# Analysing the Usage of Character Groups and Keyboard Patterns in Password

J. Kävrestad, J. Zaxmy and M. Nohlberg

School of Informatics, University of Skövde, Skövde, Sweden
e-mail: joakim@kavrestad@his.se, johan.zaxmy@his.se, marcus.nohlberg@his.se

## Abstract

Even with the advances in different methods for authentication, passwords remain the most common approach for authentication as well as for encryption of user data. Password guessing attacks have grown to be a vital part of computer forensics as well as penetration testing. In this paper, we seek to provide a statistical analysis of password composition by analyzing what character sets that are most commonly used in over 1 billion leaked passwords in over 20 different databases. Further, we use a survey to analyze if users that actively encrypt data differ from the norm. The results of this study suggest that American lowercase letters and numbers are the, by far, most commonly used character sets and that users who actively encrypt data use keyboard patterns and special characters more frequently than the average user.

## Keywords

Passwords, password guessing, keyboard patterns, encryption, brute force

## 1 Introduction

It is a well-established fact that passwords are the most common form of authentication in modern computer systems (Nielsen, Vedel, & Jensen, 2014; Woods & Siponen, 2018). Furthermore, a modern user has many different accounts that need to be passwords protected, and a general password guideline is to have different passwords for different accounts (Ur et al., 2015). While that recommendation seeks to protect the user in case of a password leak, it is hard for users to follow and leads to users making simple and easy to guess modifications to their passwords, if they use unique passwords at all (Ur et al., 2015). There are indeed other means of authentication and strategies for managing passwords, including biometric solutions and password managers. However, those solutions have failed to gain significant interest among the general public leaving passwords as the most common form of authentication (Ruoti, Andersen, & Seamons, 2016).

As of today, any organization is depending on IT and thus, secure computer systems. A common attack vector, used by attackers, is to target user account passwords in order to get access to computer systems. A way to protect said systems is to actually attack them in order to find weaknesses, a process called penetration testing. As described by Denis, Zena, and Hayajneh (2016), attempting to crack passwords in a common part of penetrating testing.

Another discipline where password cracking is a common practice is within digital forensics. Digital forensics seeks to examine digital devices in order to uncover data that can be used as evidence in a criminal investigation (Kävrestad, 2018). A typical scenario is that the data that is of interest to a forensic examiner is protected by a password in some way, for instance using encryption (Karie & Venter, 2015; Vincze, 2016). Freiling, Groß, Latzo, Müller, and Palutke (2018) even call disk encryption the most common threat to evidence acquisition for hard drives today.

As described by Tatlı (2015), password cracking can be performed using a brute force attack or a dictionary attack. A brute force attack is an attack where all possible passwords are tested until the correct one is found. When performing a dictionary attack, a dictionary containing different possible passwords is created, and then, the possible passwords are tested.

While pure brute force attacks and dictionary attacks can be seen as to opposite attacks, there are several ways to leverage statistical insight about how users create passwords in smart brute force attacks. As exemplified by Weir (2010), smart brute force attack can be executed by limiting the key space by excluding characters that are rarely used, or by focusing the attack on common grammatical structures.

The aim of this paper is to analyze how frequently different character sets and keyboard patterns are used in passwords by automatically analyzing over 1 billion passwords that leaked online. The generated data will provide an overview of the average computer users password behavior in regard to the aspects of interest for this study. Further, as argued by Parker, Ophoff, Van Belle, and Karia (2015), it is reasonable to assume that users that employ encryption are more security aware than the average user. Based on that assumption, this paper also surveys the difference in usage of character sets and keyboard patterns between ordinary users and user that actively encrypts their data. This is done by means of a survey where over 600 respondents answered questions about the structure of their personal e-mail account password and whether or not they actively encrypt their data.

The results of this study will add to the knowledge base about user-generated passwords and provide input into the execution of dictionary attacks and smart brute force attacks. Further, the study identifies differences between overall users and users that intentionally encrypts data and can, as such, provide valuable input into password guessing attacks against encrypted data.

## 2 Methodology

The study was completed in two steps where the general usage of character sets and keyboard patterns was examined by analyzing dumps of passwords that were found openly on the internet. The second part of the study was conducted using an online survey where respondents were asked if they intentionally encrypt their data, then they were asked a series of questions about the structure of their personal e-mail account password. The remainder of this section provides a more detailed description of each step in the research process.

