Project in computer security

Meeting 3

Suggested missions from previous meeting

1. Investigating the features in Skype&type project.
2. Mapping the exact algorithms that Skype&type uses.
3. Collecting data from 3 more people. (We want 5 sets of recordings)
4. Record tests (Regular computer typing, words, passwords etc.)
5. Preparing a full demo:
    a. Train a model for all the collected data.
    b. Guess the Recorded tests.
    c. Estimate the accuracy of the model.
6. Suggest a new Definition Of Done for the project.

Investigating the features in Skype&type project

S&T design has a modular approach and is divided to 4 main parts:

- Listener - Gets the wav files and passes it to the Dispatcher.
- Dispatcher - Extracts keystrokes and features.
- Model - The trained classifier.
- Output - Handles the output.

Algorithms in S&T

Default learning algorithm: **Logistic Regression**. Which as we saw from the papers should give the best results for the problem.

Potentially you can use any algorithm in the **sklearn** library (it's a modular implementation). But it will require some development to make it work because every algorithm requires different parameters.

Collecting good data is HARD

- To collect large amounts of good training data we should **work manually.**
- We can go up to **5000 samples** if we really want to, **is it enough?**
- Our demos and estimations have trained on **approximately 2000 data samples.**
- We have 3 target recordings:
  - “hello world hello world”
  - “308561562”
  - “one ring to rule them all one ring to find them one ring to bring them all and in the darkness bind them”

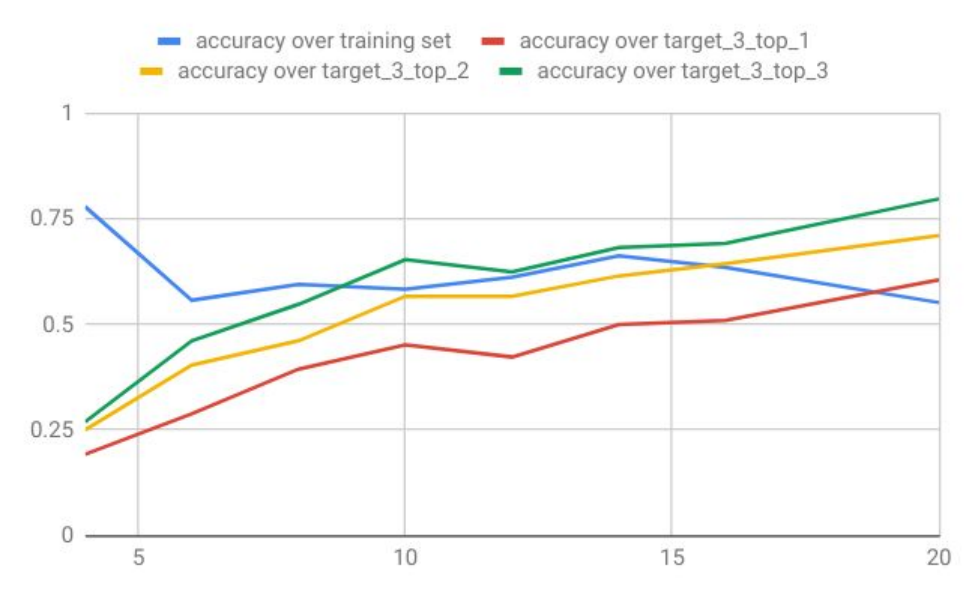
More data ?= Accuracy

DEMO

More data ?= Accuracy

More data ?= Accuracy

More data ?= Accuracy



Future plans proposal

Our conclusions: **Collecting good data is HARD!, and More data == Accuracy!**

We want to make the data collection process more robust and easy!

