Assignment 1 Part 2

Technique

The technique used to crack the password was to use a wordlist from the internet which has most used passwords and passwords breached from popular data breaches.

One few of the popular wordlists i used was Rockyou.txt and 10-million-password-list-top-1000000.txt

The hashes in the linkedin data breach were 40 characters long. And based on the length we can deduce that the hashing algorithm used is SHA1. Because, any input when hashed using SHA1 gives 40 characters long hash.

Similarly in case of formspring data breach the hashes were 64 characters long. SHA256 is the hashing algorithm which gives 64 characters long hash.

Based on these info wordlist was encrypted SHA1 and SHA256 respectively.
Then the encrypted wordlist was then used to see if any of the hashes match with the hashes present in the input file. This was also done as a separate step to save precessing time. As encrypting everything on the go will take a lot of time.

The hashes from the input file were sorted using python's inbuilt sort method which uses Timsort method.
Then each entry from the encrypted wordlist is searched against the sorted hash list to check if there is any match. Binary search was used to make search faster. Once there is a match program can retrieve the original password from the list and print it in output file.

Other Techniques

The other technique considered was to bruteforce techniques where different combinations of alphanumeric characters are hashed and are matched against the input file hashes. But since we do not know the length of the original password in the hash file. The brute force technique will result in lot of combinations and will take longer time to crack

The wordlist passwords retrieved from the internet were in plain text which was then converted to SHA1 and SHA256 hashes respectively to be used in the password cracking program.

Comparison

The SHA1 passwords were pretty straightforward to crack and took less than 5 mins to scan the while file and get the matches.

The SHA256 was much more difficult to crack and took more than 20 mins to process all the inputs in the file.

In conclusion, SHA256 provide more security compared to SHA1.