## 2.1 Automatic Password Classification

The automated password classification was completed by a Python3 script. The script first analyzed what character sets that the examined passwords contained and reported how frequently the different character sets and compositions of character sets was used. The character sets were defined in two different ways, called ASCII and UTF8. The UTF8 definition followed the definition of UTF8 as described in "C0 Controls and Basic Latin" and "C1 Controls and Latin-1 Supplement" (unicode.org, 2018). The character sets were defined using regular expressions as follows:

- Lowercase letters (55 characters): [\u00DF-\u00F6\u0061-\u007A\u00F8-\u00FF]
- Uppercase letters (53 characters): [\u0041-\u005A\u00C0-\u00D6\u00D8-\u00DE]
- Numbers (10 characters): [0-9]
- Special Characters (60 characters): [\u0021-\u002F\u003A-\u0040\u005B-\u0060\u007B-\u007E\u00 A1-\u00BF\u00F7\u00D7]

As it is reasonable to assume that some of the characters defined as letters in UTF8 may, in fact, be treated or seen as special characters by computer systems, we also defined the character sets to resemble the American Standard Code for Information Interchange (ASCII). In this definition, the only characters defined as letters was those defined as letters in "C0 controls and Basic Latin" (a to z and A to Z), all characters in "C1 Controls and Latin-1 Supplement" was considered special characters (unicode.org, 2018). This definition was expressed using regular expressions as follows:

- Lowercase letters (26 characters): [\u0061-\u007a]
- Uppercase letters (26 characters): [\u0041-\u005A]
- Numbers (10 characters): [0-9]
- Special Characters (116 characters): [\u0021-\u002F\u003A-\u0040\u005B-\u0060\u007B-\u007E\u00A1-\u00FF]

The script also analyzed if the password was a keyboard pattern. A keyboard pattern is, in this paper, defined as a sequence of adjacent keys entered on a keyboard, similar to Wheeler (2016). As demonstrated by Ur et al. (2016), there are other ways to define keyboard pattern. However, since this paper uses a script for automatic detection of a pattern-based password, a strict definition was deemed necessary. For this purpose, a dictionary listing all characters as keys and adjacent keys as values was created. In this study, only the QWERTY keyboard layout was considered. Further, a combination of the American and Swedish keyboard layout was used, where the American keyboard layout was given precedence whenever conflicts emerged. Passwords of only one character were not considered patterns.

As for the databases analyzed, the main database is called "Exploit.In" and is a compilation of passwords from different leaks that first appeared in 2016 (haveibeenpawned.com, 2018). There are several articles and blog posts that report the database as real, and containing valid password (Casal, 2017; Hunt, 2017). The

database dump used in this paper contained 1,012,024,699 lines, and each line was considered one password. The origin of the passwords in this database is unknown, and that is certainly a problem in terms of validity. Further, it is impossible to know if a single user is present one or more times in the database, making bias in the database possible. While these validity concerns are problematic, we argue that using this database is interesting because it is very large and there are reports of it containing real passwords. Also, to increase the validity of the study, data from ten additional databases were taken into account in this project for comparison, as suggested by Lincoln and Guba (1985).

Schrittwieser, Mulazzani, and Weippl (2013) discuss ethical considerations in research, and one important practice is to ensure that no one is harmed by the research. This study examines real passwords from real users. All passwords used in this study was found using regular search engines, thus they are available to anyone, and because of that, we argue that using them in the study is not problematic. However, to maintain some degree of integrity for the password holders we choose not to specify where the databases have been downloaded from, or any of the individual passwords analyzed. Also, except for "Exploit.In", we choose not to publish the names of the other databases as their names reveal what website they are from. The case with "Exploit.In" is that the name does not reflect where the passwords originate from, and therefore, there are no ethical issues in publishing it.

### 2.2 Password Statistics Survey

In order to survey if users that actively decide to encrypt their data differ from other users, an online survey was used. The survey was designed to let the respondents report what strategies and character sets they used when designing the password to their personal e-mail account. The respondents were also asked to report if they actively encrypted their data or not. Dichotomous questions were used throughout the survey to enforce the respondents to make a hard choice or opt out from answering.

The survey was distributed through Amazon Mechanical Turk, a platform that allows a requester to create a task and pay users (called mturks) to complete that task (Buhrmester, Kwang, & Gosling, 2011). This method allows for the collection of a large sample in a relatively short amount of time. Also, research into data quality suggests that using mturks to gather survey data produces high-quality data samples (Buhrmester et al., 2011; Kees, Berry, Burton, & Sheehan, 2017).

