International Journal of Data Science and Advanced Analytics (ISSN: 2563-4429)

Password Generator and Storage

Goh Wei Qian[1], Intan Farahana Binti Kamsin[2], Zety Marlia Zainal Abidin[3], Hemalata Vasudavan[4]
[1,2,3,4]Asia Pacific University of Technology and Innovation, Technology Park Malaysia, Bukit Jalil, Kuala Lumpur, Malaysia
[1]tp060087@mail.apu.edu.my, [2]intan.farahana@staffemail.apu.edu.my, [3]zety@staffemail.apu.edu.my,
[4]hemalata@staffemail.apu.edu.my

***Abstract*— In the era of social web, most of the people have their own social media account, and some even have multiple accounts. They need to maintain each of the username and password, which is not easy for everyone. So, they will use the same username and password or select the very weak password to remember easily. To solve this serious problem, this research proposes a system that can generate and store the password. The users only need to remember one master password to manage all passwords. This gives users more confidence to create strong passwords and change passwords frequently without forgetting. This research will use stratified methods for sampling and questionnaires for data collection. In conclusion, the system will allow users to use social media in a safer environment. In future recommendation, we hope that in the future most social media will adopt biometric passwords instead of current text passwords.**

***Index Terms*— Hashing, Encrypt, Two-factor Authentication, Password Strength.**

1. Introduction

As the amount of information stored online continues to grow, social media security is more important than ever. The Internet has made the world more connected. In most cases, this is a good thing. But all these connections also create unprecedented access to information about people and businesses. This can turn into a very bad thing when hackers and scammers get involved. Despite social media giants such as Facebook and Twitter doing their best to protect customer information, hacking still occurs frequently. Even if it is not a completely hack, social media can still help criminals steal other people's identities through shared information. Social media becomes a major channel for cybercriminals. On July 15, 2020, a number of celebrities including Elon Musk, Bill Gates and so on, post a tweet that doubled the price to collect Bitcoin (within half an hour, send Bitcoin to an account and they will pay back double the amount). Within half an hour, the account received bitcoin transfers with a total value of over $100,000 [2]. People realize that this is the biggest account hacking incident that has ever happened on a social media platform. Many users use the same passwords on multiple sites, so if an attacker gets a user's social media passwords, it is likely that they also have passwords for financial sites, including bank accounts, retirement accounts, and credit card accounts [26]. Giving up the use of social media is not a reasonable option. While social networks aren't always safe and secure, there are a few safeguards the users can take to protect them and the organization from some of the most common social media security threats. In this proposal, we propose a solution to protect social media by strengthening passwords and store the passwords. Therefore, it is recommended to use a different password for each website.

2. Literature Review

2.1. Research Domain

2.1.1. Hash algorithm in password store

Password plaintext storage in the database faces many threats, including the application level, the database level, the operating system level, the computer room level, and the employee level [10]. It is very difficult to prevent 100% from being stolen by attackers.

AES, which is the most common symmetric encryption algorithm. This algorithm requires the keys used for encryption and decryption to be stored properly. Although keys are generally stored separately from user information, and there are some mature software or hardware based key storage solutions in the industry, it is impossible to ensure those keys are 100% secure [11][13]. Therefore, AES is not a good way to store passwords, because once the key is leaked, the user's plaintext password is also leaked.

SHA-256, which is a 256-bit long hash value generated by a hash algorithm, called a message digest. Attackers can use lookup tables or rainbow tables to crack user passwords [8][25]. Note that it is cracking passwords, not cracking SHA-256. The reason passwords can be cracked is that user passwords often require brain memory and manual input, so they are not too complicated, and often have a limited length. For example, an 8-digit password has only 8^{10} possibilities. Attackers can calculate the SHA-256 of 0-99999999 in advance and store as a lookup table. When a given SHA-256 needs to be cracked, it can be queried from the table [4]. If the number of passwords has 10 digits, and includes numbers and letters, there are a total of $(10 + 26 + 26)^{10}$ possibilities. Because the data is too large, attackers will use rainbow tables to crack [1][6][33].

