Collaborative Research: SaTC: CORE: Medium: Defending Against Social Engineering Attacks with In-Browser AI

Roberto Perdisci – Lead PI (perdisci@uga.edu)

Nick Nikiforakis – PI (nick@cs.stonybrook.edu)

Phani Vadrevu – PI (phani@cs.uno.edu)

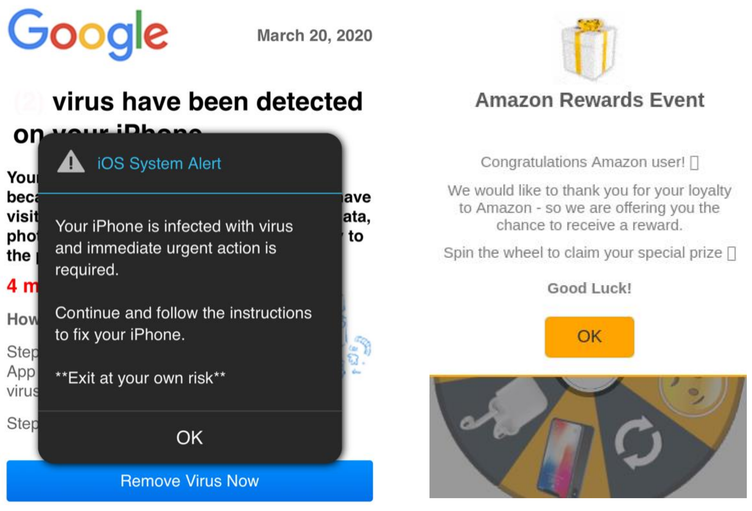


NSF Grants No. 2126641, 2126654, 2126655

Challenge

- *Social engineering* (SE) attacks exploit humans' decision-making process.
- They use a large variety of deception and persuasion tactics.
- They are at the basis of many web-based scams, phishing attacks, and malware distribution.
- They cause significant monetary loss.
- Existing defenses (e.g., URL blockers) are mostly ineffective.

