The Effect of New NIST Password Guidelines on Password Creation For Web Applications and a Comparison to The Past

A Technical Report submitted to the Department of Computer Science

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Vinh Do
Spring, 2021

Technical Project Team Members
Vinh Do

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Signature ___*Vinh Do*__________________________________ Date __5/5/21___
Vinh Do

Approved ____ Date __________
Aaron Bloomfield, Department of Computer Science

Approved ____ Date __________
Brad Campbell, Department of Computer Science

The Effect of New NIST Password Guidelines on Password Creation For Web Applications and a Comparison to The Past

The Password Paradigm Shift

Vinh Do
CS Department
University of Virginia
Charlottesville, Virginia, USA
vtd4bc@virginia.edu

ABSTRACT

In 2017, the National Institute of Standards and Technology (NIST) released new guidelines for password and security requirements for applications. These guidelines will be the focus of the research paper, which will analyze the password requirements that various online applications impose in 2020 and how these restrictions compare to not only each other but also to the password requirements of applications before these new guidelines were released. Previous research has put forth a method of mathematically calculating the minimum password strength of various web applications in order to compare sets of requirements and one conclusion from these findings is that increasing the minimum password length has greater benefits than requiring special characters or numbers. Arming users with knowledge on how to create stronger passwords that are also less complex would result in fewer password leaks and stronger systems overall. The research will be conducted through accessing different genres of modern applications with the intention of creating new accounts to determine their password requirements, and analyzing how these results compare to previously published findings. Strict password requirements impose an unnecessary burden on users, limiting the design space for possible strong passwords. Removing these requirements and replacing them with suggestions to the user on how to make a stronger password allows for greater creativity when creating a password which overall results in a stronger and more memorable secret.

1 INTRODUCTION

The average digital user has 25 password protected accounts and will use eight every day [10]. Keeping track of 25 distinct sets of characters and remembering which accounts each password belongs to can be nearly impossible. Even just generating 25 unique passwords that do not relate to each other, as is often advised to maximize security, can be a difficult task. Various solutions to this problem involve using a password manager or writing the password down on a physical piece of paper; however, these methods introduce even greater risks than simply having a weak password. Using a password manager bottlenecks all of a user’s passwords into relying on a single password to keep them safe, once again bringing us back to the original problem, but this time with even more to lose [1]. Physically writing down a password is not safe because it could easily be stolen or lost; any potential passerby such as a co-worker or a malicious guest could swiftly read the note without anyone ever noticing. Without a viable way to remember passwords, users may resort to techniques such as just using the same password for all of their systems or simply cycling through every possible password they can think of until the login succeeds. Creating a simpler password that is used uniformly across all systems has similar challenges to using a password manager and leaves users vulnerable to a brute force attack or a dictionary attack [9]. Cycling through a set of passwords costs users time and effort and if done on an insecure website, leaves users vulnerable to keyloggers or other similar threats. The often-recommended technique of using a mnemonic can make for a strong password; unfortunately, this technique is too well known, so attackers are still able to exploit these types of passwords, despite a user’s best efforts [10]. Some users even believe that any password that is easy to remember is automatically an insecure password, however this is not necessarily the case [4]. Ideally, a user would be able to create multiple passwords which are hard to crack, but also memorable enough to distinguish from each other in order to minimize the time lost when simply trying to log into an account.

2 BACKGROUND

In 2017, the National Institute of Standards and Technology (NIST) released new guidelines for password requirements with the goals of increasing the strength of passwords, as well as removing unnecessary security measures in order to increase the usability of certain applications [2]. The NIST is a non-regulating government agency that decides on acceptable standards for how to use technology; federal agencies and contractors or businesses that wish to interact with the federal government are required to be NIST compliant [7]. The new guidelines require applications to have an eight character minimum for new passwords and to support having passwords at least 64 characters long. The NIST also condemned the practice of requiring special characters and numbers, while simultaneously requiring that all UNICODE characters (including emojis) be acceptable characters if the user desires to use them [8]. One common barrier to generating strong passwords is the inability of users to see what characters are being typed as they type them, which discourages users from creating longer and more complex passwords [8]. With this limitation in mind, NIST now requires applications to allow users to paste in their password to prevent user typos from interfering with password creation [7]. Another suggested feature from the NIST is the Show Password option which allows users to see their password as they type it, allowing them to check for mistakenly typed characters. Also, the guidelines called for a stop to forcing users to change their passwords periodically in the absence of a breach or other security risk. Studies found that this practice placed an unnecessary mental burden on users, who would grow tired of having to put forth effort to create another secure password and would eventually succumb to poor password creation and management practices [4].