We will achieve this goal with:

- GUI that manages all the process in one program.
- Hint the program for errors
- Interactive data collection

The user interface

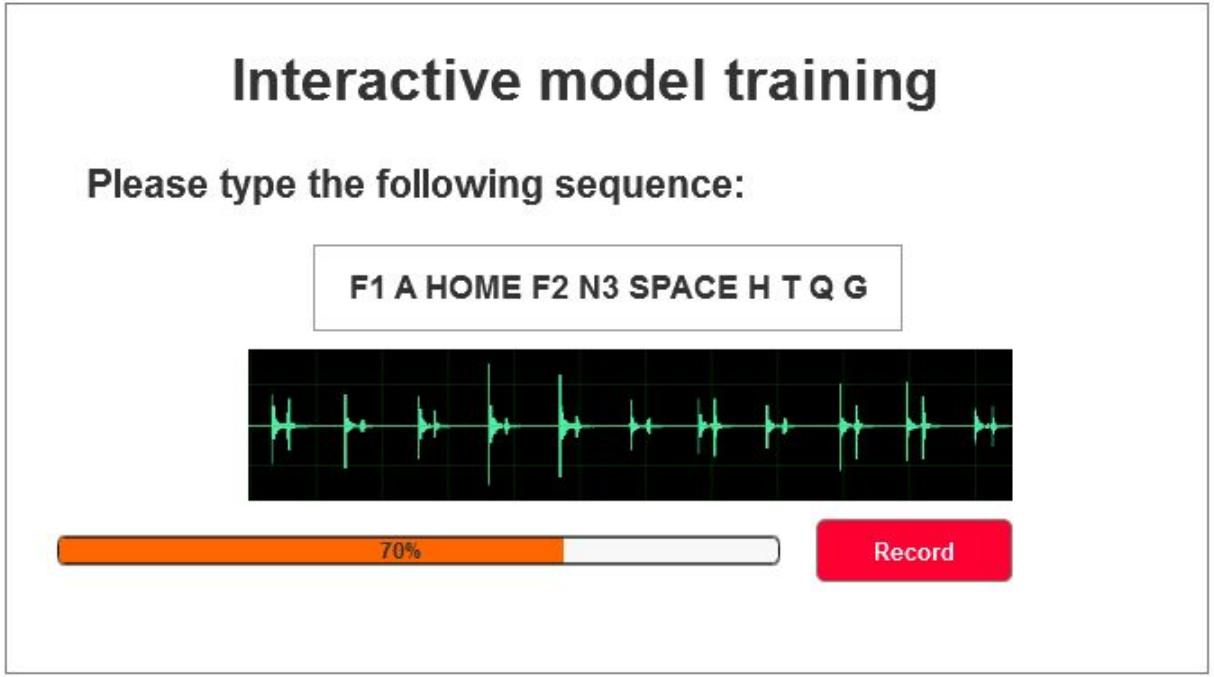
Acoustic Keylogger

Name	Accuracy	Options
K260	★★★★★	Classify Update Delete
Dell Inspiron 5373	★★★★☆	Classify Update Delete
MacBook Pro 2017	★★★☆☆	Classify Update Delete
Dell XPS 15 2018	★★☆☆☆	Classify Update Delete

