



DESIGN AND DEVELOPMENT OF A COLLISION RESISTANT MODIFIED SHA1

Esmael V. Maliberan

Department of Information Technology, Surigao del Sur State University, Tandag City, Surigao del Sur, Philippines

E-Mail: malibs_28@yahoo.com

ABSTRACT

The Secure Hash Algorithm is best used in checking data integrity and authenticating digital media due to its fast hash calculating power. However, the SHA1 hashes are highly vulnerable to different attacks most especially collision attack because of the undersized hash length that compromises the integrity and confidentiality of data during transit. Presented in the current study is a new design to extend the hash size of SHA1 to secure a client-server communication through login authentication. Experimental results revealed that on five datasets that have collision in SHA1, none of them collided using the enhanced SHA1. Furthermore, the enhanced SHA1 is more resistant to brute force attack over the original SHA1 in terms of cracking time which is estimated to be sexagintillion years on a standard desktop PC. The results also presented that the produced hash codes of the enhanced SHA1 were resistant against a rainbow table and dictionary attack. By increasing the hash value of SHA1 from 160-bits to 1280-bits using AND and XOR operators, this paper has shown that execution of the enhanced SHA1 in digesting the passwords in login authentication delivers a safe way in a web transaction particularly in web-based environment.

Keywords: brute force attack, client-server communication, hashing, modified SHA1, rainbow table.

INTRODUCTION

Internet security is one of the interested topics in the field of Information Technology, most especially in a web-based environment [1], [2]. In the past few years, encryption and decryption are developing into main interest in data transmission [3]. One solution for preserving the transmitted data is a hash function. It is a portion of varied major Internet security claims [4]. It maintains integrity and confidentiality of the message and therefore these hashing algorithms are utilized mainly for file verification and user authentication [5], [6]. Hashing algorithms are additionally important elements in various security applications and practices [7]. SHA1 generates a 160-bit hash message [8], [9] based on principles of those utilized by Ronald L. Rivest who invented MD4 and MD5. It is the most commonly used hash algorithm, with its wide variety of applications [10], [11], [12] due to its adaptability, power [13], and agility [14]. The algorithm has been made as part of the capstone project of the government of United States and is usually applied to secure web transactions and other requests [15], [16]. While SHA1 is widely used, it does not seem to give adequate distributional avalanche effect of the input differences in the compression function [17], [10] by which this will lead to the probability of producing the same hash code of the two different input inside its compression function [18], [19]. Therefore, an improved diffusion needs to be configured to distribute the output at every stage inside its compression function [18], [20], [21]. The initial collision attack in SHA1 algorithm was attained by researchers, which produced two distinct PDF files. The study of [22] delivered another identical free-start pair for SHA-1 in a collision within its compression function. Its undersize hash value would make it vulnerable to numerous attacks including a rainbow table, a dictionary, brute-force. Therefore, the SHA1 algorithm for client-server communication and login authentication

is no longer safe. Some modifications and improvements were developed to address these problems [23], [24]. [25] Developed a new technique to improve the MD5 algorithm with the compression function of SHA1, which formed 256-bit hash code. On the other hand, [26] proposed a modification to the SHA-1 hash function. Rather than logical functions, the authors used the typical distributed pseudo-random function. These changes produced absolute hash values for original messages and contributed a hash value requirement to prevent brute attack. A number of hash functions were made, many of them were from the MD4 algorithm developed by Rivest [27]. Meanwhile, [28] argued that the enhancement of the SHA-1 was appropriate. According to the author, hackers and cryptanalysts found attacks on SHA-1 in 2005 which meant that the algorithm may not be safe enough to use. [29] the CPSO (Canonical Particle Swarm Optimization) method was used to boost the Stable Hash Algorithm-1. The idea was to forecast the block of control, which apprehend the password and provided the password stream with the same size of a log-list. [30] modified the method and formula to digest or decode the last SHA1 message by XORing, refining and expanding the mathematical formulation. The enhanced SHA-1 conserved its unique rounds covering of 80 rounds and an output of 160-bit hash value. [31] tested many collision attacks on the initial 80-step of the algorithm and proven a thorough adapted form of the procedure that weakened the probability of collision and elevated the time-complexity required to recognize an exponential factor of 2 in the collision. Presented in this paper is the modification of the algorithm by expanding its hash value from its original 160-bit to 1280-bit using the combined XOR and AND operator to prevent known attacks such as collision, dictionary, brute-force and rainbow table.



