Ya (Grace) Xiao

Email: yax99@vt.edu **Web:** https://people.cs.vt.edu/~yax99/ **Tel:** (+1) 540-739-4136

TECHNICAL SKILLS

Security: Secure Coding, Vulnerability Detection, Automatic Repair, Insider Threat Detection, Anomaly Detection

Engineering: Python, Java, C++, Soot, LLVM, Tensorflow, Keras,Pyro, Java Cryptography Architecture, Spring Security

Applied Machine Learning: Word Embedding, Natural Language Modeling, Clustering

EDUCATION

Virginia Tech, Blacksburg, VA Aug 2017 - Present
Ph.D. candidate in Computer Science, GPA: 3.94/4.0
Thesis Title: Neural Network based Methodologies and Comparisons for Securing Cryptography 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

Oracle Labs, Brisbane, Australia Summer 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 - Present
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.
- **An Empirical Study for Existing Java Security API Misuse Detection Tools**
  - Compare five security API screening tools (CogniCrypt, CryptoGuard, FindSecBugs, SonarQube, Xanitizer) on three security vulnerability benchmarks (CryptoBench, MUBench, OWASP).
  - Interact with developers through pull requests for fixing vulnerabilities.
- **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 an LSTM based sequence model to identify anomalies of CPS systems.
  - Implemented in Python using Tensorflow and Keras.
- **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

PUBLICATIONS

[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.

[arxiv'21] **Ya Xiao**, Salman Ahmed, Wenjia Song, Xinyang Ge, Bimal Viswanath, Danfeng (Daphne) Yao. "Embedding Code Contexts for Cryptographic API Suggestion: New Methodologies and Comparisons".

[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.

[arXiv'20] **Ya Xiao**, Yang Zhao, Nicholas Allen, Nathan Keynes, Danfeng (Daphne) Yao and Cristina Cifuentes. "Industrial Experience of Finding Cryptographic Vulnerabilities in Large-scale Codebases".

[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.

[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.

[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.

[arxiv'21] Ying Zhang, Mahir Kabir, **Ya Xiao**, Danfeng (Daphne) Yao, Na Meng. "Data-Driven Vulnerability Detection and Repair in Java Code".

[Under Review] 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". (Submitted to Transactions on Software Engineering)

[Under Review] Ying Zhang, Mahir Kabir, **Ya Xiao**, Danfeng (Daphne) Yao and Na Meng. "Automatically Detecting Security-API Misuses in Java Programs: Are We There Yet?". (Submitted to Transactions on Software Engineering)

[Under Submission] Miles Frantz, **Ya Xiao**, Tanmoy Pias and Danfeng (Daphne) Yao. "Detection and Benchmark for Python Cryptographic Misuses".

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.

PATENT

[Under Review] Danfeng (Daphne) Yao, Salman Ahmed, **Ya Xiao**. "High-accuracy Insider Threat Detection and Reasoning with Probabilistic Evidence". (Submitted)