Salt, which is a randomly generated string and is added to each key during the hashing process. It changes from the original HASH(key) to HASH(key + salt). Salt prevents attackers from using precomputed rainbow tables, as different salt can result in different hash value even if the password is the same [14][20]. Bcrypt, which is a cryptographic hash designed based on the eksblowfish algorithm. It can dynamically adjust the cost (the number of iterations) to adjust the calculation speed, so even if the computer power continues to increase in the future, it can still resist brute force attacks [18][22].

Therefore, this research will focus on the Bcrypt algorithm because it satisfies two most important conditions. First, the algorithm needs to be irreversible, so as to effectively prevent

password leakage. Second, the algorithm needs to be slow, and slowness is an effective way to deal with brute force cracking.

2.1.2. Use of insecure password

In the Internet age, people need to deal with a variety of passwords, including payment passwords, social media passwords, e-commerce platform passwords, and so on. Password is the biggest barrier to protect user security and privacy [31].

Position	Password	Time to crack it	Number of users
1 ▲ (2)	123456	< 1 sec	2,543,285
2 ▲ (3)	123456789	< 1 sec	961,435
3 • (New)	picture1	3 hrs	371,612
4 ▲ (5)	password	< 1 sec	360,467
5 ▲ (6)	12345678	< 1 sec	322,187
6 ▲ (17)	111111	< 1 sec	230,507
7 ▲ (18)	123123	< 1 sec	189,327
8 ▼ (1)	12345	< 1 sec	188,268
9 ▲ (11)	1234567890	< 1 sec	171,724
10 • (New)	senha	10 sec	167,728

Figure 1: List of the 2020 most commonly used passwords [15]

In 2020, the network security company NordPass counted nearly 275.7 million passwords on the Internet and released a list of the 2020 most commonly used passwords TOP200 [15]. These passwords all have a common feature, that is, most of them can be cracked within 10 seconds, and even the newly listed “picture1” will not take more than 3 hours to crack. In addition to these simple numbers or letter combinations, passwords related to personal information such as birthdays, mobile phone numbers and addresses are also minefields for setting passwords.

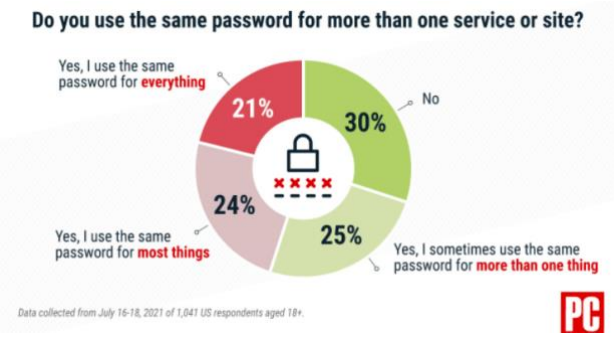

Figure 2: Survey of using same password [9]

In 2021, PCMag conducted a new survey of 1,041 U.S. residents aged 18 or older. The findings show that 70% of adults are still using the same password for more than one thing [9]. Using reuse passwords is very dangerous. When a data breach occurs or a password for a website is cracked, an attacker can use the password to try on multiple websites and cause unimaginable losses [12][24][26].

Therefore, this research will avoid the use of these insecure passwords, in order to prevent criminals from “credential stuffing” or being directly cracked, which will bring danger to privacy and property security.

2.1.3. Pseudorandom password generate

True random numbers produce unpredictable and invisible results [23]. The random function in the computer is replicated using a specific algorithm, with a clear and apparent output. As a result, the "random number" generated by the computer's random function is a pseudorandom number rather than a true random number [30].

Middle square method, which is a pseudorandom number generation method. In 1946, while analyzing neutron collisions at the Los Alamos laboratory, S. Ulm, N. Metropolis and John Von Neumann proposed this method. In practice, this is not a good method as it is usually short-cycle and has some serious drawbacks. When repeated many times, this method will start generating the same number repeatedly or looping to the previous number in the sequence and looping indefinitely [3][28].

