Carnegie Mellon University
CyLab

Student Seminar Series

Modeling Security Weaknesses to Enable Practical Runtime Defenses

Friday, May 24th 2019
12:00pm – 1:00pm
CIC 2201

Speaker: William Melicher

William Melicher is a PhD candidate from Carnegie Mellon University where he is advised by Lujo Bauer. William has broad interest in security and privacy research, and has worked on projects on usable security, online privacy, security applications of machine learning, and web security. He has received several awards, including two best paper awards for work on passwords, at USENIX Security and CHI, and the IEEE Cybersecurity Award for Practice. During his doctoral work, William spent two summers working at Google on the identity and privacy teams.

Talk Abstract:

Security weaknesses are often caused by patterns in human behaviors. However, it can be difficult to identify such patterns in a practical, yet accurate way. In order to fix security weaknesses, it is crucial to identify and detect them. Useful systems to model security weaknesses must be accurate enough to guide users' decisions, but also be lightweight enough to produce results in a reasonable time frame. In this thesis, we show how machine learning techniques allow us to detect security weaknesses that result from patterns in human behavior faster and more efficiently than current approaches, enabling new, practical run-time defenses. We present two applications to support this thesis.

First, we use neural networks to identify users' weak passwords and show how to make such models practical for fully client-side password feedback. One problem with current password feedback is that users can get either quick but substantially incorrect feedback by using heuristics that have little relation to password strength, or accurate but slow feedback by simulating adversarial guessing using large models. In contrast, we found that our models of password guessing are both more accurate and smaller than previous ones, which enables us to more practically estimate resistance to password-guessing attacks in real time on client machines.

Second, we use deep learning models to identify client-side cross-site scripting vulnerabilities in JavaScript code. We collected JavaScript functions from hundreds of thousands of web pages and using a taint-tracking-enabled browser labeled them according whether they were vulnerable to cross-site scripting. We trained deep neural networks to classify source code as safe or as potentially vulnerable. We demonstrate how our models can be used as a lightweight building block to selectively enable other defenses.

This is a practice talk for William's PhD thesis defense.

Please direct any questions regarding this seminar to Sekar Kulandaivel (skulanda@andrew.cmu.edu) or Shunsuke Aoki (shunsuka@andrew.cmu.edu).