2.1 Related Work

While the NIST has set minimum standards for applications to use for their passwords, these are not strict requirements imposed on every piece of software and many applications do deviate from these guidelines with their own password composition policies. While there are some applications that take stricter approaches by increasing the minimum characters needed or including a visual indication of how strong an entered password is, other applications use loosened security as a way to incentivize users to sign up for their services.

3 STUDY DETAILS

This research replicates a study conducted by Peter Mayer, Jan Kirchner, and Melanie Volkamer [5] which analyzed password requirements from 2010 and 2016, measured minimum password strength of popular websites in America and Germany, and then compared the results. The study was presented at the Proceedings of the Thirteenth Symposium on Usable Privacy and Security in 2017. The goal of the study was to determine how password strength was changing over time and to determine if certain factors such as the function of a website, the existence of viable alternatives, or the presence of advertisements would affect the different password requirements, referred to in the study as Password Composition Policies (PCPs), As all the data is from 2010 and 2016, the PCPs were all gathered from before the release of the new NIST guidelines. The goal of my research is to reproduce the study using modern applications, now that a couple years have passed for developers to comply with the NIST guidelines if they so choose. Then similar to how the original study compared the PCPs of 2016 and 2010, I will compare data from 2020 to 2016.

The Mayer study's [5] data comes from collecting data from a set of the overall most popular websites at the time and then the most popular websites of specific genres, such as banking websites and university websites. I used a slightly different strategy from the Mayer study [5] when choosing which accounts to record data for; I visited websites that I might personally encounter in my daily life, regardless of popularity. This modification was intended to better match the experience of a single digital user, rather than the experience of users overall in order to better determine how these PCPs would affect a user in daily life. To create the accounts I decided to try to use the same password for every single one in order to simplify the process and be able to analyze the differences between what was required of my passwords with an actual example. After actually creating accounts for all of my selected websites, I went back to the Mayer study [5] and recorded data from some of the websites that I had not thought of but were familiar to me or that I might possibly visit in the future, replicating edge case scenarios where I visit a novel website. All the websites I decided upon were separated into categories based on function, except for the latter addition of websites taken from the Mayer study [5] which were all grouped into one category. The categories were email, social media, retail, online games, video entertainment, gig economy, and miscellaneous. The email category was not intended to be a large source of data, rather a necessity due to email being almost ubiquitously required when creating a new account for any platform. The motivation behind this division was from the original study's intention to determine whether or not the function of a website would affect the PCP. One major component that my research leaves out from the study is the comparison of PCPs to a set of German websites. This omission is due to the fact that I would not be able to read password requirements if they were described in German, thus imposing a language barrier on any potential data collection.

3.1 Procedure

The first step was creating accounts on all the websites in order to collect data on their PCPs. To collect the data, I went through each category and created a new account for every website, each time taking note of the minimum password requirements and creating as minimal a password as possible, in the sense that I would only use stronger password creation techniques if they were explicitly required. From the start, my chosen password was "pjlmfmciast", an acronym for "Please just let me finish my capstone I am so tired." If fewer characters were required, I would only include as many as necessary. As I encountered stricter requirements, I added the rules to capitalize the first letter, to append the number 2 at the end, or to append the @ symbol at the end if any of these rules were necessary to create the account. If there was a choice between any of the three or multiple, I would use the strategies above in the order listed for consistency.