Linear Congruential Generator was proposed by D.h.Lehmer in 1949. The characteristic of using this method to generate pseudorandom numbers is that it is very easy to implement and the generation speed is fast. However, the disadvantages are also obvious. The maximum period of 32-bit numbers can only be up to 2^{32}, which cannot meet the requirements of applications that require high-quality pseudorandom numbers [5][21].

Blum Blum Shub was proposed by Lenore Blum, Manuel Blum and Michael Shub in 1986. This method has high computational complexity and can be effectively applied to encryption tasks with low key generation rate [32]. It also can slow down dictionary and brute force attacks and can adaptively adjust the guaranteed minimum time to generate keys for specific authentication and encryption tasks [16].

Therefore, this research will focus on the Blum Blum Shub pseudorandom generator because it satisfies two kinds of unpredictability. First, forward unpredictability. Without knowing the seed, no matter how many bits in the sequence are known, the next bit cannot be predicted. Second, backwards unpredictability. The seed value cannot be inferred from any value produced.

2.2. **Similar Systems**

2.2.1. Chrome Password Manager

The chrome password manager can automatically handle all passwords entered by the user and save them.

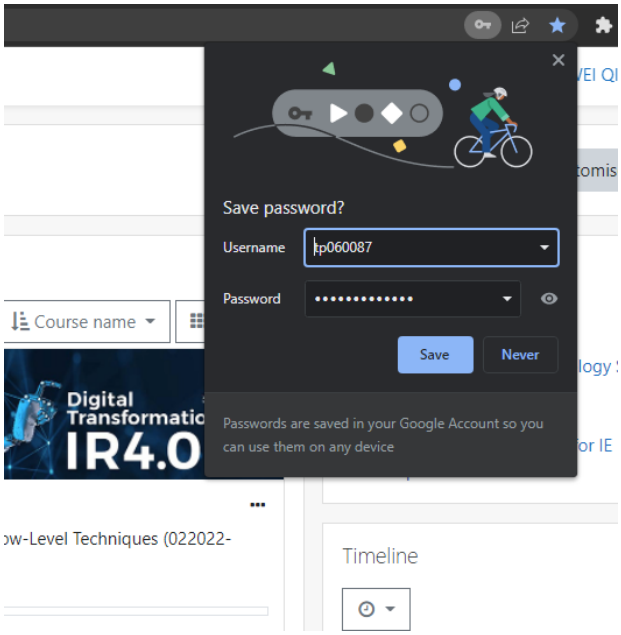

Figure 3: Prompt of saving password in Chrome

Every time a user visits a new website that requires login details, the chrome browser automatically asks if they want to save the password. When the user visits these pages again, the chrome browser will automatically fill in the username and password, so the user does not need to do anything.

Chrome password manager very convenient and user friendly. When this feature is turned on, chrome browser will automatically store and fill in user account username and password the next time user visits these sites.

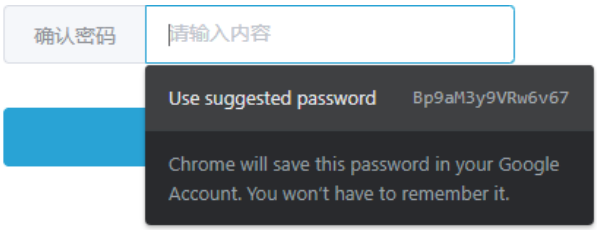

Figure 4: Suggested password of Chrome

Chrome password manager also have useful password generator features. Besides that, user's account details are automatically synced as long as the same browser is used. Best of all, this service is completely free.

However, chrome password manager has limited security features and functionality. This password manager will encrypt the passwords, but it cannot determine whether the passwords need to be strengthened. Also, this password manager does not have the option to add a master password to increase the level of security. This password manager also does not work in multiple browsers.

2.2.2. Vov Password Generator

Vov Password Generator is an excellent password generator that quickly generates passwords for user's accounts using any character set. The features of Vov Password Generator are fast installation and no configuration required, so the user can start generating passwords right away.