Before the survey was executed, a pilot test was performed to ensure that the survey was understandable. 30 persons were paid to complete the pilot on Amazon Mechanical Turk, and two additional participants performed the survey supervised by the researchers. Following the pilot, it was evident that some mturks just clicked through the questions, thus two control questions were added and later used to exclude respondents that gave incorrect responses to the control questions.

In analyzing the survey, the respondents were split into two groups based on their response to the question about encryption. Mean values showing the percentage of respondents that used patterns and the different character sets was calculated. Further,

Spearman´s rank correlation, as described by Mukaka (2012), was used to evaluate the relationship between active use of encryption and use of patterns and character sets. To ensure the validity of the results, Kendall's rank correlation was also used.

# 3 Results

This section presents the results of the two parts in the research process, respectively

## 3.1 Analysis of Password Databases

Using a script written in python3, 11 databases of leaked passwords were processed. The script reported how many of the passwords that could be considered to be patterns and how many passwords that used uppercase letters (UC), lowercase letters (LC), numbers (No) and special characters (SP). The results are presented in Table 1, below. "Exploit.In" is database number 11.

| No | Char def | Passwords | LC | No | SP | UC | Pattern |
|---|---|---|---|---|---|---|---|
| 1 | ASCII | 37 126 | 95.69 | 84.73 | 10.69 | 6.79 | 0.003 |
| 1 | UTF8 | 37 126 | 95.69 | 84.73 | 10.68 | 6.79 | 0.003 |
| 2 | ASCII | 38 820 | 89.23 | 53.67 | 1.65 | 8.43 | 1.05 |
| 2 | UTF8 | 38 820 | 89.23 | 53.67 | 1.65 | 8.43 | 1.05 |
| 3 | ASCII | 95 072 | 77.58 | 54.14 | 1.03 | 6.43 | 1.18 |
| 3 | UTF8 | 95 072 | 77.58 | 54.14 | 1.03 | 6.43 | 1.18 |
| 4 | ASCII | 184 365 | 86.46 | 54.12 | 2.10 | 10.10 | 0.88 |
| 4 | UTF8 | 184 365 | 86.46 | 54.12 | 2.10 | 10.10 | 0.88 |
| 5 | ASCII | 192 831 | 84.66 | 38.53 | 0.65 | 3.05 | 1.59 |
| 5 | UTF8 | 192 831 | 84.66 | 38.53 | 0.63 | 3.05 | 1.59 |
| 6 | ASCII | 226 082 | 92.85 | 29.61 | 0.84 | 3.78 | 0.72 |
| 6 | UTF8 | 226 082 | 92.85 | 29.61 | 0.83 | 3.78 | 0.72 |
| 7 | ASCII | 375 853 | 85.00 | 64.84 | 0.41 | 6.52 | 0.71 |
| 7 | UTF8 | 375 853 | 85.00 | 64.84 | 0.41 | 6.52 | 0.71 |
| 8 | ASCII | 720 303 | 97.44 | 97.71 | 6.61 | 10.86 | 0.23 |
| 8 | UTF8 | 720 303 | 97.45 | 97.71 | 6.60 | 10.91 | 0.23 |
| 9 | ASCII | 313 2006 | 84.82 | 69.26 | 2.77 | 0.00 | 0.29 |
| 9 | UTF8 | 3 132 006 | 84.82 | 69.26 | 2.77 | 0.00 | 0.29 |
| 10 | ASCII | 3 431 316 | 75.82 | 70.64 | 3.98 | 18.75 | 0.29 |
| 10 | UTF8 | 3 431 316 | 75.88 | 70.64 | 3.75 | 19.26 | 0.29 |
| 11 | ASCII | 1 012 024 699 | 84.31 | 69.64 | 5.85 | 6.59 | 2.83 |
| 11 | UTF8 | 1 012 024 699 | 84.31 | 69.64 | 5.83 | 6.59 | 2.83 |

**Table 1: Result of automated analysis of databases**

Looking at the database called "Exploit.In", 2.83% of the analyzed passwords are considered pattern. The results from the other databases range from 0.003% to 1.59%. Looking at the analysis for character sets, the frequencies are 97.5% to 75.8% for lowercase letters and 97.7% to 29.6% for numbers. Looking only at the database called "Exploit.In", that is the by far largest database, 84.31% of the passwords contain

lowercase letters and 69.64% of the passwords contain numbers. The least frequently used characters sets are uppercase letters and special characters. The frequency of special characters ranged from 0.4% to 10.69% while uppercase letters range from 0% to 19.26%. Zooming in on "Exploit.In", 5.83% of the passwords contain special characters and 6.58% contain uppercase letters. It is also noticeable that the results are almost identical for both definitions of character sets.

