Tool for forensic analyses of digital traces

Author: Mária Hatalová, Advisor: Václav Matyáš
Faculty of Informatics, Masaryk University

Motivation

The thesis studied forensic analysis and its use by expert witnesses that work for the Police of the Czech Republic. The goal of the thesis was to address the needs of expert witnesses by simplifying and speeding up their work and reducing the amount of manual work that needs to be done for solving forensic tasks assigned by the police.

Typical forensic task and forensic tools

- We described a typical forensic task as it is assigned by the police. The expert witness is given copies of data carriers seized throughout house search and his task is to find all files that contain specific keywords. The expert witness has to create an expert testimony – a pdf document - containing the files with the keyword occurrences together with their metadata.
- We analyzed available tools that could be used for processing typical forensic tasks. The tool Autopsy seemed to be a good candidate, however, when it was tested on a real case, it turned out as not suitable for this purpose due to multiple issues that occurred.
- Due to not having found a suitable tool the rest of the thesis was devoted to the IT Forensic Tool (ITFT), that was being developed by the technical consultant of the thesis.

IT Forensic Tool v2.0

- As the part of the thesis, we created a new version of ITFT - IT Forensic Tool v2.0. We discussed the requirements for further functionality that should be added to ITFT and proposed and implemented a solution to each of the requirements.
- We integrated a PostgreSQL database with the aim to simplify the way how the tool stores the information about forensic tasks.
- We speeded up the keyword search over the database using suitable indexes (Figure 1).
- To be able to automatically extract and analyze text from images, Optical Character Recognition (OCR) was incorporated into ITFT. Out of several OCR tools we picked Tesseract, which was extracting text from images most accurately (Figure 2).
- For a number of most common file formats, we managed to automate the pdf report generation.

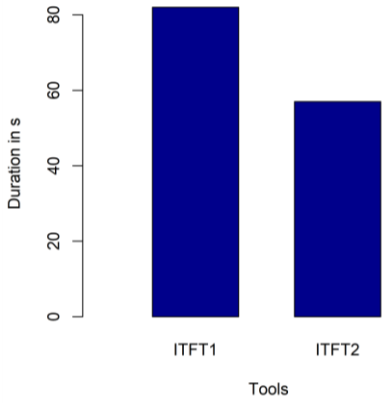

Figure 1: Duration of keywords export

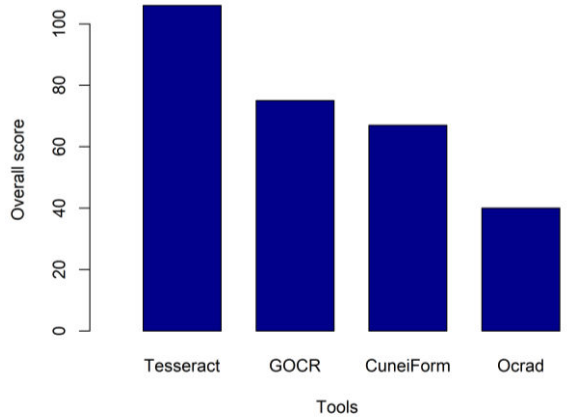

Figure 2: Output accuracy score

Conclusions and future work

- In the thesis, we managed to simplify and speed up the work of expert witnesses that work for the Police of the Czech Republic. We also reduced the amount of manual work required for solving typical forensic tasks.
- The IT Forensic Tool v2.0 was successfully tested on a real case assigned by the Police of the Czech Republic.
- We proposed future improvements for the ITFT tool, such as rewriting the tool in an object-oriented language, creating means to handle unexpected interruptions, and more.