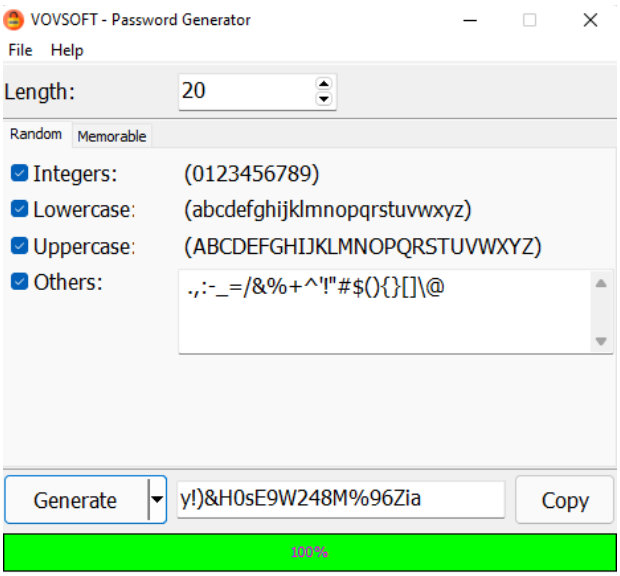

Figure 5: Vov Password Generator interface

The user can choose the length of the generated password and the characters to include. But the biggest drawback of this generator is that it cannot store the passwords, so the user must record these passwords separately.

2.3. **Comparison table**

	Chrome Password Manager	Vov Password Generator
Password generator	√	√
Custom password type	×	√
Check password strength	×	√
Password storage	√	×
Password storage encryption	AES-256	×
Autofill	√	×
Master key	√	×
Two-Factor authentication	×	×
Vault timeout action	×	×
Platform	Chrome	Windows, Android
Paid plan	×	×

Table 1: Comparison Table of Similar Systems

The table shows that both systems have password generators, but the chrome password manager does not have the ability to customize password types and check the password strength. The Chrome password manager can store the password and password storage encryption methods is AES-256. It also has autofill and need to use master key to log in. But both systems did not have two-factor authentication and vault timeout action. The Chrome password manager is a web-based system and only support the chrome browser. The Vov password manager is a desktop/mobile application and support the windows and android system. Both systems are free.

2.4. **Conclusion**

In conclusion, there is a limitation based on the previous similar systems. Therefore, this research will develop a system that can generate and store passwords, allow users to customize password types and automatically check password strength. Then, store the generated password using the Brcypt algorithm. When logging into the system, the user must use the master key and two-factor authentication. The system will automatically log out at the time set by the user to ensure the security of the account.

3. **Problem Statement**

A brute force attack is the repeated attempt of different password combinations to find the correct one, to crack a password or username, or to find the key used to encrypt a message. According to the Verizon 2021 Data Breach Investigations Report, hacking is still the main attack vector which include the brute force passwords attack. There is over 89% of breaches caused by hacking, it is included brute force attack or use of lost or stolen credentials [27]. Brute force attacks do not rely on vulnerabilities within a website, but they rely on users with weak or guessable credentials. They unlike many other tactics used by attackers. Brute force attacks very popular because they are simple and many potential targets. Attackers can perform multiple simultaneous attacks to increase their chances of success [29]. As COVID-19 leads to the rise of remote work, many brute force attacks have been attempting to remotely manage Windows systems that using the Windows Remote Desktop Protocol (RDP) [19]. According to the ESET threat report, RDP brute force attacks are the one of the most common attack vectors used to compromise corporate networks. RDP brute force attacks have been growing between 2020 and 2021, especially in the last four months of 2021, from 55 billion in T2 in 2021 to 206 billion in T3 in 2021, a substantial increase 274% [7]. There is another brute force attack occurs in 2018, Magento which is a popular e-commerce platform, its admin panel is compromised. There are over 1000 account credentials found on dark web. The aim of the attackers is to grab account holders' credit card numbers and infect their devices with malware for cryptocurrency mining [17]. Therefore, it is imperative to protect accounts with strong passwords to reduce data breaches, financial losses and other incidents.

