Yao. "Methodologies for Quantifying (Re-)randomization Security and Timing under JIT-ROP", ACM SIGSAC Conference on Computer and Communications Security (CCS), Virtual Conference, November 2020.

[CCS'19] Sazzadur Rahaman, **Ya Xiao**, Sharmin Afrose, Fahad Shaon, Ke Tian, Miles Frantz, Murat Kantarcioglu, and Danfeng (Daphne) Yao. "Cryptoguard: High precision detection of cryptographic vulnerabilities in massive-sized java projects", ACM SIGSAC Conference on Computer and Communications Security (CCS), London, UK, November 2019.

[DTRAP'22] **Ya Xiao**, Yang Zhao, Nicholas Allen, Nathan Keynes, Danfeng (Daphne) Yao and Cristina Cifuentes. "Industrial Experience of Finding Cryptographic Vulnerabilities in Large-scale Codebases". ACM Digital Threats: Research and Practice (DTRAP), 2022

[IDSC'19] **Ya Xiao,** Qingying Hao, Danfeng (Daphne) Yao. "Neural Cryptanalysis: Metrics, Methodology, and Applications in CPS Ciphers", IEEE Conference on Dependable and Secure Computing (IDSC), Hangzhou, China, November 2019.

[SecDev'22] **Ya Xiao,** Yang Zhao, Nicholas Allen, Nathan Keynes, Danfeng (Daphne) Yao, and Cristina Cifuentes.

"Industrial Strength Static Detection for Cryptographic API Misuses". In Proceedings of the IEEE Secure Development Conference (SecDev). Practitioners' Session. Atlanta, GA. Oct. 2022.

[TSE'22] Sharmin Afrose, **Ya Xiao**, Sazzadur Rahaman, Miller Barton. Danfeng (Daphne) Yao. "Development of Benchmarks for Java Cryptographic APIs and Evaluation of Static Vulnerability Detection Tools". IEEE Transactions on Software Engineering, 2022.

[ICPC'22] Ying Zhang, **Ya Xiao**, Md Mahir Asef Kabir, Danfeng (Daphne) Yao, Na Meng. "Example-based Vulnerability Detection and Repair in Java Code". IEEE/ACM International Conference on Program Comprehension (ICPC), 2022.

[CSET'19] Xiaodong Yu, **Ya Xiao**, Danfeng (Daphne) Yao and Kirk Cameron. "Comparative Measurement of Cache Configurations Impacts on Cache Timing Side-Channel Attacks", The 12th USENIX Workshop on Cyber Security Experimentation and Test (CSET), Santa Clara, CA, August 2019.

[TSE'22] Ying Zhang, Md Mahir Asef Kabir, **Ya Xiao**, Danfeng (Daphne) Yao, Na Meng. "Automatic Detection of Java Cryptographic API Misuses: Are We There Yet?" IEEE Transactions on Software Engineering (TSE), 2022.

[CSUR'20] Yuan Luo, **Ya Xiao**, Long Cheng, Guojun Peng, and Danfeng (Daphne) Yao. "Deep Learning-Based Anomaly Detection in Cyber-Physical Systems: Progress and Opportunities", ACM Computing Surveys (CSUR), 2020.

MAGAZINES

[IEEE S&P'22] Danfeng (Daphne) Yao, Sazzadur Rahaman, **Ya Xiao**, Sharmin Afrose, Miles Frantz, Ke Tian, Na Meng, Cristina Cifuentes, Yang Zhao, Nicholas Allen, Nathan Keynes, Barton P. Miller, Elisa R. Heymann, Murat Kantarcioglu, and Fahad Shaon. "Being the Developers' Friend: Our Experience Developing a High-Precision Tool for Secure Coding."

IEEE Security & Privacy. 2022.

[IEEE Computer'22] Salman Ahmed, **Ya Xiao**, Taejoong (Tijay) Chung, Carol Fung, Moti Yung, and Danfeng (Daphne) Yao. "Privacy Guarantees of BLE Contact Tracing: A Case Study on COVIDWISE." IEEE Computer. February 2022.

TUTORIALS

[ESORICS'21] **Ya Xiao**, Miles Frantz, Sharmin Afrose and Danfeng (Daphne) Yao "Tutorial: Principles and Practices of Secure Cryptographic Coding in Java" (90 minutes Tutorial), European Symposium on Research in Computer Security (ESORICS), September, 2021.

[SecDev'20] **Ya Xiao**, Miles Frantz, Sharmin Afrose, Sazzadur Rahaman and Danfeng (Daphne) Yao. "Tutorial: Principles and Practices of Secure Cryptographic Coding in Java" (90 minutes Tutorial), IEEE Secure Development Conference (SecDev). September, 2020.