EntroPass

Project Plan

A project plan, created for the EntroPass project.

Jochem Stevense
591323

Flexible Project
Hogeschool van Arnhem en Nijmegen
The Netherlands
2021
Version 0.1

Contents

1 Introduction

This project plan has been created for the Hogeschool van Arnhem en Nijmegen as part of the Flexible Project.

Cyber security is an increasingly important discipline in the professional world, with its weakest link being human error and negligence. The use of weak, easy to guess passwords is still the most convenient way for an attacker to gain access to a device. Currently, there is a great number of tools to use to guess passwords. However, these tools make use of passwords that have been used often and are not very valuable if an attacker is trying to target a specific user, where personal information might allow for faster results.

This project is being conducted by a single Embedded Systems Engineering student and is meant to lay the foundation for the development of a open-sourced password guessing program, using personal information, designed to be used for security tests and to raise awareness to the importance of strong randomised passwords. The project will be evaluated by Ton Ammerlaan and Remko Welling.

The program is being designed to produce word lists with possible passwords for a specific person. These produced word lists are intended to be used to demonstrate how long it would take for a program to guess a specific persons password, while being used on a "regular" machine. The program aims at reducing the computational power needed by regular *Dictionary Attacks*, where the attacker simply tries all passwords in a word lists, based on most used passwords.

The competences to be developed during this project, include the ability to document code clearly and document the workings of the program for both developers and non-developers to understand the program.

The project has the following criteria to be evaluated on:
Documentation

- The documentation should include a to do list and/or advise chapter to allow future developers to continue the coding of the project.
- The documentation should include guidelines for other programmers to follow, in order to maintain uniform styling in the code.
- The code should include comments to explain itself.

Functionality

- The program should be able to guess passwords, based on personal information.
- The program should be able to output these passwords into a word list.
- The user should be able to see the amount of time it took generate the word list.
- The user should be able to input a password for comparison, to see how long if takes for the program to guess this password.
- The program should be prepared to be expanded for the use of guessing usernames, based on personal information.

The stakeholders for this project are the following:

- Ethical hackers and Security researchers.
- Companies, looking to test their internal security.
- Personal users.
- Hogeschool van Arnhem en Nijmegen.
- The research team.

The next chapters provide the project plan for the personal password guessing project, hereafter named the EntroPass project.

The first chapter will go over the problem, specifying the existing base for the project and the challenge it means to solve after the completion of this project. After this, the project will be analysed to formulate the goals and sub-goals and to create a base to start researching. The next chapter is the research plan, continuing from the project analysis and formulating research questions based on the goals and sub-goals, while also providing a justification for the research. Finally, the limitations and conditions will be dealt with.

2 Problem orientation

This project is being developed following a number of encounters with the use of poor passwords, based on personal information, making them easy to guess for people in possession of this personal information, combined with knowledge on often used password formats. After some research, it appeared that there are no tools available to demonstrate the problem with these passwords.

Companies often spend a lot of time and effort in mitigating vulnerabilities in their systems, while not sufficiently addressing the problem of weak passwords, based on personal information. Passwords based on personal information, can easily be guessed by an attacker, since most of this information is freely available online, for example on social networks.

The reason why this password guessing tool will be created, is not to aid attackers in malicious attacks, but to allow researchers, penetration testers, companies and even personal users to prepare for these type of attacks.

3 Project Analysis

The goal of the project is to create a program that automates the password guessing process, using personal information, provided by the "attacker".

This is a different approach than is traditionally used to gain access to a person account. The most used attacks use tool to perform Brute Force attacks or Dictionary attacks. Brute Force attacks use processing power to generate all different possible combinations possible and uses these to try to login.

A Dictionary attack uses a list of words to try and login. Both of these attacks use a lot of preocessing power and are not normally specified to be used for a specific person. These attacks will try to gain access to a account, while monitoring the amount of processing time needed. Once the set threshold for processing time is passed, the attacker usually targets another account. This project is meant to target a single user, by using personal information of this person and generating password out of this information.

Doing this for a single user should require far less processing time, while requiring more for a large number of users, since the generated passwords should be different for all of these users.

The program is to be used by ethical hackers, penetration testers, enthousiasts and other non-malicious users. The main target group will be users

with basic technical knowledge of security practices and tools.

The program should use personal information to be able to guess the password(s) of a specific person and write this into a word list, to be used by Fuzzing/Brute Force attack tools. This is meant for penetration testers and security researchers to be used to try and gain access to a specific account, with consent of the target.

To be able to make the program efficient in guessing personal passwords, the program should be able process personal information and use this information to generate different possibilities in the form of keywords, combined with often used practices in weak passwords. To be able to this, these often used practices should be mapped.

4 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?*

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

4.1 Research methodology

The research methodology will deal with the methods, used to answer the sub-questions and the main question, formulated in chapter 4. 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.

5 Limitations and Conditions

This project is bound to certain limitations and conditions. The project is reliant on a non-existent budget and should be available freely for interested parties. Also, the research and development should take around 80 hours to complete, meaning that the project might well be a proof of concept, instead of a full fledged application, only to be developed further afterwards.

The condition, set by the development team itself, is that the project is freely available to all who wish to use it for non-malicious intent, with the condition that it may not be used for financial profits in any way. Other developers are free to alter it and use it for inspiration, as long as the the developed programs remain free to use and open-sourced. For this reason, the project includes the GNU General Public License, which can be found in the source repository.

The code will be made available through Github, along with the documentation. A manual for users will be published in the form of a *README* file in the same Github repository, and the project will be concluded with a presentation of the final results.

6 Planning

Date	Activity	Deadline	Involved members
January 2021	Preparations project	31^{st} of January	Project team
February 2021	Research	28^{th} of February	Project team
March 2021	Design program, start documentation	31^{st} of March	Project team
April 2021	Finalise program	30^{st} of April	Project team
May 2021	Present results to evaluating lecturers	31^{st} of May	Project team, evaluating lecturers