Next, the script analyzed how frequently different compositions of character groups was used in the passwords. This analysis is summarized in Table 2, where Exploit.In is presented separately and an average value for all other databases is presented for comparison.

| Composition | Average ASCII | Average UTF8 | Exploit.In ASCII | Exploit.In UTF8 |
|---|---|---|---|---|
| **All groups** | 0.32 | 0.34 | 0.36 | 0.36 |
| **All except LC** | 0.08 | 0.08 | 0.06 | 0.06 |
| **All except No** | 0.11 | 0.25 | 0.09 | 0.09 |
| **All except SP** | 3.81 | 3.82 | 3.55 | 3.55 |
| **All except UC** | 1.37 | 1.34 | 2.61 | 2.6 |
| **UC and LC** | 1.13 | 1.15 | 1.1 | 1.1 |
| **UC and SP** | 0.05 | 0.03 | 0.03 | 0.03 |
| **UC and No** | 2.57 | 2.57 | 0.89 | 0.89 |
| **LC and No** | 46.97 | 46.98 | 48.38 | 48.39 |
| **LC and SP** | 1.13 | 0.94 | 2.35 | 2.34 |
| **SP and No** | 0.2 | 0.19 | 0.27 | 0.27 |
| **Only LC** | 28.02 | 28.06 | 25.87 | 25.88 |
| **Only No** | 12.99 | 12.99 | 13.52 | 13.52 |
| **Only SP** | 0.05 | 0.02 | 0.08 | 0.08 |
| **Only UC** | 1.22 | 1.24 | 0.5 | 0.5 |

**Table 2: Frequency of compositions of character sets, expressed in percent**

As shown by Table 2, the average percentages between "Exploit.In" and the other analyzed databases are similar. Passwords containing lowercase letters and numbers make up almost half of the analyzed passwords, one-fourth of the passwords only contain lowercase letters and one eight contain only numbers. In total, these three compositions make up over 80% of the analyzed passwords.

### 3.2 Surveying the difference between the common user and active encryptors

This section covers the implementation of the survey as well as the analysis of the data gathered. The survey was designed and distributed through the Amazon Mechanical Turk system and accepted responses for 10 days. The survey itself contained two sections, one with demographic questions where the respondents were asked about

their occupation and if they actively encrypted their data or not. In the second part of the survey, the respondents were asked to answer yes/no questions about their personal e-mail account password. The survey also contained two control question asking the respondents to solve very simple mathematical questions.

In total, 651 respondents answered the survey, but 47 responses were removed from the dataset due to incorrect answers to the control questions. The first part of the analysis was to analyze the respondent's use of different character sets and keyboard patterns in their passwords. The results of that analysis combined with the results for the database called "Exploit.In" are listed in Table 3 below.

| Dataset | LC | UC | Sp | No | Use of patterns |
|---|---|---|---|---|---|
| Survey (all) n=604 | 96% | 90% | 75% | 96% | 10% |
| Survey (enc) n=212 | 93% | 90% | 83% | 93% | 17% |
| Survey (non_enc) n=392 | 97% | 91% | 70% | 97% | 5% |
| Exploit.In n=1012024699 | 84% | 7% | 6% | 70% | 3% |

**Table 3: Percentage of yes responses**

As seen in Table 3, the results of the survey suggest that all over 90% of all respondents use lowercase letters and numbers in their passwords. This is a similar result to the previous automatic analysis of leaked password databases. However, a high number of respondents (90%) also claim to use uppercase letters and 75% report using special characters. The tendencies that special characters and uppercase letters are used less frequently than lowercase letters and numbers is still present. A reason for why uppercase letters and special characters are getting high numbers in the survey can be that large e-mail providers encourage or even enforce the use of different character sets in passwords. Another explanation can be in the demography of the respondents where 35% report actively encrypting their data and 24% report being a student in, holding a degree in or working in IT. Thus, it is reasonable to assume that the security awareness in the group of users that responded to the survey is higher than in the general population.