4 RESULTS

Originally, 52 websites were selected for accounts to be made, however 3 were excluded from the data collection because their password composition policies were unreachable and 10 more were added from taking from the Mayer study [5], leaving us with a total of 56 data points. Of these 56 websites, 20 started off with instructions on how to create a password that satisfied their specific requirements, while the rest only offered that information when the entered password failed to meet their hidden requirements. Only 14 (25%) of the visited websites had a policy with a minimum character count of eight and did not require a special character or number, thus complying with the NIST guidelines. The lowest number of minimum characters allowed was from Netflix with a count of four and the highest minimum was from Yahoo at nine characters. With the maximum being Yahoo's nine character minimum, it would seem as if websites are hesitant to require any more characters for their passwords, in order to avoid inconveniencing potential new users. One service, Slack, did not require a password at all. It solely relied on a security code that was sent to the entered email address. 20 out of 46 accounts initially displayed the website's password requirements (the last 10 accounts were excluded from this statistic because they were not part of the original data collection, so this particular data for those 10 is unknown). This information matters because a study by Yıldırım and Mackie [10] explored this concept by asking users to create a password when given a minimum character length and advice on multiple ways to construct a secure and memorable password. The study found that just given these suggestions, users were able to make passwords of greater strength compared to their counterparts. A majority of these 20 websites were from the retail and online games genre, while none of them came from the social media genre. Further research would have to be done in order to determine if there are underlying causes for this relation or if it is simply a coincidence.

Five password composition policies made up a majority of the data. The most common set of requirements was an eight character minimum and at least one uppercase letter, resulting in the password of "Pjlmfmci" being used a total of 15 times, or 26.8% of the time. With 14 occurrences was the NIST compliant set of requirements with the password of "pjlmfmci." At 12 occurrences was both the eight character minimum and one number, as well as the eight character minimum, one number and a capital letter, thus resulting in the passwords "pjlmfmc2" and "Pjlmfmc2." The fifth most popular requirement set occurred 11 times and was a 6 character minimum with the password "pjlmfm". All other sets of password requirements were significantly scarcer. Overall, I ended up having 15 unique passwords. Off the top of my head, I have no recollection of which passwords belong to which websites, so if I wanted to access one of these accounts, I could potentially have to cycle through 15 different passwords before finally inputting the correct one. Based on the findings, there are five most likely PCPs, so users could be expected to try passwords up to five times before gaining access to their account if they do not successfully recall the password on the first try. This poses an inconvenience to users who simply want to be able to log into whatever software they want without the hassle of having to piece together what their password could be. A study by Chiasson et al. [1] tried to determine the memorability of different types of passwords. After waiting a short period of time, 68% of participants were able to successfully recall a recently created password on the first try. This statistic leaves over a quarter of users with the process of having to cycle through passwords until they find the correct one, which results in 88% of participants with access to their account [1]. Oftentimes, the participants would get confused and mix up passwords they made for other accounts because there were not distinguishing factors between the passwords they created, although some did use the same password for multiple accounts. Given that a user will likely need a password at least eight times every day, this inconvenience reduces the usability of every application and reduces quality of life for the users.

Out of six functional categories, only two demonstrated any sort of PCP trend. Out of 10 social media accounts, only Facebook required any special password rules such as uppercase letters or special characters. Also, all four of the gig economy accounts had the exact same PCP, requiring a minimum of eight characters and no additional rules beyond that. A similar trend was observed with other applications that could be associated with encouraging rash consumer decisions or purchases that could be made from temptation, such as Amazon, Innisfree, and Sephora. These three retail options had some of the lowest minimum character counts of six, five, and six, only being beaten by Netflix which had a minimum requirement of four. Although this is a small sample size, these lax policies may be due to not wanting to discourage users from creating an account to order something online, thus increasing sales. The miscellaneous category was not included in this analysis because the relation between the accounts in that category was not related to the purpose of the application.

One measurement of password strength used by the Mayer paper [5] was the mathematical strength of the password composition policy. This value is calculated by finding the log base 2 of the size of the charset required by the password and then multiplying this value by the minimum number of characters required. By applying the same formula to our data, we can see that the average PCP has gone down compared to the 2016 data. The average in 2010 was 35.7 and the average in 2016 was 41.4. The calculated average for my dataset is 38.6 and while this would imply that the systems of today are more vulnerable than before, this is not necessarily the case. This formula does not take into account that users are inconvenienced by having requirements of special characters and numbers and case sensitivity, which makes them more likely to create a weak password [1]. A study by Chiasson et al. [1] has shown that these restrictions limit user creativity in creating a password, resulting in strategies such as simply adding the required extra characters onto the beginning or the end of another password. This simple strategy is well-known to attackers, which narrows down the number of possible combinations they need to try, ultimately resulting in a more easily guessed password [4]. An optimal password would have a relation to the application that is not commonly understood, but is understood by the user and have the whole password mean something significant to the user in order to rely on cued recall, rather than uncued recall, as cued recall has been proven to be a more reliable method of remembering information [1]. The Yıldırım and Mackie study [10] also discovered that memorability of the password on the first try increased significantly. The most important finding was that almost all of the participants said that they found this technique useful and would use it in the future. Converting users to start using practices such as these in order to make stronger passwords is crucial to ensuring that they not only continue but also help spread the knowledge which increases the speed of adoption. A smaller and more immediately attainable feature that helps users make stronger passwords is the presence of a password strength meter, as demonstrated in the Yıldırım and Mackie study [10]. These meters encourage users to create stronger passwords, however too strict of a meter will discourage users due to the increased inconvenience of having to come up with an even more elaborate or longer password. Since the common thought is that complex and less memorable passwords are harder to crack [4], users who still wish to make a password given a strict password meter will create something less memorable, potentially influencing them to take other measures to remember the password, such as writing it down or repeating a password of another account. The goal of this research is to open up the space of password creation to allow users to have the freedom to create more complex passwords that are more memorable to them, in turn strengthening their passwords.

