Assessing the Human Ability to Recognize Synthetic Speech

Daniel Prudký, xprudk08@stud.fit.vutbr.cz

Supervisor: Ing. Anton Firc

Motivation

Artificial intelligence is associated with threats of all kinds, and deepfakes are one of them. There are many attack scenarios in which deepfakes have been used. People should be able to defend themselves against this spread of fraudulent information and media, they should know how to recognise or verify such things and how to deal with them. However, we do not know whether we will ever be able to do this.

Goals

The aim of this work is to determine whether people are able to recognize synthetic speech.

Experiment

In experiment were respondents confronted with voice deepfakes without being told about them and observe whether they recognize it as a deepfake or at least notice something strange.
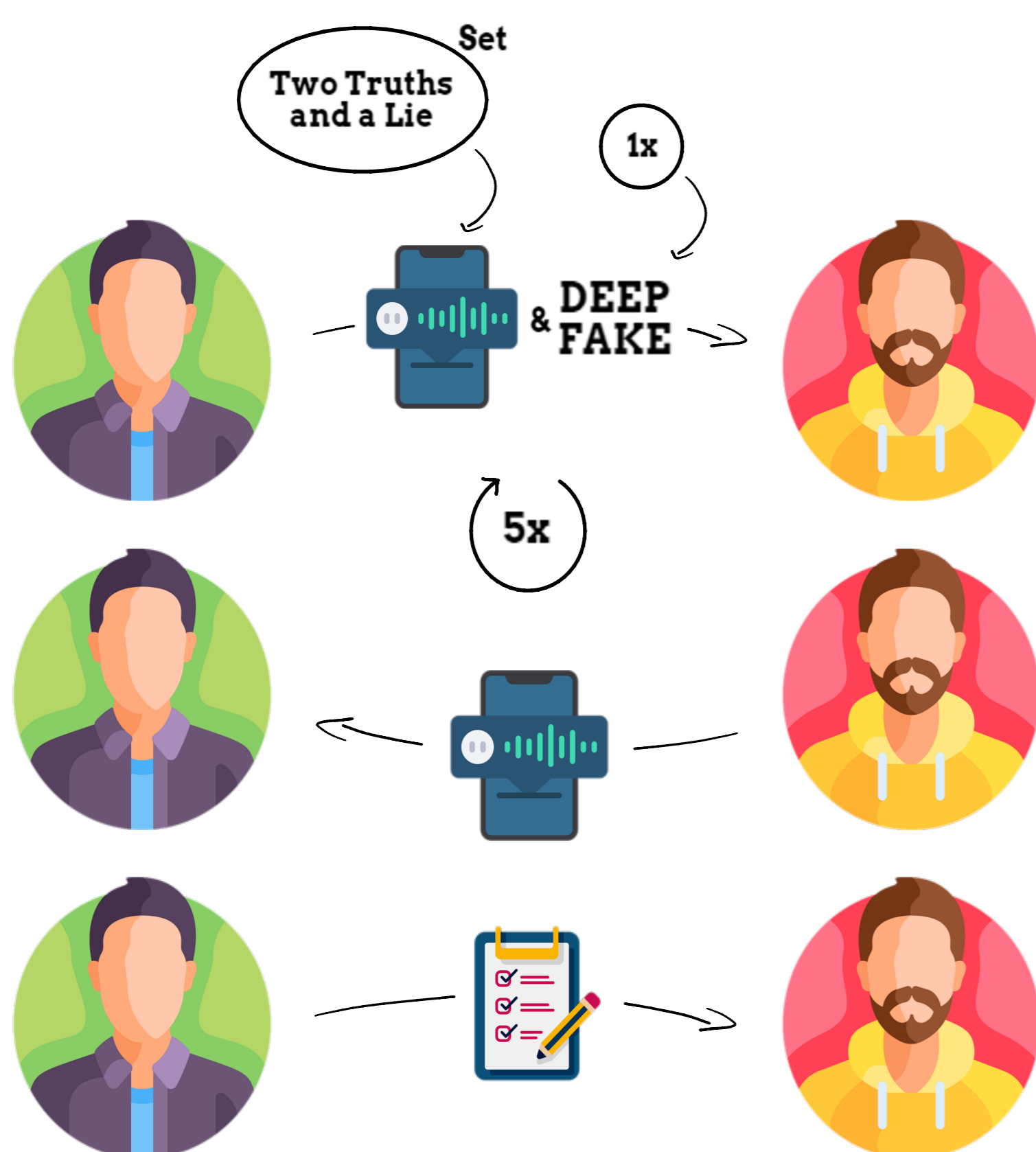

Figure 1: Experiment diagram

Experiment was hid behind a cover story and under this curtain, we asked the participants simple questions using voice messages in a classic chat. In between the real voice messages, we also sent one pre-prepared deepfake recording of my voice and then watched to see if the participants recognized the deepfake or at least noticed anything odd.

