Application of Combinatorics in Helping to Determine the Security of a Password

Dafi Ihsandiya Faraz - 13516057
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
13516057@std.stei.itb.ac.id

***Abstract*—Despite the presence of other identity verification methods such as fingerprint recognition, facial recognition, and security tokens, people still use passwords as as their preferred method to verify their identity on all sorts of platforms. However, the passwords that they use tend to be very simple and easy to guess, resulting in a higher probability of a compromise. This paper is intended to increase the awareness of the importance of a strong password by using combinatronics as a basis.**

***Keywords*—Password, Security, Probability, Combinatorics**

I. Introduction

A password is a series of characters that is used to verify the identity of a person in order to gain access to a device, an account, or other resources. Since a password consists of a finite set of characters, consisting of a finite number of letters, numbers, and symbols, we can make use combinatorics to determine the probability of a given password to be guessed.

It is to be noted that the security of a password cannot be determined solely by mathematical combinatorics, it may also be determined by the user's psychological tendencies and other factors. Due to the purpose of this paper being an essay to cover a topic that involves discrete mathematics, this paper will not cover the psychological aspects of a password.

This paper also involves theoretical cases that may or may not happen in real life, and is used only for the purpoose of further developing the concepts and theories that is explained in this paper.

II. Basic Theories

A. Combinatorics

Combinatorics is a branch of mathematics studying the enumeration, combination, and permutation of sets of elements and the mathematical relations that characterise their properties [1].

In this paper, we will make use of combinatorics in helping us determine the security strength of a password. The security strength of a password will be determined using the mathematical concept of probability, which represents the chance of an occurrence in the form of a fraction.

B. Probability

A probability is a branch of combinatorics that focuses on determining the likelihood of a given occurrence. The probability of a given occurrence can be determined using the following formula:

$$Probability = \frac{number\ of\ favorable\ outcomes}{number\ of\ possible\ outcomes}$$

Fig 2.1 Probability formula
Source: https://www.mathplanet.com/education/algebra-2/discrete-mathematics-and-probability/probabilities *accessed on December 3rd 2017.*

C. Permutations

A permutation is a possible arrangement of a set, which consists of a distinct number of elements. The number of permutations for a set of elements can be determined by the length of the set, using the mathematical formula:

$${}_{n}P_{k} \equiv \frac{n!}{(n-k)!}$$

Fig 2.2 Permutation Equation (Uspensky 1937, p.18)

With P being the number of permutations possible, n being the number of characters to choose from, and k being the length of the set.

For example, here are all the permutations for the word 'MATH':

- M,A,T,H
- A,M,T,H
- T,M,A,H
- M,T,A,H
- A,T,M,H
- T,A,M,H
- T,A,H,M
- A,T,H,M
- H,T,A,M
- T,H,A,M
- A,H,T,M
- H,M,T,A
- M,H,T,A
- T,H,M,A
- H,T,M,A
- M,T,H,A
- T,M,H,A
- A,M,H,T
- M,A,H,T
- H,A,M,T
- A,H,M,T
- M,H,A,T

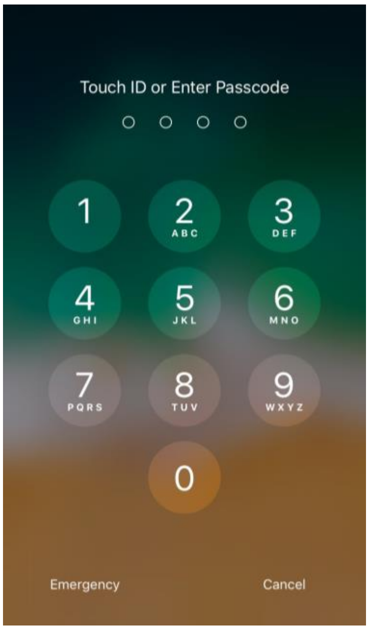
- H,A,T,M
- H,M,A,T

As shown in the list above, despite the length of the word being only four characters long, the number of permutations this word has is quite large. Based on the equation in Figure 2.2, it can be inferred that the longer the length of a set, the larger the amount of permutations the set has. For this reason, lots of websites now require a minimum number of characters for your password in order to increase the difficulty for other people to guess your password correctly.

III. Different Types of Passwords

A. 4-digit Password

The 4 digit password is the most common password type, it can be found on mobile phones, atm machines, and even bicycle locks. The 4 digit password consists of four numerical characters ranging from 0 to 9, and is one of the most unsafe password type there is due to its characteristics.