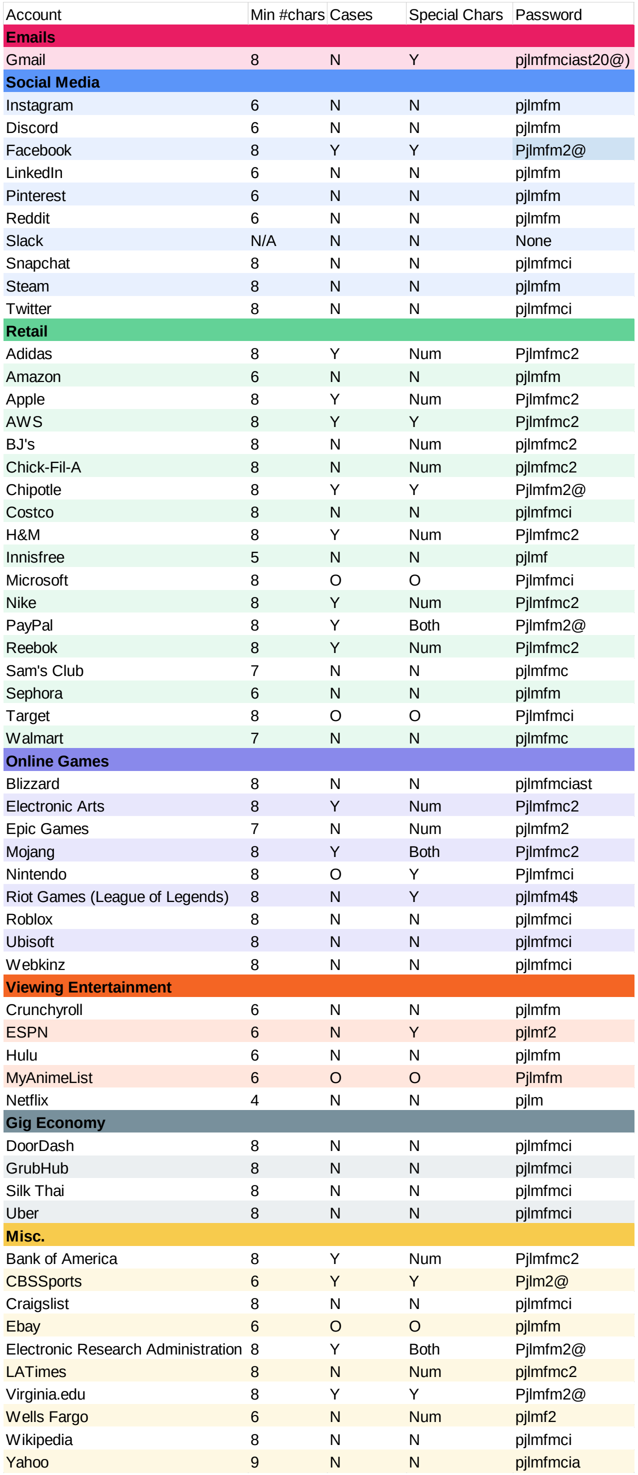
5 CONCLUSION

I replicated a research experiment which looked at different password creation rules and compared them to their counterparts from six years ago. With up-to-date data and a new set of guidelines published for all developers to use if they wished, my research demonstrated that using the definition of password strength from this four-year-old study does not accurately reflect how uncrackable a password is. This formula which is used by Mayer et al. [5] is out of date because the paradigm of password strength has shifted. The idea of using special characters and numbers to increase the maximum possible combination of chars that a hacker would need to cycle through is powerful, but led to users often making weaker passwords as a result of the added complexity [8]. The new strategy of password creation is to create longer passwords that users are still able to recall without the aid of external tools, therefore removing the need to protect another system that could be holding the password. The solution to this problem would be to follow the NIST guidelines and then provide helpful and specific tips to users as they are making their password to aid them in more easily creating a strong password. The most effective method would be to provide several detailed descriptions of how to transform any phrase or idea into a secure and unique password as shown by the Luna study [4]. Compared to the vague restrictions of including a special character and a number, users can make a password more memorable and unique to them, since special characters are not naturally involved in human life. Another powerful feature is the password strength meter, which informs users if their passwords are not actually as secure as they think they are [10]. Simply letting the user know that their password is not secure is enough to convince them that they should pick a new one until they reach a secure password they are happy with. Even if there is no meter, informing users the importance of creating a strong password can convince them to create something stronger [10].

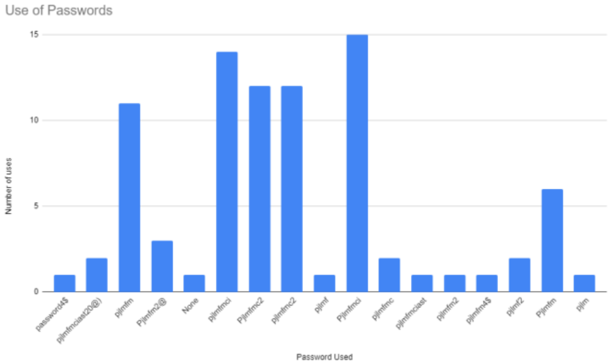