Results

The results of the experiment are quite clear.

During the conversation, none of the respondents reacted in any way to the fraudulent deepfake message. And only one mentioned the message afterwards as something strange in the conversation.
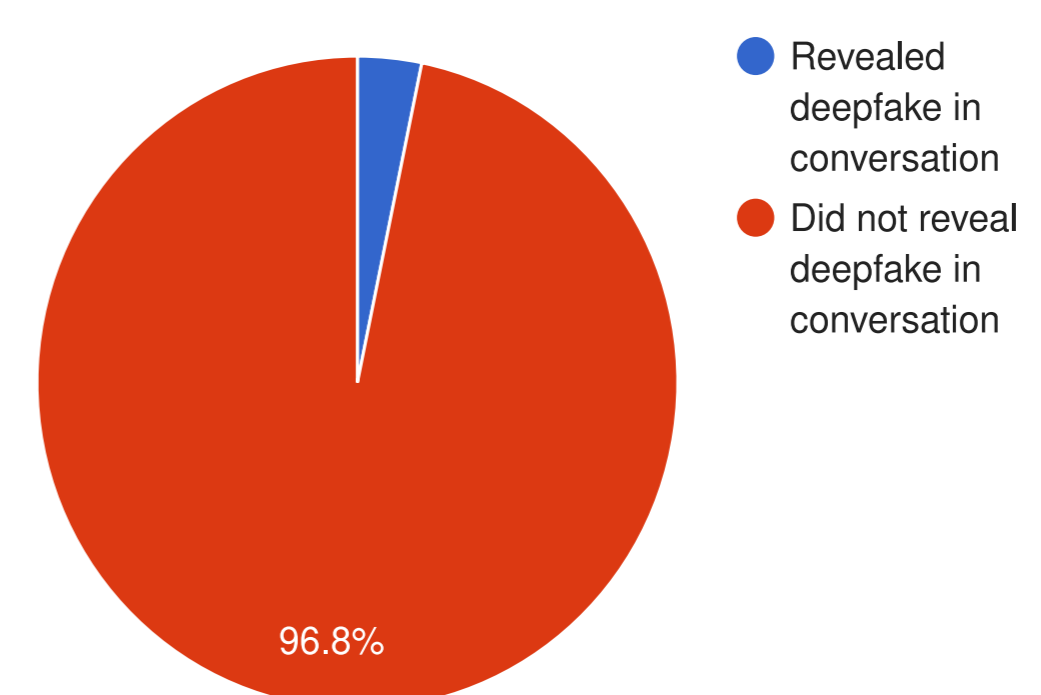