Touch ID or Enter Passcode

Fig 3.1 A 4-digit password on a smartphone

The 4-digit password has a relatively small amount of possible combinations. Each digit can only range from 0 to 9, which means that the number of characters to choose from is only 10. The possibility of a 4-digit password can be simply be calculated by multiplying the possible characters of each of the number slot:

10	10	10	10

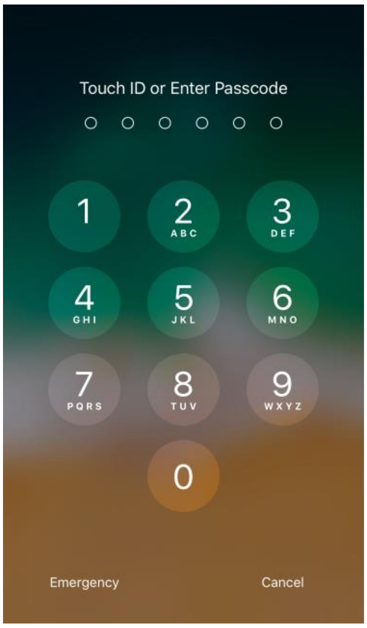
10*10*10*10 = 10,000 possibilities

Let us say that your friend has a 4-digit password on his phone, and you saw lots of fingerprints on the numbers 7, 8, 1, and 0. Since your friend uses his phone a lot and his phone is set to ask for a password every time he turn on his phone, it can be assumed that those four numbers are the numbers for his password. By using the permutation equation from figure 2.2:

$$\frac{4!}{(4-4)!} = \frac{4!}{1} = 24\ permutations$$

There are only 24 possible combinations of the numbers 7,8,1, and 0, meaning that we have a 1 out of 24 chances to guess his password correctly. This high chance of guessing the password correctly is one of the reason most devices now recommend a 6-digit password or even an alphanumeric password. Even though the 4-digit password is the least secure form of a password, lots of people still use it due to its simplicity.

B. 6-digit Passwords



Touch ID or Enter Passcode

Fig 3.2 A 6-digit password on a smartphone

The 6-digit password, as the name suggests, consists of 6 characters that consists of the numbers 0 to 9, and is meant to be a replacement of the less secure 4-digit password. The chances of guessing a 6-digit password is 100 times lower than guessing that of a 4-digit password.

10	10	10	10	10	10

10*10*10*10*10*10 = 1,000,000 possibilities

Let us say that now your friend has a 6-digit password on his phone, and his fingerprints are present on the numbers 7, 8, 9, 0, 1, and 2. With the same conditions as the problem in section A, the number of possible permutations for this set of numbers is:

$$\frac{6!}{(6-6)!} = \frac{6!}{1} = 720\ permutations$$

As you can see, the number of permutations for a 6-digit password is significantly larger than that of a 4-digit password. Now it would take a maximum of 720 tries for you to guess his password correctly.

C. Alphanumeric Passwords

Alphanumeric Passwords, as given by the name, consists of alphabet and numeric characters. Normal alphanumeric passwords don’t allow the use of symbols or special characters, therefore there are a total of 62 (10 numbers, 26 lowercase letters, 26 uppercase letters) to choose from. Alphanumeric passwords don’t usually limit the length of the password, therefore we have 62 to the power of n possibilities, whereas n is the length of the passwords.

62	62	62	62	62	…

62*62*62*62*62*… = 62^n possibilities

As you can see, alphanumeric passwords significantly decrease the probability of anyone guessing your password, and as you increase the length of the password, the chances of your password being guessed is lowered by 62 times.

For comparison purposes, we will use the same case as section A and B, with the characters being addc28. What you may realize now is that there is now two identical characters in the password. When you look at his screen, there are only 5 fingerprints, however you are sure that he typed 6 characters for his password. So now, the only characters you know he typed in is a, d, c, 2, and 8.

The password may have been aadc28, addc28, adcc28, adc228, or adc288. To calcuate all the possibilities of the password, we can use the method from section A and B for each of the probable passwords above, and adding the possibilities together.

$$\frac{6!}{1}+\frac{6!}{1}+\frac{6!}{1}+\frac{6!}{1}+\frac{6!}{1}=5*\frac{6!}{1}=3600\ permutations$$

You may notice that by increasing the repetition of characters in a password, it may decrease the chance of anyone getting your password right. If the characters you in the theoretical case above were a,d,c, and 2, and you still saw that he typed in 6 characters. The password possibility would increase again, therefore lowering the chance of guessing the password correctly. Now, the possible password may be aaadc2, adddc2, adccc2, adc222, aaddc2, addcc2, adcc22, and so on.