These techniques are all easy to implement and have powerful effects in strengthening the systems they are put in. Users are always willing to secure their information, especially in this digital age where privacy is a highly contested value in society. Simply informing users of how to protect themselves goes a long way in password security. The NIST guidelines help to build this possible scenario by recommending that applications lift their outdated requirements, however it seems that the stigma against plain text passwords has not yet been overcome in the three years since the release of these new guidelines. Moving forward, hopefully these requirements are replaced with the alternative methods described in this paper in order to better protect everyone's data.

6 FUTURE WORK

For future work, you could expand the set of websites to include more genres, since I only selected genres that applied personally to me. An additional step that could have been taken would be to have a social security number, credit card, phone number, and address ready to use to sign up for accounts. In order to protect my personal information, I chose not to disclose any of these pieces of information to any website for this project. Of the X websites visited, Y did not allow users to create a new account without at least one of these pieces of information. In some cases, this restriction prevented the collection of data due to never reaching the password creation step. Encountering this barrier removed potential genres of websites that could be visited for data, such as bank accounts or some university logins. With these paths closed off to me, the diversity of the types of websites that I could visit was limited, possibly skewing the results. Also as described in the Results section, another branch of research that could be done would be to investigate whether certain types of websites are more likely to provide their password requirements from the start or it is simply coincidence. One hypothesis would be that websites involving the exchange of money are more likely to speed up the password creation process in order to minimize the overhead that a user faces before they could potentially change their mind on purchasing a product.