RESEARCH METHOD

Research Design

Simulations were done using Delphi programming language which was directly linked to MYSQL as database. XAMPP server 2.5 and MS SQL server was used for storing the hash value of user passwords. All processes in this study were conducted on a personal computer running on a Celeron processor with 4Gb of memory and Windows 8, 64-bit operating system.

Five pairs of files that generated identical hash value in the original SHA1 were used for testing and verification. These file pairs are man.jpg and woman.jpg, text1.bin and text2.bin, set.bin and set1.bin, try.exe, and hi.exe, and f1.bak and f2.bak. The SHA1 hashing algorithm follows a Merkle-Damgard Construction. To address the flaws of SHA1, modification to the SHA1 output was made as shown in Figure-1.

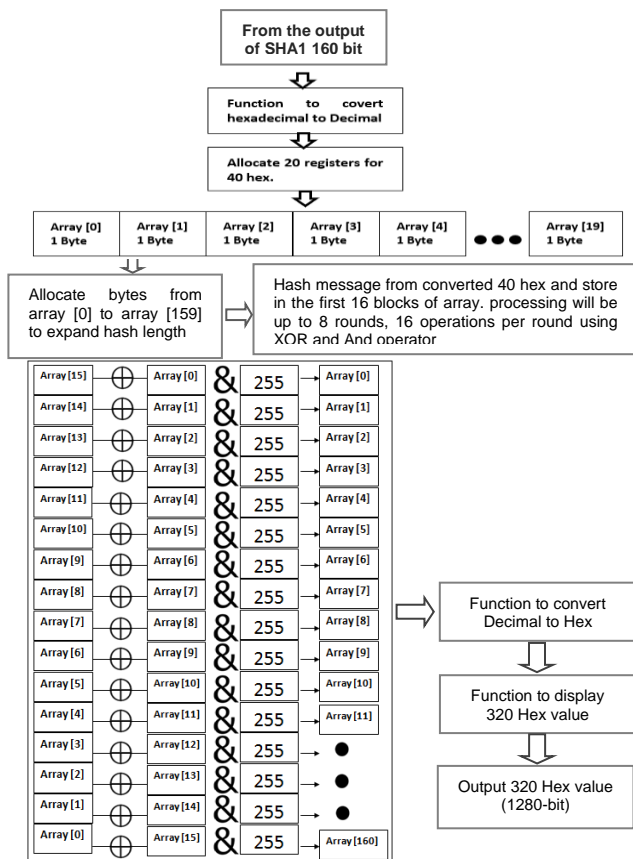


Figure-1. Architecture of the Modified SHA1.

Modification Process

The modification process of the study is shown in Figure-1. When a password is hashed using SHA1, its output will then be converted into a hex value encompassing 40 characters. There will be a function initiated to change this hex value into decimal and assign an array of 20 blocks containing 1 byte for each block. Two (2) characters will be assigned and allowed for every block. Additional function is made to assign an array of 160 blocks with a size of 1 Byte. It is designed to expand

the hash size to 1280 bit as its proposed hash value. The message will be digested or hashed during this stage using XOR and AND operators and automatically stored it in an array of 20 blocks. In designing for modification, XOR was carefully considered due to its best bit shuffling properties [32], [33] which is proposed by many cyphers [34]. This operator is used in most of the cryptographic algorithms which has proven to be highly effective [35]. Moreover, the AND operator was added to 255 decimals since it is a required value of 1 byte to be stored in an array. The process will include up to eight rounds with 16 operations per round. Finally, a function is executed to convert the decimal value to hexadecimal and outputs the 320 hex hash value or 1280-bit hash.

Against Collision Attack

Five datasets that collided in SHA1 were used as input to the program to test whether it still collided in the enhanced SHA1. Every pair of files was uploaded into the simulation program that generates hash value using SHA1 and the enhanced SHA1 as shown in Figure-2.