Figure 2: Graph of deepfake message detection in conversation

On the other hand, after the experiment was revealed, in the subsequent identification of the deepfake message, only that specific message was correctly marked by 83.9% of the respondents.
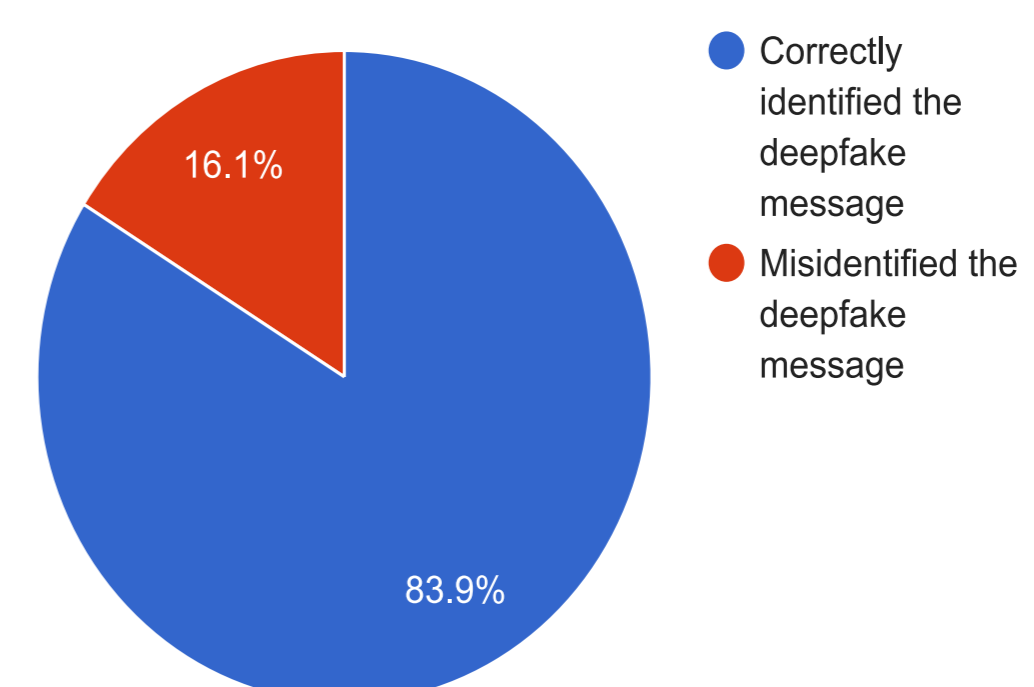

Figure 3: Graph of accuracy of identifying deepfake message among others

The experiment found that 83.9% of respondents had at least heard of deepfakes, primarily from social media, educational videos or simply stumbling across them online.
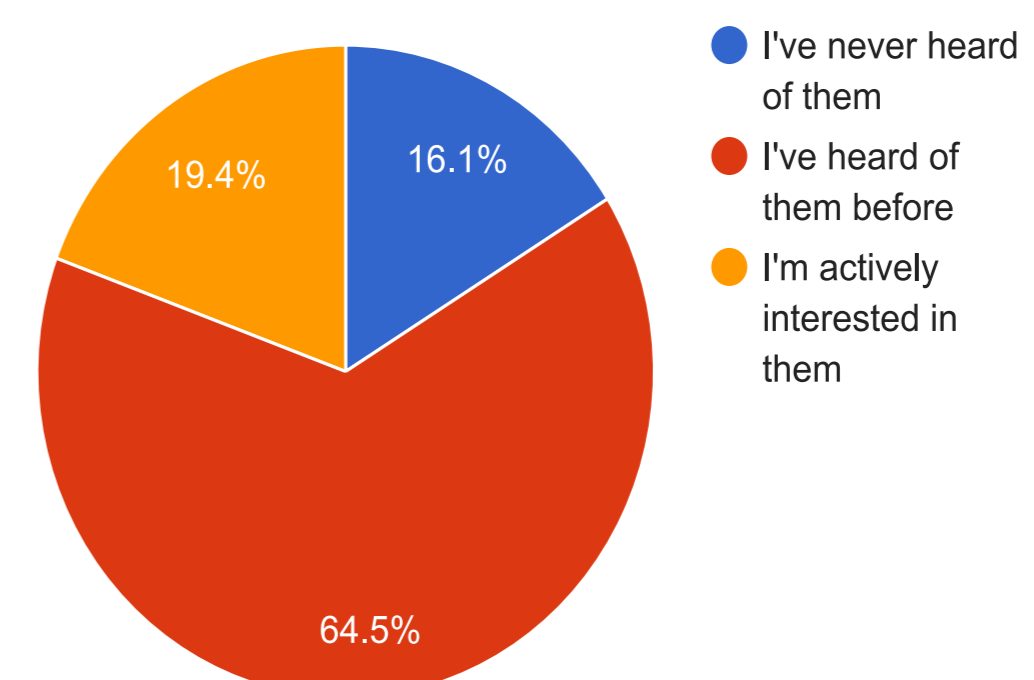

Figure 4: Graph of people's awareness of deepfakes

WORK HAS SHOWN THAT THE HUMAN ABILITY TO RECOGNIZE VOICE DEEPFAKES IS NOT AT A LEVEL WE CAN TRUST. IT IS VERY DIFFICULT FOR PEOPLE TO DISTINGUISH BETWEEN REAL AND FAKE VOICES, ESPECIALLY IF THEY ARE NOT EXPECTING THEM.