Train new model interactively

Train new model from files

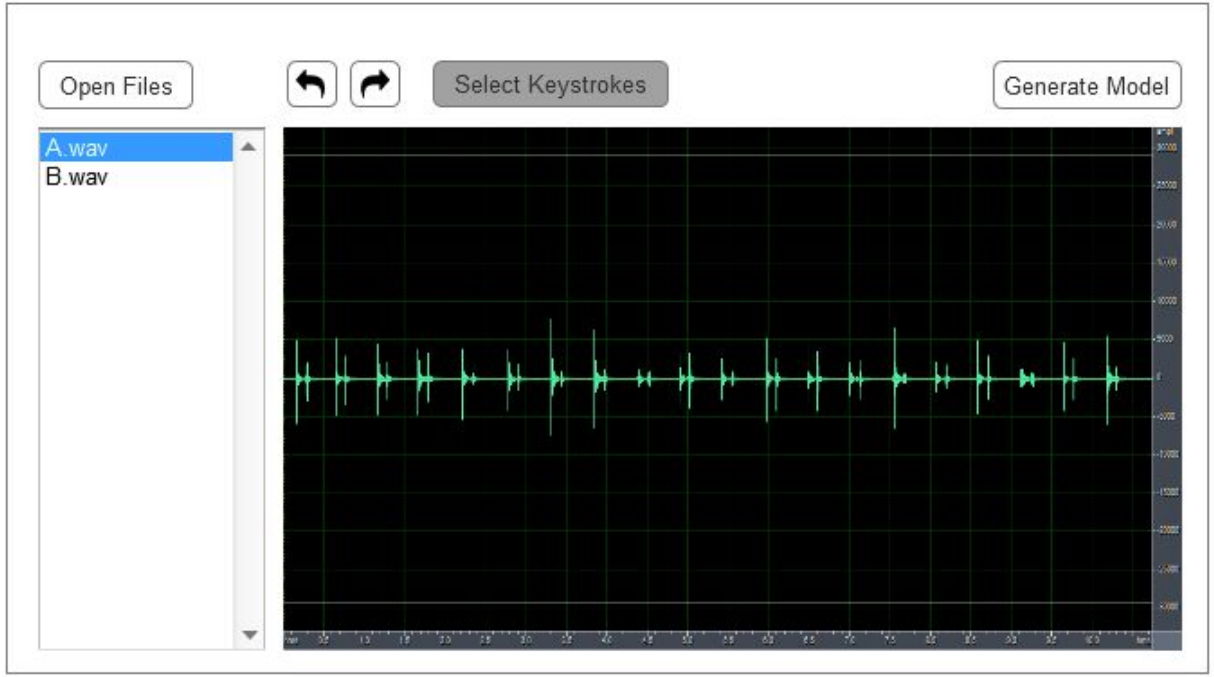
The user interface



The user interface

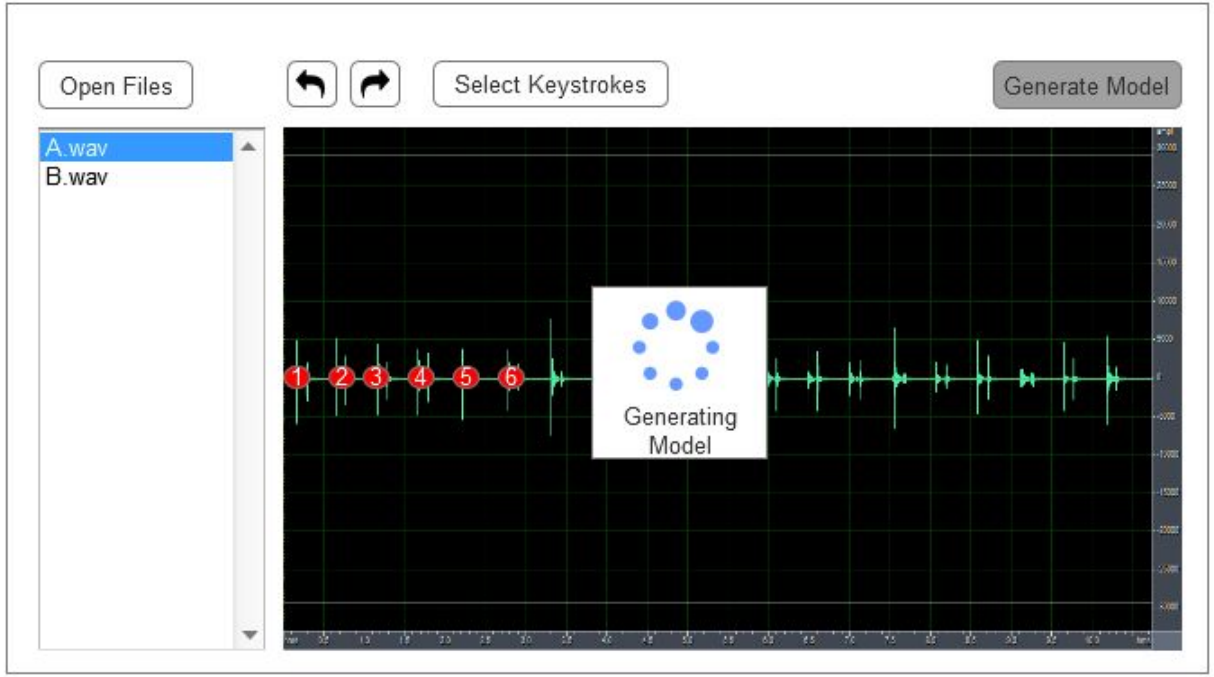
The user interface

The user interface



The user interface

The user interface



The user interface