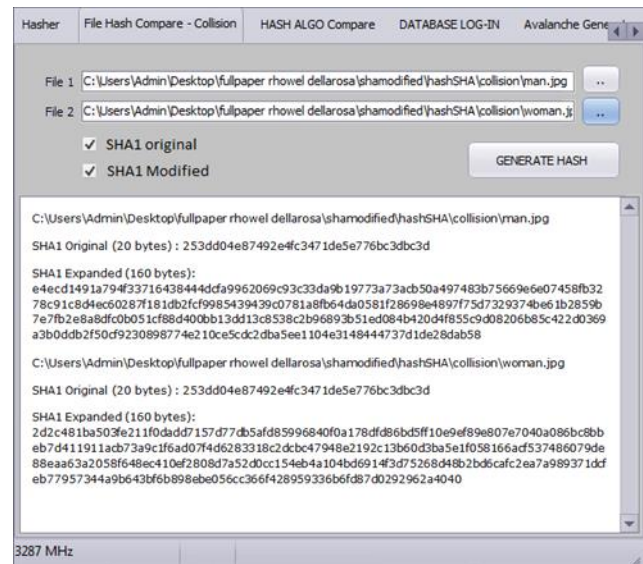


Figure-2. Testing the enhanced SHA1 against collision attack.

Brute Force Attack

To measure the resilience of the modified against brute force attack, a password cracking tool (<https://thycotic.force.com/support/s/article/Calculating-Password-Complexity>) was used and the time needed to crack passwords were recorded. The possible combination based on character set is given by Equation 1. The cracking time, on the other hand, was computed using Equation 2.

$$\text{Possible combinations} = \text{possible number of characters}^{\text{Password length}} \quad (1)$$

$$\text{Cracking time} = \frac{(\text{hash calculation per millisecond} * \text{possible password combination})}{2} \quad (2)$$



Dictionary Attack

To determine resilience against a dictionary attack, an online cracking tool called CrackStation (<https://crackstation.net>) that contains 1,493,677,782 words of possible passwords with equivalent hashes was used. The tool utilizes huge pre-computed lookup tables to break password hashes and these tables store a plot for that particular hash between the password hash and the right password. The hash codes are listed and indexed for searching the database immediately for a given confusion. If the message is found, the password can be obtained in a split second. It is applied only to messes without salt and supports current hash algorithms such as MD4, MD5, sha256, sha512, and others. Ten trials were done using the tool with different hash values generated from the enhanced SHA1 to guarantee precise resilience to a dictionary attack.

Rainbow Table Attack

To test the reliability of the modified SHA1 algorithm against a rainbow table attack, a pre-computed rainbow table was used. The table is made up of a database with a large number of hash function inputs and outputs. The aim of this test was to look for and compare a password with its hash value within the table. The Rainbow Crack tool was used to test a rainbow table attack (www.rainbowcrack.com).

Application of the ENHANCED SHA1

A registration form and a login form were built for a user to register and communicate with the application server using Delphi Language. To use the system, a user was required to register a name, email address, and password which were then submitted to a server. The method is processed by a Delphi script, which then saves it to the database. Each time a user entered his email address and password, the modified SHA1 algorithm hashed the password into a 1280-bit hash length and compared its value to the stored hash value of the user's original password. The user's password was hashed into a 1280-bit length during registration, and stored in one password table in the server database. If the two passwords had the same hash values, the login authentication was verified and the user is given access to the server.

RESULTS AND ANALYSIS

Resistance to a Collision Attack

Five pairs of files were run and tested in a simulated program using the Delphi language. Shown in Figure-3 is the first pair of files run and tested using the program. Based on the experiment conducted, results show that on the first pair of images namely "man.jpg" and "woman.jpg", the said files which collided using SHA1 produced different hash values when hashed by the modified SHA thus preventing collision attack. Shown in Figure-4 is the second pair of files run and tested using the simulation program.

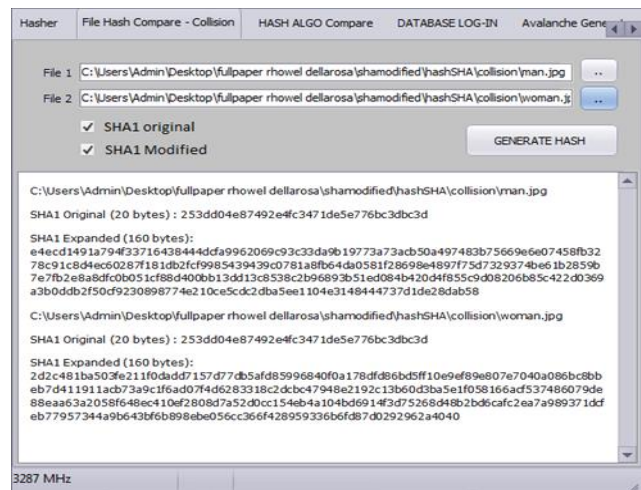


Figure-3. Result for testing a collision attack.