D. Complex Alphanumeric Passwords

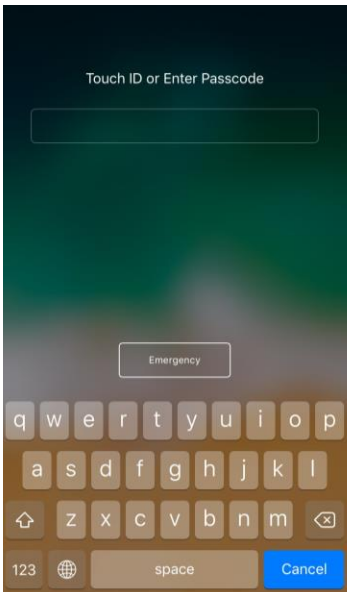


Fig 3.3 Password with complex alphanumeric input on a smartphone

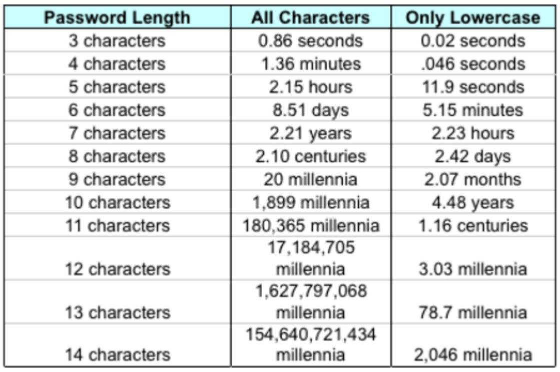
Complex alphanumeric passwords are alphanumeric passwords with the addition of special characters and symbols such as dots, commas, hashes, etc. With the addition of special characters, the chances of guessing the password correctly has significantly decreased. Lots of websites now require at least one special character to be included in your character due to this fact. This is currently the safest type of password possible, as there is an abundant amount of characters to choose from, thus significantly lowering the chance of anyone guessing it correctly.

The total possible character combination for the password uses the same concept as the alphanumeric passwords from section c, however, rather that 62 possible characters, now the total possible character is represented by ‘c’ (as the total possible character varies from one platform to another).

c	c	c	c	c	…

c*c*c*c*c*... = c^n possibilities

With complex alphanumeric passwords, adding just a few more characters to the password is shown to significantly slow down any attempts to crack the password by sheer brute force (brute force cracking method without using any dictionaries or other forms of information). This is is further supported by an article from the website lifehacker.com, which shows the amount of time it takes to crack a complex alphanumeric password with a specified amount of characters:



Password Length	All Characters	Only Lowercase
3 characters	0.86 seconds	0.02 seconds
4 characters	1.36 minutes	.046 seconds
5 characters	2.15 hours	11.9 seconds
6 characters	8.51 days	5.15 minutes
7 characters	2.21 years	2.23 hours
8 characters	2.10 centuries	2.42 days
9 characters	20 millennia	2.07 months
10 characters	1,899 millennia	4.48 years
11 characters	180,365 millennia	1.16 centuries
12 characters	17,184,705 millennia	3.03 millennia
13 characters	1,627,797,068 millennia	78.7 millennia
14 characters	154,640,721,434 millennia	2,046 millennia

Fig. 3.4 Brute Force Duration Table. Source: https://lifehacker.com/5505400/how-id-hack-your-weak-passwords , accessed on December 4th 2017.

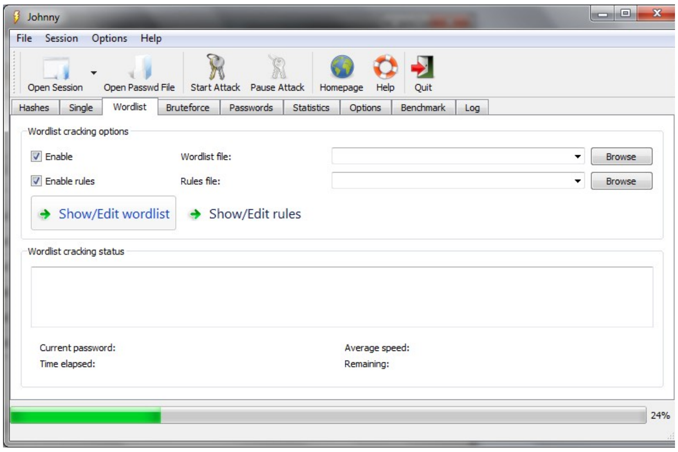
IV. Brute Force Cracking

Brute force (also known as brute force cracking) is a trial and error method used by application programs to decode encrypted data such as passwords or Data Encryption Standard (DES) keys, through exhaustive effort (using brute force) rather than employing intellectual strategies[2].