Looking at the relationship between active use of encryption and the use of different character sets and patterns, Spearman's rank correlation was used. In this case, yes responses were coded as 1, and no responses were coded as 0. Spearman's rank correlation returns a value between 1 and -1 where 1 is a perfect correlation. In this case, a perfect correlation between the use of encryption and the use of keyboard patterns would mean that all respondent's that actively encrypt data also use keyboard patterns. Kendall's rank correlation was also used to validate the results. The result of these tests is presented in Table 5 below.

| Question | Kendall´s | Sig. (2-tailed) | Spearman's | Sig. (2-tailed) |
|---|---|---|---|---|
| **Pattern** | .194** | 0.000 | .194** | 0.000 |
| **Special** | .149** | 0.000 | .149** | 0.000 |
| **Uppercase** | -.015 | 0.705 | -.015 | 0.705 |
| **Lowercase** | -.068 | 0.092 | -.069 | 0.092 |
| **Numbers** | -.096* | 0.018 | -.096* | 0.018 |

**Table 4: Correlation tests. * marks correlations that are significant at the .05 level and ** marks correlations that are significant at the 0.01 level (n=604)**

As seen in Table 4, both tests generate almost identical values. There are three significant correlations reported. For patterns and special characters, a correlation of 0.194 and 0.149 respectively are reported. These correlations are significant with a p-value of 0.01. Also, there is a negative correlation of -0.096 for numbers that are significant with a p-value of 0.05. The results mean that respondents that actively encrypt their data use patterns and special characters more frequently than users that do not encrypt their data. Further, they use numbers slightly less.

# 4 Conclusions

The aim of this paper was to map the use of different character groups and keyboard patterns in passwords and to analyze differences in the use of those password characteristics between users that actively encrypt their data and the general user.

The study suggests that numbers and lowercase letters are used far more often than uppercase letters and special characters. The study also revealed that just above 25% of the analyzed passwords contained only lowercase letters, and about 13% contained numbers. In conclusion, the character sets numbers, and lowercase letters cover over 80% of the analyzed passwords. In this study, the character sets were defined in two different ways called ASCII and UTF8. In the datasets used in this study, the definition of the character sets had a negligible impact on the results suggesting that non-American letters are rarely used in passwords. The implication of this is that the paper suggests that non-American letters can be omitted in passwords guessing attacks, and that insight would heavily reduce the key space.

The script also analyzed the frequency of keyboard patterns in the password databases. In this study, just under 3% of the analyzed passwords were considered patterns. At this point, it should be mentioned that the definition of a keyboard pattern is a limitation in this study as it is possible to define a keyboard pattern in different ways. For instance, one could argue that using the characters at the top corners of the keyboard would also be a pattern. However, it was deemed necessary to use a strict definition since a broader definition of a pattern would also risk yielding false positive hits. Also, using the Swedish keyboard layout was a mistake left from development.

A survey was used to analyze if the user that actively encrypt their data differ from the general user with regards to the use of different character sets and keyboard patterns.

It showed that users that actively encrypt their data is more prone to use pattern-based passwords and special characters, and somewhat less prone to use numbers in their passwords. Further, the survey differed from the leaked databases in that the respondents reported using special characters and uppercase letters more frequently than what was found when analyzing the password databases. This can be due to the fact that large e-mail providers emphasize or even enforce the use of different character groups when generating passwords. This suggests that the context where a password is created have a great impact on the password structure. This notion is further discussed by Shay et al. (2016) that demonstrates that different sites use very different password policies and that that affects the password composition

The results of this study provide insight into how users construct passwords and password statistics in general. As such, it can serve as a background for practitioners when executing password guessing attacks. In particular, this study suggests that most passwords are made up of lowercase letters, numbers or a combination of those. However, the study also suggests that the context where a password is created has a significant impact on the resulting password structures. As such, this study emphasizes that leveraging knowledge about the password that needs to be cracked is important. While the usefulness of leveraging personal information in password guessing attacks has been discussed time and again (Hitaj, Gasti, Ateniese, & Perez-Cruz, 2017; Houshmand & Aggarwal, 2017), this study also suggests that it is important to analyze the site or service that the password is for and adjust the attack according to the password policies of that platform. Further, the study suggests that users that do encrypt their data are more prone to use keyboard patterns and special characters than the regular user. The results from this study are also important for the research community, as they bring new insights into a key security behavior, selection of passwords, of both security minded, and regular users. This insight can be used for further research into awareness training, specifically when it comes to influencing an improved password quality.

Looking at the limitations of using leaked password databases, that has to be considered uncontrolled data sources, one obvious direction for future work would be to repeat this study using data from controlled sources, for instance, the password database in one or more organizations. Another direction for future work could be to continue the endeavor for granular password statistics by looking at statistics for individual characters.