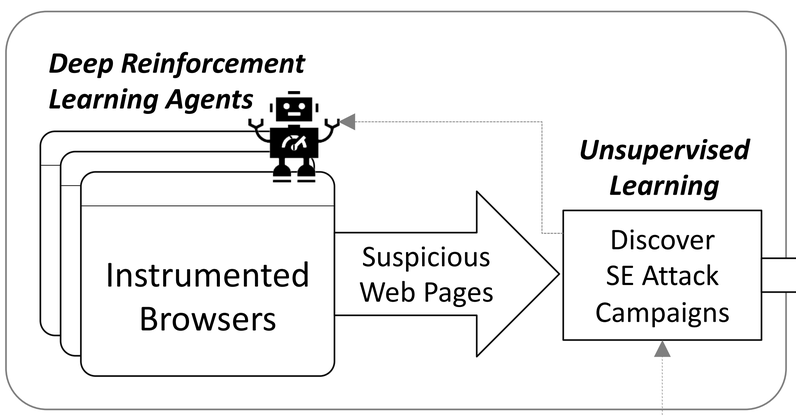
SE attack examples



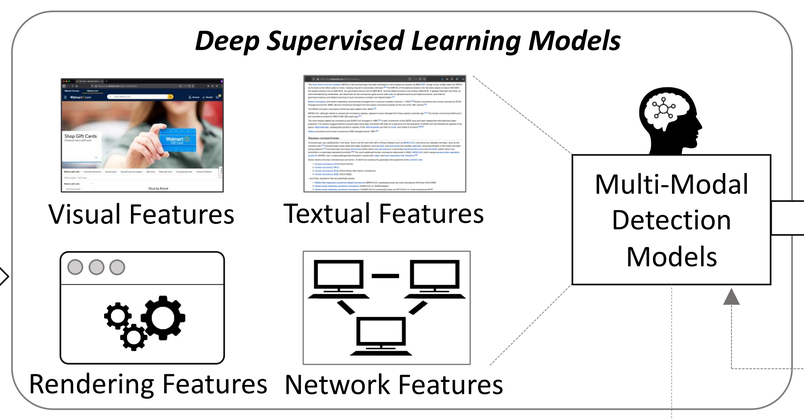
Technical Solution + Impact

- **SE Scout**: automatically discover and collect datasets of attack examples.
- **SE Learn**: Learn how to detect future SE attacks using multiple ML models.
- **SE Shield**: Deploy efficient ML models within the browser.
- ***Privacy***: ML models will run locally, without sharing users' browsing data.
- ***Impact***: Block SE attacks encountered by users while browsing.

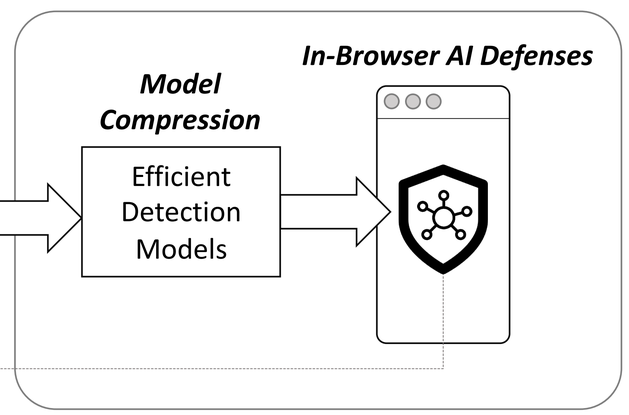
SE Scout

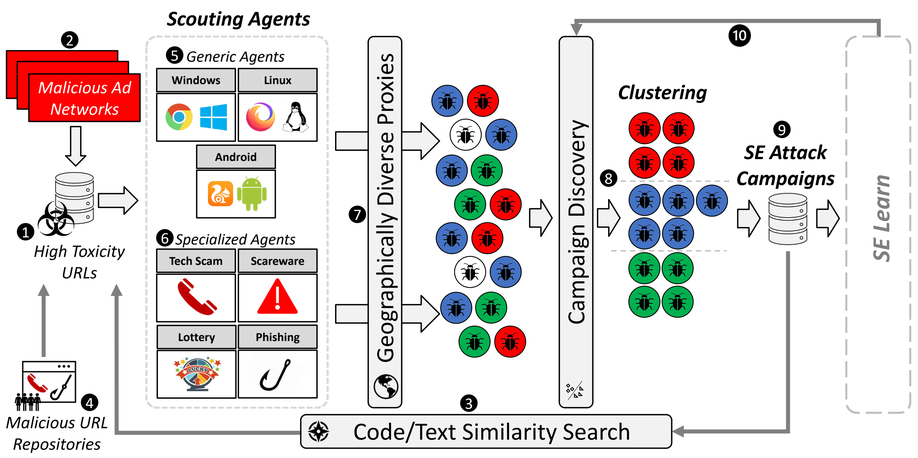


SE Learn

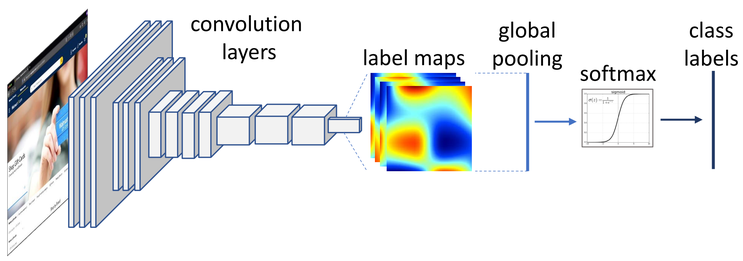


SE Shield





Fully-Convolutional Network



Broader Impact

- **Educating senior citizens**: Senior citizens are one of the main targets of SE attacks. We plan to work with AARP and other organizations to educate senior citizens on how to identify and defend themselves against web-based scams.