Just as a criminal might break into, or "crack" a safe by trying many possible combinations, a brute force cracking application proceeds through all possible combinations of legal characters in sequence. Brute force is considered to be an infallible, although time-consuming, approach[2].

Basic brute force cracking tries all the possible combinations of characters, if the password is a 4-digit password, then the brute force method will try to input the password as follows:

0000, 0001, 0002, 0003, 0004, 0005, 0006, 0007, 0008, and so on until it reaches 9999. The same method goes for 6-digit passwords and alphanumeric passwords. Since alphanumeric passwords can ususally be of any length, we can decrease the chances of the brute force cracking to succeed by increasing the number of characters in the password.



Wordlist cracking options

☑ Enable Wordlist file:

☑ Enable rules Rules file:

Wordlist cracking status

Current password: Average speed:

Time elapsed: Remaining:

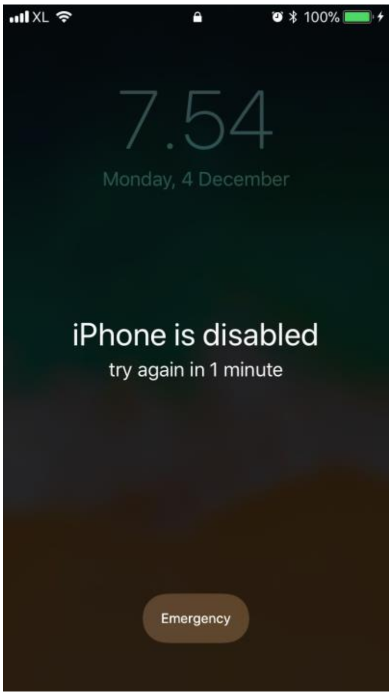
24%

Fig 4.1 John the Ripper: A Brute Force Program for Windows, Linux, DOS, and OSX. Source: https://fossbytes.com/best-password-cracking-tools-2016-windows-linux-download/ , accessed on December 4th 2017.

Brute force programs nowadays make use of dictionaries which contains words from all sorts of languages, in an order that puts the more frequently used word at the front. This results in a significantly higher chance of getting the password right.

From the way this program works, it can be inferred that the use of meaningful words in a password is to be avoided whenever possible. A long, random, and high complexity password would prevent, or at least reduce the probability of these brute force cracking programs to work.

To prevent brute force programs to work, several platforms limit the number of attempts to enter a password, an example of this safety feature can be found on numerous smartphones, which prevents the person using it from entering a password for a specified amount of time after a number of password attempts have been made.



iPhone is disabled

try again in 1 minute

Fig 4.2 An example of the brute force prevention on a smartphone

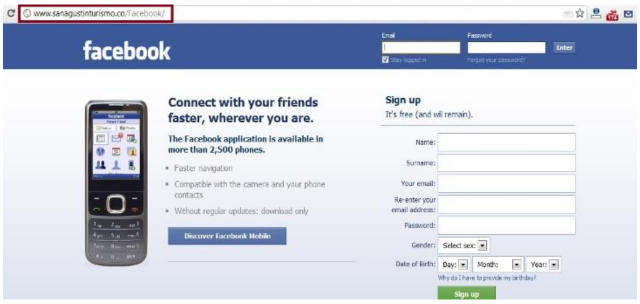
V. Other Forms Of Password Exploitation

A. Rainbow Table Attack

A rainbow table is a list of pre-computed hashes - the numerical value of an encrypted password, used by most systems today - and that's the hashes of all possible password combinations for any given hashing algorithm mind. The time it takes to crack a password using a rainbow table is reduced to the time it takes to look it up in the list. However, the table itself will be huge and require some serious computing horse power to run, and it's useless if the hash it is trying to find has been 'salted' by adding random characters to the password before applying the hashing algorithm. [3].

B. Phishing

Phising is an attempt to gain access to your password by creating a fake link or site that looks similar to the login screen you expect to see when logging in to the site that you use. This form of password exploitation tricks you into entering your personal information, including your password, by exploiting your carelessness. This is a very common form of password exploitation and can be found everywhere, including that spam folder in your email.



Connect with your friends faster, wherever you are.

The Facebook application is available in more than 2,500 phones.

- Faster navigation
- Compatible with the camera and your phone contacts
- Without regular updates: download only

It's free (and will remain).

