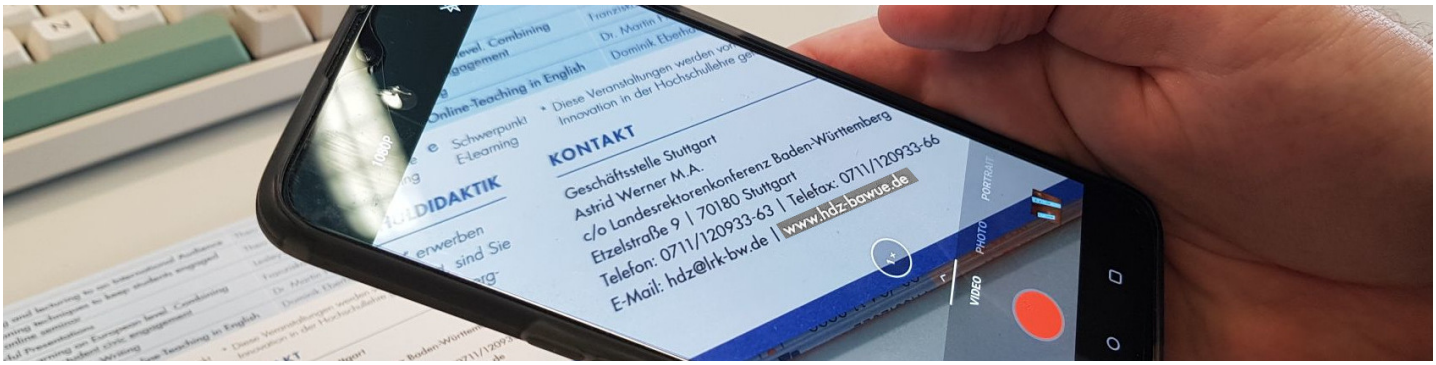
Karlsruhe Institute of Technology

Research Group Intelligent System Security (ISEC)
Institute of Information Security and Dependability
Karlsruhe Institute for Technology

Research Group Security · Usability · Society (SECUSO)
Institute of Applied Informatics and Formal Description Methods
Karlsruhe Institute for Technology

Model Extraction and Social Engineering Attacks on Apple LiveText

Bachelor's Thesis Proposal



Background

Optical Character Recognition (OCR) are tools that convert images of typed, handwritten or printed text in other formats. OCRs are quite common, nowadays, especially for blind and visually impaired users (e.g., text-to-speech tools). 'LiveText' is an OCR implemented in Apple iPhones and iPads mounting Apple A12 Bionic chips and successors, i.e., 2018+ iPhones and 2019+ iPads. On top of recognizing characters, LiveText highlights hyperlinks and enables the users to directly open websites by pointing the camera to printouts of the domain. As all OCRs, LiveText makes mistakes and misinterprets characters. This motivates a new real world attack vector, e.g., for phishing attacks, to take users to malicious versions of legitimate websites. However, this problem is difficult to investigate, due to the closed source nature of the LiveText system. Hence, in the first step, a surrogate model needs to be extracted. Next, the student runs adversarial attacks on it and verifies the transferability of the attacks on the original LiveText system. Finally, she develops a social engineering attack that takes advantage of the flaws and evaluates it in a user study.
The goal of this multi-disciplinary thesis is two-folds:

- Review literature on model extraction attacks and conduct a model extraction on Live Text to construct a surrogate model. Implement transferable attacks against this model. Verify the transferability of the attack on the original Live Text system quantitatively.
- Review literature on phishing attacks. Identify, motivate and evaluate social engineering attacks, scenarios and attacker models by taking advantage of the flaws in the Live Text recognition model (e.g., homographic attacks, typo-attacks, pixel-based attacks, etc.). Verify in a user study the scenarios' applicability, qualitatively and quantitatively.

Requirements

- Basic understanding of Machine Learning and Apple programming
- Interest in Social Engineering and Phishing Attacks
- Interest in multi-disciplinary research
- Language: English

Related work to start with

1. Heise.de, 24.01.2022: *Mehr Fehler mit QR-Codes, auch bei OnePlus - Google bietet Pixel-Update*
2. Heise.de, 20.01.2022: *Googles Kamera verfaelscht Links in QR-Codes*
3. CSO Germany, 13.04.2021: *7 new social engineering tactics threat actors are using now*
4. Trend Micro, 09.02.2022: *Hidden Scams in Malicious Scans: How to Use QR Codes Safely*

Supervisors & Contact

- Maximilian Noppel, M.Sc.,, ISEC, maximilian.noppel@kit.edu
- Mattia Mossano, M.Sc. M.A., SECUSO, mattia.mossano@kit.edu