# 5 References

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on psychological science, 6*(1), 3-5.

Casal, J. (2017). 1.4 Billion Clear Text Credentials Discovered in a Single Database. *4iQ*.

Denis, M., Zena, C., & Hayajneh, T. (2016). *Penetration testing: Concepts, attack methods, and defense strategies.* Paper presented at the Systems, Applications and Technology Conference (LISAT), 2016 IEEE Long Island.

Freiling, F., Groß, T., Latzo, T., Müller, T., & Palutke, R. (2018). Advances in Forensic Data Acquisition. *IEEE Design & Test, 35*(5), 63-74.

haveibeenpawned.com. (2018). Exploit.In. Retrieved from https://haveibeenpwned.com/

Hitaj, B., Gasti, P., Ateniese, G., & Perez-Cruz, F. (2017). Passgan: A deep learning approach for password guessing. *arXiv preprint arXiv:1709.00440*.

Houshmand, S., & Aggarwal, S. (2017). *Using Personal Information in Targeted Grammar-Based Probabilistic Password Attacks*, Cham.

Hunt, T. (2017). Password reuse, credential stuffing and another billion records in Have I been pwned. Retrieved from https://www.troyhunt.com/password-reuse-credential-stuffing-and-another-1-billion-records-in-have-i-been-pwned/

Karie, N. M., & Venter, H. S. (2015). Taxonomy of challenges for digital forensics. *Journal of forensic sciences, 60*(4), 885-893.

Kees, J., Berry, C., Burton, S., & Sheehan, K. (2017). An analysis of data quality: Professional panels, student subject pools, and Amazon's Mechanical Turk. *Journal of Advertising, 46*(1), 141-155.

Kävrestad, J. (2018). *Fundamentals of Digital Forensics: Theory, Methods, and Real-Life Applications*: Springer.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry* (Vol. 75): Sage.

Mukaka, M. M. (2012). A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal, 24*(3), 69-71.

Nielsen, G., Vedel, M., & Jensen, C. D. (2014). *Improving usability of passphrase authentication.* Paper presented at the Privacy, Security and Trust (PST), 2014 Twelfth Annual International Conference on.

Parker, F., Ophoff, J., Van Belle, J.-P., & Karia, R. (2015). *Security awareness and adoption of security controls by smartphone users.* Paper presented at the Information Security and Cyber Forensics (InfoSec), 2015 Second International Conference on.

Ruoti, S., Andersen, J., & Seamons, K. (2016). *Strengthening password-based authentication.* Paper presented at the Twelfth Symposium on Usable Privacy and Security ({SOUPS} 2016).

Schrittwieser, S., Mulazzani, M., & Weippl, E. (2013). *Ethics in security research which lines should not be crossed?* Paper presented at the Security and Privacy Workshops (SPW), 2013 IEEE.

Shay, R., Komanduri, S., Durity, A. L., Huh, P. S., Mazurek, M. L., Segreti, S. M., . . . Cranor, L. F. (2016). Designing password policies for strength and usability. *ACM Transactions on Information and System Security (TISSEC), 18*(4), 13.

Tatlı, E. I. (2015). Cracking more password hashes with patterns. *IEEE Transactions on Information Forensics and Security, 10*(8), 1656-1665.

unicode.org. (2018). Unicode 11.0 Character Code Charts. Retrieved from http://unicode.org/charts/

Ur, B., Bees, J., Segreti, S. M., Bauer, L., Christin, N., & Cranor, L. F. (2016). *Do Users' Perceptions of Password Security Match Reality?* Paper presented at the Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.

Ur, B., Noma, F., Bees, J., Segreti, S. M., Shay, R., Bauer, L., . . . Cranor, L. F. (2015). *I added '!' at the end to make it secure": Observing password creation in the lab.* Paper presented at the Proc. SOUPS.

Weir, C. M. (2010). Using probabilistic techniques to aid in password cracking attacks.

Wheeler, D. L. (2016). *zxcvbn: Low-Budget Password Strength Estimation.* Paper presented at the USENIX Security Symposium.

Vincze, E. A. (2016). Challenges in digital forensics. *Police Practice and Research, 17*(2), 183-194.

Woods, N., & Siponen, M. (2018). Too many passwords? How understanding our memory can increase password memorability. *International Journal of Human-Computer Studies, 111*, 36-48