

APPENDIX

PRIMARY STUDIES

- P1 J. Walden, J. Stuckman and R. Scandariato, "Predicting vulnerable components: Software metrics vs text mining," *International Symposium on Software Reliability Engineering*, 2014.
- P2 Y. Zhou, S. Liu, J. Siow, X. Du and Y. Liu, "Devign: Effective vulnerability identification by learning comprehensive program semantics via graph neural networks," *Conference on Neural Information Processing Systems*, 2019.
- P3 J. Tian, W. Xing and Z. Li, "BVDetector: A program slice-based binary code vulnerability intelligent detection system," *Information and Software Technology*, 2020.
- P4 C. Theisen and L. Williams, "Better together: Comparing vulnerability prediction models," *Information and Software Technology*, 2020.
- P5 P. Morrison, K. Herzig, B. Murphy and L. Williams, "Challenges with applying vulnerability prediction models," *Symposium and Bootcamp on the Science of Security*, 2015.
- P6 A. Fidalgo, I. Medeiros, P. Antunes and N. Neves, "Towards a Deep Learning Model for Vulnerability Detection on Web Application Variants," *International Conference on Software Testing, Verification and Validation Workshops*, 2020.
- P7 X. Ban, S. Liu, C. Chen and C. Chua, "A performance evaluation of deep-learned features for software vulnerability detection," *Concurrency and Computation: Practice and Experience*, 2019.
- P8 Z. Li, D. Zou, J. Tang, Z. Zhang, M. Sun and H. Jin, "A comparative study of deep learning-based vulnerability detection system," *IEEE Access*, 2019.
- P9 T. Nguyen, T. Le, K. Nguyen, O. de Vel, P. Montague, J. Grundy and D. Phung, "Deep Cost-Sensitive Kernel Machine for Binary Software Vulnerability Detection," *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 2020.
- P10 G. Lin, J. Zhang, W. Luo, L. Pan, O. D. Vel, P. Montague and Y. Xiang, "Software Vulnerability Discovery via Learning Multi-domain Knowledge Bases," *IEEE Transactions on Dependable and Secure Computing*, 2019.
- P11 M. Gegick, L. Williams, J. Osborne and M. Vouk, "Prioritizing software security fortification through code-level metrics," *ACM Workshop on Quality of protection*, 2008.
- P12 M. Jimenez, M. Papadakis and Y. L. Traon, "Vulnerability prediction models: A case study on the Linux Kernel," *International Working Conference on Source Code Analysis and Manipulation*, 2016.
- P13 S. Neuhaus, T. Zimmermann, C. Holler and A. Zeller, "Predicting vulnerable software components," *ACM Conference on Computer and Communications Security*, 2007.
- P14 S. Liu, G. Lin, Q. Han, S. Wen, J. Zhang and Y. Xiang, "DeepBalance: Deep-Learning and Fuzzy Oversampling for Vulnerability Detection," *IEEE Transactions on Fuzzy Systems*, 2020.
- P15 M. Zagane, M. K. Abdi and M. Alenezi, "Deep Learning for Software Vulnerabilities Detection Using Code Metrics," *IEEE Access*, 2020.
- P16 H. Perl, S. Dechand, M. Smith, D. Arp, F. Yamaguchi, K. Rieck, S. Fahl and Y. Acar, "VCCFinder: Finding potential vulnerabilities in open-source projects to assist code audits," *ACM SIGSAC Conference on Computer and Communications Security*, 2015.
- P17 Y. Joonseok, R. Duksan and B. Jongmoon, "Improving vulnerability prediction accuracy with Secure Coding Standard violation measures," *International Conference on Big Data and Smart Computing (BigComp)*, 2016.
- P18 V. Nguyen, T. Le, T. Le, K. Nguyen, O. De Vel, P. Montague, L. Qu and D. Phung, "Deep Domain Adaptation for Vulnerable Code Function Identification," *International Joint Conference on Neural Networks*, 2019.
- P19 I. Chowdhury and M. Zulkernine, "Using complexity, coupling, and cohesion metrics as early indicators of vulnerabilities," *Journal of Systems Architecture*, 2011.
- P20 H. K. Dam, T. Tran, T. Pham, S. W. Ng, J. Grundy and A. Ghose, "Automatic Feature Learning for Predicting Vulnerable Software Components," *IEEE Transactions on Software Engineering*, 2018.
- P21 A. Meneely and L. Williams, "Strengthening the empirical analysis of the relationship between Linus' Law and software security," *International Symposium on Empirical Software Engineering and Measurement*, 2010.
