Weilin Xu

University of Virginia
Department of Computer Science
85 Engineer's Way
Charlottesville, VA 22904

Phone: (434) 466-7432

Email: xuweilin@virginia.edu
Homepage: https://xuweilin.org

Research Interest

Adversarial Machine Learning
I'm interested in creating robust machine learning-based classifiers. My research has developed a generic method for generating adversarial examples using genetic programming, and a general technique named *Feature Squeezing* to harden deep learning models by eliminating unnecessary features.

Education

Ph.D. in Computer Science, University of Virginia (UVa) May 2019 (expected)
Co-advisors: Prof. David Evans and Prof. Yanjun (Jane) Qi

B.E. in Computer Science, Beijing University of Posts and Telecommunications (BUPT) June 2012

Research Experience

Research Assistant, Security Research Group and Machine Learning Group, UVa 2014.5 – present
I have been the lead PhD student of two projects: *Genetic Evasion* which attacks state-of-the-art malware classifiers; and *Feature Squeezing* which detects adversarial examples. Details at https://EvadeML.org.

Engineer, Network and Information Security Lab, Tsinghua University 2012.7 – 2013.8
I led a team constructing a nationwide honeynet system.

Student Developer, The Honeynet Project, Google Summer of Code 2012 2012.5 – 2012.8
I created 6Guard, an IPv6 attack detector that was later widely deployed on CNGI-CERNET2.

IPv6 Expert, Nmap Security Scanner, Google Summer of Code 2011 2011.5 – 2011.8
I created the efficient IPv6 host discovery techniques and improved the IPv6 support of Nmap.

Lead Student, Student Innovation Lab, BUPT 2010.4 – 2011.4
I led a team developing an embedded IPv6/IPv4 Dual Stack NAT router based on OpenWrt.

Research Assistant, State Key Lab of Networking and Switching, BUPT 2009.10 – 2010.1
I developed a generic file system encryption feature on the Linux Kernel VFS layer.

Publications

Reviewed Proceedings

Weilin Xu, David Evans, Yanjun Qi. Feature Squeezing: Detecting Adversarial Examples in Deep Neural Networks. To appear in *the* 25^{th} *Network and Distributed System Security Symposium (***NDSS'18***)*, San Diego, Feburary 2018. (*acceptance rate 21.5%*)

Weilin Xu, Yanjun Qi, David Evans. Automatically Evading Classifiers: A Case Study on PDF Malware Classifiers. In *the* 23^{rd} *Network and Distributed System Security Symposium (***NDSS'16***)*, San Diego, Feburary 2016. (*acceptance rate 15.4%*)

Other Publications

Qixue Xiao, Kang Li, Deyue Zhang, **Weilin Xu**. Security Risks in Deep Learning Implementations. *arXiv preprint arXiv:1711.11008* 2017.

Weilin Xu, David Evans, Yanjun Qi. Feature Squeezing Mitigates and Detects Carlini/Wagner Adversarial Examples. *arXiv preprint arXiv:1705.10686* 2017.

Weilin Xu, Andrew Norton, Noah Kim, Yanjun Qi, David Evans. Poster: EvadeML-Zoo: A Benchmarking and Visualization Tool for Adversarial Machine Learning. USENIX Security 2017.

Weilin Xu, Yanjun Qi, David Evans. Poster: Automatically Evading Classifiers. IEEE S&P 2015.

Research Community Service

External reviewer for IEEE S&P (Oakland) 2015, 2017, USENIX Security 2016, ACM AsiaCCS 2016, NDSS 2017, ACM CCS 2017.

PC member for NIPS Machine Learning and Computer Security (MLSec) Workshop 2017.

Invited Talks

Attack and Defense in Adversarial Machine Learning
09/12/2017: Tsinghua University, Beijing, China.
09/13/2017: Internet Security Conference, Beijing, China.
09/14/2017: Beijing University of Posts and Telecommunications, Beijing, China.
09/15/2017: Baidu, Inc., Beijing, China.
09/18/2017: ShanghaiTech University, Shanghai, China.
09/27/2017: SangFor, Inc., Shenzhen, China.

Automatically Evading Classifiers
03/31/2016: Beijing University of Posts and Telecommunications, Beijing, China.

Teaching Experience

TA, CS6316: Machine Learning, Prof. Yanjun Qi, University of Virginia, Fall 2016.
I assisted Prof. Qi in customizing the Tensorflow Playground as an interactive tool for teaching machine learning concepts.

TA, CS4414: Operating Systems, Prof. David Evans, University of Virginia, Fall 2013 and Spring 2014.
I assisted Prof. David Evans in developing an undergraduate operating system course (focus on system programming), which is the first course to use the Rust programming language in the world.

References

Professor David Evans, Ph.D. evans@virginia.edu
Department of Computer Science, University of Virginia
Charlottesville, VA 22904

Assistant Professor Yanjun (Jane) Qi, Ph.D. yanjun@virginia.edu
Department of Computer Science, University of Virginia
Charlottesville, VA 22904