4. **Research Aims**

This research aims to propose an effective system to protect the privacy data and property of all users.

5. **Research Objectives**

- To create the function of analyzing strength of the password.
- To create strong password generator to make sure the passwords are strong enough.
- To store passwords for user to prevent forgetting.
- To implement two-factor authentication for login.

6. **Research Significance**

The results of this research will help internet users reduce the use of insecure passwords. The users can generate and store the password in this system. Generated passwords can make it harder for attackers to guess and can also be stored so users would not forget. For researchers, this research will provide some solutions to the use of insecure passwords on social media or other online platforms.

7. **Methodology**

The target users of the system are individuals, including any age group and occupation. Every user with an account and password is a target user of the system. This is because it is not easy to use complex passwords and remember, and even one person can have multiple accounts and passwords at the same time.

The sampling method used in this research is stratified sampling. It divides a population into smaller subgroups called stratification. Stratification is based on common attributes or characteristics of members, such as occupation or age. This research will randomly survey 200 users those have several accounts and passwords and divide them into different subgroups, including gender, age range, ethnicity, and occupational background. A random sample of these subgroups will be drawn and pooled for analysis. This method provides better population coverage, saves a lot of time and cost, and greatly improves efficiency.

This research will use questionnaires to collect data. Questionnaires can not only save manpower, material resources, financial resources and time, but also can survey many people in a short period of time by carried out questionnaires in groups or by mail. Therefore, questionnaires are highly efficient and suitable for computer processing of data, saving the cost and time of analysis. Questionnaires are also unified. This is helpful for comparative analysis of the respondents under the same situations and can also analyze respondents with different consciousness from the society.

The questionnaire will contain about 5 category questions for grouping (age, gender, occupation, race and number of accounts.), 10 to 20 questions to collect opinion data, and 1 to 2 subjective questions to obtain additional opinion data. Before

starting a questionnaire, a pilot test must be conducted. A pilot test can save time and cost by reducing the risk of errors or problems.

Once the questionnaires results have been collected, analyze the data collected from the questionnaires to assess the feasibility of this research.

8. Overview of the Proposed System

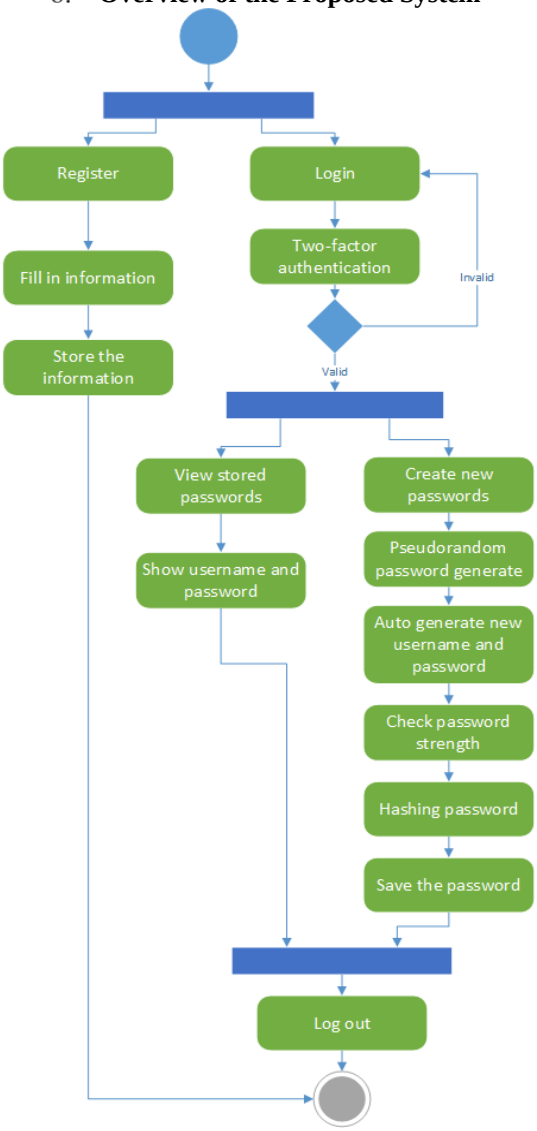

Figure 6: Activity diagram of the proposed system

The activity diagram shows the flow of the proposed system. The user must register an account before they log in. When logging in, the user must use two-factor authentication to prevent others from logging in without access rights. In this proposed system, the user can create new passwords and view the stored passwords.