Account	Min #chars	Cases	Special Chars	Password
Emails				
Gmail	8	N	Y	pjlmfmciast20@)
Social Media				
Instagram	6	N	N	pjlmfm
Discord	6	N	N	pjlmfm
Facebook	8	Y	Y	Pjlmfm2@
LinkedIn	6	N	N	pjlmfm
Pinterest	6	N	N	pjlmfm
Reddit	6	N	N	pjlmfm
Slack	N/A	N	N	None
Snapchat	8	N	N	pjlmfmci
Steam	8	N	N	pjlmfm
Twitter	8	N	N	pjlmfmci
Retail				
Adidas	8	Y	Num	Pjlmfmc2
Amazon	6	N	N	pjlmfm
Apple	8	Y	Num	Pjlmfmc2
AWS	8	Y	Y	Pjlmfmc2
BJ's	8	N	Num	pjlmfmc2
Chick-Fil-A	8	N	Num	pjlmfmc2
Chipotle	8	Y	Y	Pjlmfm2@
Costco	8	N	N	pjlmfmci
H&M	8	Y	Num	Pjlmfmc2
Innisfree	5	N	N	pjlmf
Microsoft	8	O	O	Pjlmfmci
Nike	8	Y	Num	Pjlmfmc2
PayPal	8	Y	Both	Pjlmfm2@
Reebok	8	Y	Num	Pjlmfmc2
Sam's Club	7	N	N	pjlmfmc
Sephora	6	N	N	pjlmfm
Target	8	O	O	Pjlmfmci
Walmart	7	N	N	pjlmfmc
Online Games				
Blizzard	8	N	N	pjlmfmciast
Electronic Arts	8	Y	Num	Pjlmfmc2
Epic Games	7	N	Num	pjlmfm2
Mojang	8	Y	Both	Pjlmfmc2
Nintendo	8	O	Y	Pjlmfmci
Riot Games (League of Legends)	8	N	Y	pjlmfm4$
Roblox	8	N	N	pjlmfmci
Ubisoft	8	N	N	pjlmfmci
Webkinz	8	N	N	pjlmfmci
Viewing Entertainment				
Crunchyroll	6	N	N	pjlmfm
ESPN	6	N	Y	pjlmf2
Hulu	6	N	N	pjlmfm
MyAnimeList	6	O	O	Pjlmfm
Netflix	4	N	N	pjlm
Gig Economy				
DoorDash	8	N	N	pjlmfmci
GrubHub	8	N	N	pjlmfmci
Silk Thai	8	N	N	pjlmfmci
Uber	8	N	N	pjlmfmci
Misc.				
Bank of America	8	Y	Num	Pjlmfmc2
CBSSports	6	Y	Y	Pjlm2@
Craigslist	8	N	N	pjlmfmci
Ebay	6	O	O	pjlmfm
Electronic Research Administration	8	Y	Both	Pjlmfm2@
LATimes	8	N	Num	pjlmfmc2
Virginia.edu	8	Y	Y	Pjlmfm2@
Wells Fargo	6	N	Num	pjlmf2
Wikipedia	8	N	N	pjlmfmci
Yahoo	9	N	N	pjlmfmcia



REFERENCES

[1] Chiasson, S., Forget, A., Stobert, E., van Oorschot, P. C., & Biddle, R. (2009). Multiple password interference in text passwords and click-based graphical passwords. Proceedings of the 16th ACM Conference on Computer and Communications Security - CCS ’09, 500. https://doi.org/10.1145/1653662.1653722

[2] Grassi, P. A., Fenton, J. L., Newton, E. M., Perlner, R. A., Regenscheid, A. R., Burr, W. E., Richer, J. P., Lefkovitz, N. B., Danker, J. M., Choong, Y.-Y., Greene, K. K., & Theofanos, M. F. (2017). Digital identity guidelines: Authentication and lifecycle management (NIST SP 800-63b; p. NIST SP 800-63b). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-63b

[3] Kumar, S. (2017, April 9). How to store a password in database? GeeksforGeeks. https://www.geeksforgeeks.org/store-password-database/

[4] Luna, K. (2019). If it is easy to remember, then it is not secure: Metacognitive beliefs affect password selection. Applied Cognitive Psychology, 33(5), 744–758. https://doi.org/10.1002/acp.3516

[5] Mayer, P., Kirchner, J., & Volkamer, M. (2017). A Second Look at Password Composition Policies in the Wild: Comparing Samples from 2010 and 2016. https://www.usenix.org/system/files/conference/soups2017/soups2017-mayer.pdf

[6] Nakov, S. (2019). Secure Hash Algorithms. Practical Cryptography for Developers. https://cryptobook.nakov.com/cryptographic-hash-functions/secure-hash-algorithms

[7] National Institute of Standards and Technology. (n.d.). [Text]. NIST. Retrieved November 6, 2020, from https://www.nist.gov/

[8] NIST Password Guidelines. (2019, November 18). Solarwinds MSP. https://www.solarwindsmsp.com/blog/nist-password-standards2

[9] Swinhoe, D. (2020, August 5). What is a dictionary attack? And how you can easily stop them. CSO Online. https://www.csoonline.com/article/3568794/what-is-a-dictionary-attack-and-how-you-can-easily-stop-them.html

[10] Yıldırım, M., & Mackie, I. (2019). Encouraging users to improve password security and memorability. International Journal of Information Security, 18(6), 741–759. https://doi.org/10.1007/s10207-019-00429-y