- P22 X. Chen, Y. Zhao, Z. Cui, G. Meng, Y. Liu and Z. Wang, "Large-Scale Empirical Studies on Effort-Aware Security Vulnerability Prediction Methods," *IEEE Transactions on Reliability*, 2020.
- P23 X. Cheng, H. Wang, J. Hua, M. Zhang, G. Xu, L. Yi and Y. Sui, "Static detection of control-flow-related vulnerabilities using graph embedding," *International Conference on Engineering of Complex Computer Systems*, 2019.
- P24 V. H. Nguyen and L. M. S. Tran, "Predicting vulnerable software components with dependency graphs," *International Workshop on Security Measurements and Metrics*, 2010.
- P25 L. K. Shar and H. B. K. Tan, "Predicting SQL injection and cross site scripting vulnerabilities through mining input sanitization patterns," *Information and Software Technology*, 2013.
- P26 S. Liu, G. Lin, L. Qu, J. Zhang, O. D. Vel, P. Montague and Y. Xiang, "CD-VulD: Cross-Domain Vulnerability Discovery based on Deep Domain Adaptation," *IEEE Transactions on Dependable and Secure Computing*, 2020.
- P27 B. M. Padmanabhuni and H. B. K. Tan, "Buffer Overflow Vulnerability Prediction from x86 Executables Using Static Analysis and Machine Learning," *Annual Computer Software and Applications Conference*, 2015.
- P28 W. Zheng, J. Gao, X. Wu, F. Liu, Y. Xun, G. Liu and X. Chen, "The impact factors on the performance of machine learning-based vulnerability detection: A comparative study," *Journal of Systems and Software*, 2020.
- P29 H. Wang, G. Ye, Z. Tang, S. H. Tan, S. Huang, D. Fang, Y. Feng, L. Bian and Z. Wang, "Combining Graph-Based Learning With Automated Data Collection for Code Vulnerability Detection," *IEEE Transactions on Information Forensics and Security*, 2021.
- P30 Y. Shin, A. Meneely, L. Williams and J. A. Osborne, "Evaluating complexity, code churn, and developer activity metrics as indicators of software vulnerabilities," *IEEE Transactions on Software Engineering*, 2011.
- P31 G. Lin, J. Zhang, W. Luo, L. Pan, Y. Xiang, O. De Vel and P. Montague, "Cross-Project Transfer Representation Learning for Vulnerable Function Discovery," *IEEE Transactions on Industrial Informatics*, 2018.
- P32 Z. Li, D. Zou, S. Xu, X. Ou, H. Jin, S. Wang, Z. Deng and Y. Zhong, "Vuldeepecker: A deep learning-based system for vulnerability detection," *Usenix Network and Distributed System Security Symposium*, 2018.
- P33 Y. Shin and L. Williams, "Can traditional fault prediction models be used for vulnerability prediction?," *Empirical Software Engineering*, 2013.
- P34 N. Saccente, J. Dehlinger, L. Deng, S. Chakraborty and Y. Xiong, "Project achilles: A prototype tool for static method-level vulnerability detection of Java source code using a recurrent neural network," *International Conference on Automated Software Engineering Workshop*, 2019.
- P35 Y. Zhang, D. Lo, X. Xia, B. Xu, J. Sun and S. Li, "Combining Software Metrics and Text Features for Vulnerable File Prediction," *International Conference on Engineering of Complex Computer Systems*, 2016.
- P36 M. Jimenez, R. Rwemalika, M. Papadakis, F. Sarro, Y. Le Traon and M. Harman, "The importance of accounting for real-world labeling when predicting software vulnerabilities," *ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 2019.
- P37 S. Moshtari, A. Sami and M. Azimi, "Using complexity metrics to improve software security," *Computer Fraud, Security*, 2013.
- P38 D. Zou, S. Wang, S. Xu, Z. Li and H. Jin, "μVulDeePecker: A Deep Learning-Based System for Multiclass Vulnerability Detection," *IEEE Transactions on Dependable and Secure Computing*, 2019.
- P39 S. Ghaffarian and H. Shahriari, "Neural software vulnerability analysis using rich intermediate graph representations of programs," *Information Sciences*, 2021.
- P40 M. Siavvas, D. Tsoukalas, M. Jankovic, D. Kehagias and D. Tzouvaras, "Technical debt as an indicator of software security risk: a machine learning approach for software development enterprises," *Enterprise Information Systems*, 2020.