9. Conclusion

Do not let fear keep you from social media. It is a great way to keep in touch with friends and family, and to see what is going on in the world. Just keep the personal information safe. Never enter anything on social media that you would not write on a public wall (even private information) to prevent identity theft. Remember, it is not how often you use social media, but how you use it. Whether the users use social media monthly or hourly, their identity can easily be stolen if they are not careful.

References

[1] Alpatskiy, M. A., Borzunov, G. I., Epishkina, A. v., & Kogos, K. G. (2020). New Approach in the Rainbow Tables Method for Human-Like Passwords. *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, 2035–2040. https://doi.org/10.1109/EIConRus49466.2020.9039311

[2] BBC. (2020, July 16). *Major US Twitter accounts hacked in Bitcoin scam*. https://www.bbc.com/news/technology-53425822

[3] Budiman, A., Bulolo, E., & Saputra, I. (2020). Middle Square Method Analysis of Number Pseudorandom Process. *The IJICS (International Journal of Informatics and Computer Science*.

[4] Cortez, D. M. A., Sison, A. M., & Medina, R. P. (2020). Cryptanalysis of the Modified SHA256. *Proceedings of the 2020 4th High Performance Computing and Cluster Technologies Conference & 2020 3rd International Conference on Big Data and Artificial Intelligence*, 179–183. https://doi.org/10.1145/3409501.3409513

[5] Cybulski, R. (2021). Pseudo-random number generator based on linear congruence and delayed Fibonacci method. *Technical Sciences*, *24*(2021). https://doi.org/10.31648/ts.7238

[6] Dat, T. N., Iwai, K., Matsubara, T., & Kurokawa, T. (2019). Implementation of high speed rainbow table generation using Keccak hashing algorithm on GPU. *2019 6th NAFOSTED Conference on Information and Computer Science (NICS)*, 166–171. https://doi.org/10.1109/NICS48868.2019.9023888

[7] ESET. (2021). *THREAT REPORT T3 2021*.

[8] Gong, C., & Behar, B. (2018). Understanding password security through password cracking. *Computing Sciences in Colleges, Evansville, United States*, *33*(5), 81–87.

[9] Griffith, E. (2021, September 21). *Stop Using the Same Password on Multiple Sites! No. Really*. The Why Axis. https://www.pcmag.com/news/stop-using-the-same-password-on-multiple-sites-no-really

[10] Hallett, J., Patnail, N., Shreeve, B., & Rashid, A. (2021). *"Do this! Do that!, And nothing will happen" Do specifications lead to securely stored passwords?*

[11] Harba, E. S. I. (2017). Secure Data Encryption Through a Combination of AES, RSA and HMAC. *Engineering, Technology & Applied Science Research*, *7*(4), 1781–1785. https://doi.org/10.48084/etasr.1272

[12] He, X., Cheng, H., Xie, J., Wang, P., & Liang, K. (2022). Passtrans: An Improved Password Reuse Model Based on Transformer. *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3044–3048. https://doi.org/10.1109/ICASSP43922.2022.9746731

[13] Liu, Y., Zhang, W., Peng, X., Liu, Y., Zheng, S., Wei, T., & Wang, L. (2019). Design of password encryption model based on AES algorithm. *2019 IEEE 1st International Conference on Civil Aviation Safety and Information Technology (ICCASIT)*, 385–389. https://doi.org/10.1109/ICCASIT48058.2019.8973003

[14] Merhav, N., & Cohen, A. (2020). Universal Randomized Guessing With Application to Asynchronous Decentralized Brute–Force Attacks. *IEEE Transactions on Information Theory*, *66*(1), 114–129. https://doi.org/10.1109/TIT.2019.2920538

