Research Report

Entropass

Jochem Stevense

A research report presented for the design & implementation of Entropass

Embedded Systems Engineering
Flexible Project
Hogeschool van Arnhem en Nijmegen
Ton Ammerlaan, Remko Welling
The Netherlands
2021
Version 0.1

Research Report

Entropass

Jochem Stevense

Abstract

put some text here

Contents

1 Introduction

2 Research

2.1 Research Plan

To be able to create a program that fulfils the requirements, a main research question has been formulated and divided into several sub-questions.

The main research question is the following:

What functionalities should a password guessing tool have to be able to utilise standard password practices in guessing the password of a specific user?

To be able to answer the main research question, it is split into sub-questions. These sub-questions are the following:

1. *What are standard password formats?*
2. *What personal information is relevant for password guessing?*
3. *How can the processing of the information be done efficiently, from a technical perspective?*
4. *How can the program remain up to date with changing password practices?*

The combination of the sub-questions will be used to answer the main question and to help design the program.

2.2 Research Methodology

The research methodology will deal with the methods, used to answer the sub-questions and the main question, formulated in chapter 2.1. Firstly, the sub-questions will be handled, after which the main research question will be dealt with.

1. *What are standard password formats?*
   To answer this question, a desk research will be conducted, using online resources, such as leaked data/passwords, to create a list of formats, as used by users in the breach. The formats will be sorted in a list of most used to least used. This will be done by a simple Python script, which will be written specifically for the use of this research. The script will be made available for interested parties with the final project.

2. *What personal information is relevant for password guessing?*
   Passwords are often based on personal information to make them easier to memorise. This personal information should be mapped to categories, to determine what a person is most likely to use for a memorable password. This will be done by desk research. This desk research will involve the use of word lists from data breaches and analysing these manually to determine what categories are popular amongst users. Categories might involve aspects like, relationships, date of birth, pet names, etc.

3. *How can the processing of the information be done efficiently, from a technical perspective?*
   To be able to answer this question, a test setup will be created, where the setup will be used to iterate random information, using different design techniques. This process will be timed to determine which of the design techniques should be used for the program.

4. *How can the program remain up to date with changing password practices?* This question will be answered by looking into sources of information for password practices, after which possibilities will be researched for gathering this information and analysing it in an automatised way.

2.3 Research Results

The results of the research will be handled per question, firstly dealing with the sub-questions. The results are listed in this paragraph. The conclusions following this research, will be dealt with in the Conclusions chapter.

1. *What are standard password formats?*
   The first step to answering the sub-question "*What are standard password formats?*", was to find useful data to process. For this reason, an online search pointed to the freely available data on Github, called SecLists. A link to this data can be found in the appendices (4.1). This data contains information about frequently used directory names, usernames, passwords and much more. In this research, only the usernames and passwords were relevant.

2. *What personal information is relevant for password guessing?*

3. *How can the processing of the information be done efficiently, from a technical perspective?*

4. *How can the program remain up to date with changing password practices?*

3 Conclusions

3.1 Further Development

4 Appendices

4.1 Appendix A: SecLists

SecLists is a collection of data which is freely available on Github. It contains passwords, usernames and much more interesting data which can be used to gain insight in the most vulnerable passwords

4.2 Appendix B: Python password format processor

A Python 3 based script was written to simplify the processing of password data.

```
print('hello world')
```