Hessisches Kompetenzzentrum
für Hochleistungsrechnen

Impact of Multipath-Effects on Acoustic Keylogging Systems

Project Manager
Vincenz Mechler

Researchers
Matthias Gazzari

Principal Investigator
Prof. Dr.-Ing. Matthias Hollick

Project Term
2022 - 2023

Clusters
Lichtenberg II Cluster Darmstadt

Software
PyTorch

Additional Software
pyroomacoustics

Institute
Secure Mobile Networking Lab

University
Technische Universität Darmstadt

Introduction

Acoustic keylogging describes a category of side-channel attacks that recover typed keystrokes from audio-streams of nearby microphones. We identify two fundamentally different approaches, namely feature-based key identification utilizing unique acoustic signatures, and time-difference-of-arrival-based spatial sound source localization. Proposed acoustic keystroke inference systems are usually evaluated under lab conditions, and results can not be transferred to real-world scenarios. While the evaluation scenarios mostly lack environmental noise, unpredictable multipath-effects, or combined keystrokes, all these are present in acoustic keystroke signals captured ’in the wild’. Furthermore, there are no common datasets, and results are hardly comparable due to different input and output formats. Related works also do not release their implementations, and often provide only incomplete specifications. To address these problems, we designed and implemented a simulation-based acoustic keylogging evaluation framework in the context of a master’s thesis, capable of generating equivalent input data for all acoustic keylogging methods. We collected an extensive dataset and evaluated the impact of multipath-effects on several acoustic keylogging systems.

Methods

In order to simulate multipath-effects typical to the environment of acoustic keylogging, we first generate high-quality synthetic acoustic keystroke signals from a clean dataset of individual keystrokes recorded under anechoic conditions. Next, we inject these samples into customizable, virtual office environments to

add controlled amounts of realistic multipath-effects. This simulation-based approach also enables the generation of arbitrary synced audio channels at known spatial locations. This allows us to satisfy the varying input requirements of different acoustic keylogging systems. We implement four acoustic keylogging models from the literature, namely:
• Two similar nearest-neighbor and clustering models using frequency distance and
cross correlation
• A convolutional neural network adapted from the speech recognition domain
• A time-difference-of-arrival-based localization model
In addition to a clean dataset without multipath-effects, we generate ten different virtual scenes with distinct room impulse responses, and propagate our samples through these environments via acoustic simulation. We then evaluate the impact of multipath effects on the recognition accuracy of these models. We train our models on data with and without multipath-effects to evaluate the impact of known vs. unknown acoustic channels on the keystroke identification accuracy. Furthermore, we explore the possibility of compensating for unknown acoustic channels by training on several unrelated scenes, from which the keylogging models might learn to transfer to the unknown test scene. The training and/or inference of these models is computationally expensive. Therefore, our evaluations were enabled by the access to the Lichtenberg HPC, which allowed for statistically significant sample sizes along all relevant dimensions.

Results

While the multipath effects of unknown scenes cause reduced accuracy across all featurebased models, the relative decrease in identification accuracy of approx. 33% maximum is less than initially expected. This indicates that models trained in one setup can be transferred to another and still produce usable results, contrary to what is often portrayed in related work. Furthermore, multipath effects of known scenes effectively enrich the features of each individual key due to its unique location and acoustic channel to the microphone. This means that in real-world scenarios the models perform up to 30% better than under anechoic conditions. Lastly, adding multiple unrelated scenes to the training data can compensate for the lack of knowledge about the test-scenes acoustic channels, depending on the model. While the nearest-neighbor-based model showed only a slight increase in performance, the convolutional neural network was able to fully compensate the unknown acoustic properties of the test-scene. The clustering-based model, on the other hand, did not show any improvement, which is to be expected from the model semantics. The localization model could not be properly evaluated due to stability problems.

Discussion

Our work explored the previously seldom regarded dimension of multipath effects in acoustic keylogging samples and its effects

on the various types of keylogging models. We showed that acoustic keylogging still poses a threat if the environment changes (e.g. by the victim unwittingly moving the microphone or changing location with a notebook), whereas previous work usually expects training data from the exact, undisturbed environment that is to be attacked.

Last Update: 2023-06-07 08:52