[15] NordPass. (2020). *Top 200 most common passwords*. NordPass. https://nordpass.com/most-common-passwords-list/

[16] Omorog, C. D., Gerardo, B. D., & Medina, R. P. (2018). Enhanced pseudorandom number generator based on Blum-Blum-Shub and elliptic curves. *2018 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, 269–274. https://doi.org/10.1109/ISCAIE.2018.8405483

[17] Pavlović, N., Šarac, M., Adamović, S., & Mravik, M. (2019). Enchantment of Magento CMS Security. *Proceedings of the International Scientific*

Conference - Sinteza 2019, 223–228. https://doi.org/10.15308/Sinteza-2019-223-228

[18] Pervan, B., Knezovic, J., & Guberovic, E. (2022). Energy-efficient distributed password hash computation on heterogeneous embedded system. *Automatika - Journal for Control, Measurement, Electronics, Computing and Communications*.

[19] Pranggono, B., & Arabo, A. (2020). COVID-19 pandemic cybersecurity issues. *Internet Technology Letters*.

[20] Rathod, U., Sonkar, M., & Chandavarkar, B. R. (2020). An Experimental Evaluation on the Dependency between One-Way Hash Functions and Salt. *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1–7. https://doi.org/10.1109/ICCCNT49239.2020.9225503

[21] Sathya K., Premalatha J., Rajasekar, V., Kumar M., M., Deepak M., & Manoj S. R. (2021). *Modified linear congruential generator to secure random number generation*. 140007. https://doi.org/10.1063/5.0068654

[22] Sriramya, P., & Karthika, R. A. (2015). PROVIDING PASSWORD SECURITY BY SALTED PASSWORD HASHING USING BCRYPT ALGORITHM. *ARPN Journal of Engineering and Applied Sciences*, *10*.

[23] Stipčević, M., & Koç, Ç. K. (2014). True Random Number Generators. In *Open Problems in Mathematics and Computational Science* (pp. 275–315). Springer International Publishing. https://doi.org/10.1007/978-3-319-10683-0_12

[24] Suscello, N. (2021). Reducing Risk of Password Reuse through Random Character Requirements and Image Prompting. *Christopher Newport University ProQuest Dissertations Publishing*.

[25] Tatli, E. I. (2015). Cracking More Password Hashes With Patterns. *IEEE Transactions on Information Forensics and Security*, *10*(8), 1656–1665. https://doi.org/10.1109/TIFS.2015.2422259

[26] Wang, K. C., & Reiter, M. K. (2019). How to End Password Reuse on the Web. *Proceedings 2019 Network and Distributed System Security Symposium*. https://doi.org/10.14722/ndss.2019.23350

[27] Widup, S., Pinto, A., Hylender, D., Bassett, G., & Langlois, P. (2021). *2021 Data Breach Investigations Report (DBIR)*.

[28] Widynski, B. (2022). *Middle-Square Weyl Sequence RNG*.

[29] Xylogiannopoulos, K. F., Karampelas, P., & Alhajj, R. (2020). A password creation and validation system for social media platforms based on big data analytics. *Journal of Ambient Intelligence and Humanized Computing*, *11*(1), 53–73. https://doi.org/10.1007/s12652-019-01172-x

[30] Yadav, A. (2013). Design and Analysis of Digital True Random Number Generator. *VCU Scholars Compass*.

[31] 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

[32] Yu, S., Krzysztof, P., Yan, L., Maksymovych, V., Stakhiv, R., Malohlovets, A., & Kochan, O. (2022). Development of Modified Blum-Blum-Shub Pseudorandom Sequence Generator and its Use in Education. *Measurement Science Review*, *22*(3), 143–151. https://doi.org/10.2478/msr-2022-0018

[33] Zhang, L., Tan, C., & Yu, F. (2017). An Improved Rainbow Table Attack for Long Passwords. *Procedia Computer Science*, *107*, 47–52. https://doi.org/10.1016/j.procs.2017.03.054