- P41 I. Abunadi and M. Alenezi, "An empirical investigation of security vulnerabilities within web applications," *Journal of Universal Computer Science*, 2016.
- P42 L. Yang, X. Li and Y. Yu, "VulDigger: A Just-in-Time and Cost-Aware Tool for Digging Vulnerability-Contributing Changes," *IEEE Global Communications Conference*, 2017.
- P43 V. Nguyen, T. Le, O. de Vel, P. Montague, J. Grundy and D. Phung, "Dual-Component Deep Domain Adaptation: A New Approach for Cross Project Software Vulnerability Detection," *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 2020.
- P44 Z. Li, D. Zou, S. Xu, H. Jin, Y. Zhu and Z. Chen, "SySeVR: A Framework for Using Deep Learning to Detect Software Vulnerabilities,"

- IEEE Transactions on Dependable and Secure Computing*, 2021.
- P45 S. Liu, M. Dibaei, Y. Tai, C. Chen, J. Zhang and Y. Xiang, "Cyber Vulnerability Intelligence for Internet of Things Binary," *IEEE Transactions on Industrial Informatics*, 2020.
- P46 R. Scandariato, J. Walden, A. Hovsepian and W. Joosen, "Predicting vulnerable software components via text mining," *IEEE Transactions on Software Engineering*, 2014.
- P47 N. Medeiros, N. Ivaki, P. Costa and M. Vieira, "Software metrics as indicators of security vulnerabilities," *Software Reliability Engineering*, 2017.
- P48 M. A. Albahar, "A Modified Maximal Divergence Sequential Auto-Encoder and Time Delay Neural Network Models for Vulnerable Binary Codes Detection," *IEEE Access*, 2020.
- P49 X. Duan, J. Wu, S. Ji, Z. Rui, T. Luo, M. Yang and Y. Wu, "Vul-sniper: Focus your attention to shoot fine-grained vulnerabilities," *International Joint Conference on Artificial Intelligence*, 2019.
- P50 Y. Fang, S. Han, C. Huang and R. Wu, "TAP: A static analysis model for PHP vulnerabilities based on token and deep learning technology," *PloS one*, 2019.
- P51 D. Cao, J. Huang, X. Zhang and X. Liu, "FTCLNet: Convolutional LSTM with Fourier Transform for Vulnerability Detection," *Trust, Security and Privacy in Computing and Communications*, 2020.
- P52 Z. Yu, C. Theisen, L. Williams and T. Menzies, "Improving Vulnerability Inspection Efficiency Using Active Learning," *IEEE Transactions on Software Engineering*, 2019.
- P53 A. Figueiredo, T. Lide, D. Matos and M. Correia, "MERLIN: Multi-Language Web Vulnerability Detection," *International Symposium on Network Computing and Applications*, 2020.
- P54 Z. Bilgin, M. A. Ersoy, E. U. Soykan, E. Tomur, P. Comak and L. Karacay, "Vulnerability Prediction from Source Code Using Machine Learning," *IEEE Access*, 2020.
- P55 X. Chen, Z. Yuan, Z. Cui, D. Zhang and X. Ju, "Empirical studies on the impact of filter-based ranking feature selection on security vulnerability prediction," *IET Software*, 2020.
- P56 J. Stuckman, J. Walden and R. Scandariato, "The Effect of Dimensionality Reduction on Software Vulnerability Prediction Models," *IEEE Transactions on Reliability*, 2017.
- P57 Y. Tang, F. Zhao, Y. Yang, H. Lu, Y. Zhou and B. Xu, "Predicting Vulnerable Components via Text Mining or Software Metrics? An Effort-Aware Perspective," *IEEE International Conference on Software Quality, Reliability and Security*, 2015.
- P58 J. Domanska, M. Siavvas and E. Gelenbe, "Efficient Feature Selection for Static Analysis Vulnerability Prediction," *K. Filus, P. Boryszko, Sensors*, 2021.
- P59 Y. Shin and L. Williams, "An initial study on the use of execution complexity metrics as indicators of software vulnerabilities," *International Workshop on Software Engineering for Secure Systems*, 2011.
- P60 N. Medeiros, N. Ivaki, P. Costa and M. Vieira, "Vulnerable Code Detection Using Software Metrics and Machine Learning," *IEEE Access*, 2020.
- P61 R. Li, C. Feng, X. Zhang and C. Tang, "A Lightweight Assisted Vulnerability Discovery Method Using Deep Neural Networks," *IEEE Access*, 2019.