Fig 5.1 An Example of a Phising Site. Source: http://4.bp.blogspot.com/-_8BYD9vs4hs/TtHsJbhHQeI/AAAAAAAADss/CoMLPkj_LU8/s1600/Untitled.jpg , accessed on December 4th 2017.

Websites with the 'https://' prefix (usually found on most login web pages), has a digital certificate, which indicates whether or not the site is valid. By checking this certificate (which is usually provided beside the address bar on your browser), you can lower the chance of being tricked by a phishing site.

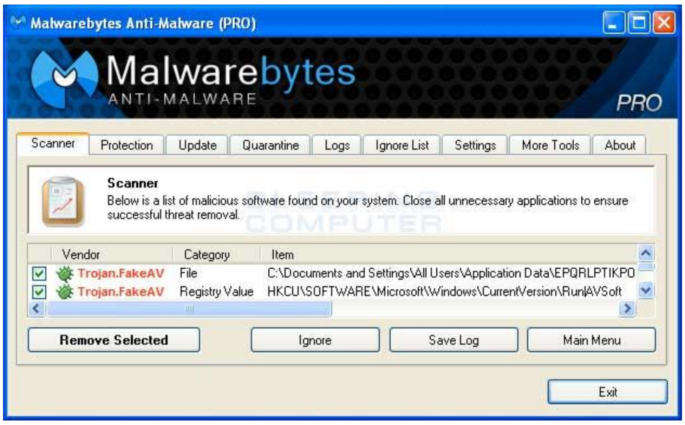
C. Social Engineering

Social engineering is basically a form of phishing that happens in the real world. The person attempting this method acts as a person with access to a certain information, in this case, a password, and pretends that he or she has forgotten his password and contacts another person with the password that he or she needed. If the attempt succeeds, the social engineer will have that very information. This form of exploit can be compared to a con artist.

D. Malware

A malware or a malicious software is a software that can access information on your computer and sends it over the internet to the malware deployer. This software can appear on your computer when you access shady websites, download files from unverified sources, opening email links that you aren't sure of, and also downloading from illegal websites such as torrents.

This form of exploit has a higher chance of happening the more you use the internet. There are several ways to prevent this malicious software from appearing on your computer. One of the way is to install an antivirus, which scans your computer for programs that are suspicious and possibly threatening. Another way is to only click on email links from people that you know, and rechecking the web address every time you enter your personal information. A login page usually uses the prefix 'https://' indicating that a secure connection is being established between your computer and the server.



Scanner

Below is a list of malicious software found on your system. Close all unnecessary applications to ensure successful threat removal.

	Vendor	Category	Item
☑	Trojan.FakeAV	File	C:\Documents and Settings\All Users\Application Data\EPQRLPTIKPO
☑	Trojan.FakeAV	Registry Value	HKCU\SOFTWARE\Microsoft\Windows\CurrentVersion\Run\AVSoft

Fig 5.2 An Antivirus detecting malware on computer. Source: https://www.bleepstatic.com/swr-guides/f/futurro-antivirus-software/mbam-futurro-antivirus-software.jpg accessed on December 4th 2017.

V. Conclusion

When creating a password, avoid using the 4-digit or even the 6-digit password, try to use a combination of letters, numbers, and symbols whenever possible. Also, do not use meaningful words in your password, as it is easily guessable by brute force programs. In addition, do not include your personal information such as your birth date, your name, or the city you came from as these informations are publicly available and may be included in the brute force dictionary, thus increasing the chance of your password to be compromised.

Although combinatorics play a significant role in securing your password, there are also other factors that affect the security of your password, by paying attention to these factors, you can significantly reduce the chance of you password being compromised.

VI. Acknowledgment

The Author would like to thank God for making this paper possible. The Author would also like to express his gratitude to Lecturer Dr. Judhi Santoso M.Sc. and Dr. Ir. Rinaldi Munir, MT. for providing the necessary resources to support the making of this paper. Last but not least, The Author would like to thank his parents for proving all the support he needs in the process.

References

[1] http://mathworld.wolfram.com/Combinatorics.html accessed on December 3rd 2017.

[2] http://searchsecurity.techtarget.com/definition/brute-force-cracking accessed on December 3rd 2017.

[3] http://www.alphr.com/features/371158/top-ten-password-cracking-techniques accessed on December 4th 2017.

Munir, Rinaldi, Matematika Diskrit, Informatika, Bandung, 2012.

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 3 Desember 2017

Dafi Ihsandiya Faraz
13516057