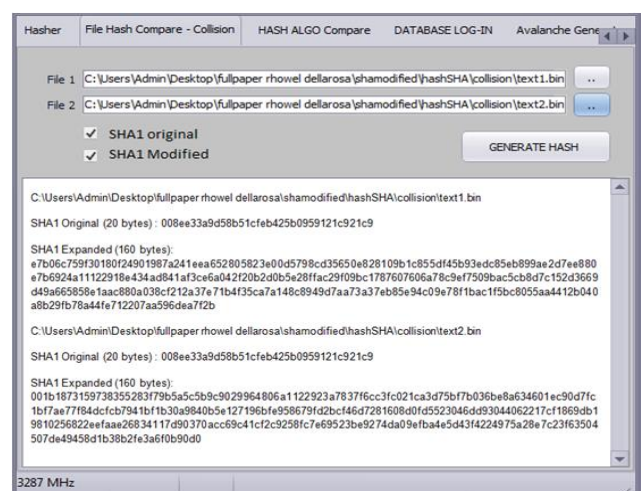


Figure-4. Result in testing collision resistance on text1.bin and text2.bin.

On the second pair of binary files namely "text1.bin" and "text2.bin," results indicate that the said files generated different hash code using enhanced SHA1 to prevent collision attack. Unlike SHA1, the modified SHA1 is collision-resistant. Shown in Figure-5 is the result of the simulation on the third pair of files. It can be observed that the results showed that on the third pair of binary files namely "set.bin" and "set1.bin", the system produced different hash values using the modified SHA1. Another test was done on the two different executable files namely the "try.exe" and "hi.exe." as shown in Figure-6.

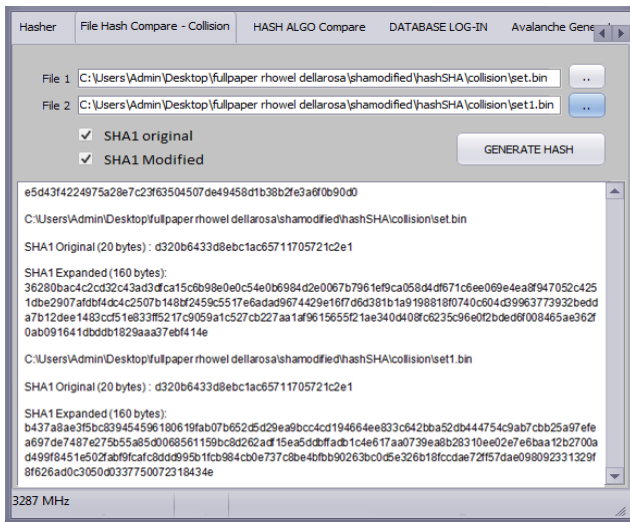


Figure-5. Result on collision attack on set1.bin and set2.bin.

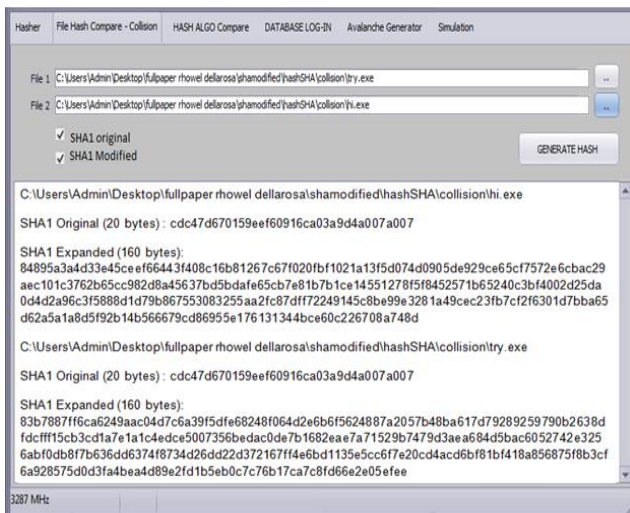


Figure-6. Result on collision attack on try.exe and hi.exe.

It can be noted that the results showed that the two executable files collided in SHA1 while it did not collide using the modified SHA1. The last test was done on two different backup files namely “f1.bak” and “f2.bak”. The results are shown in Figure-7. Consistent with the previous results, the system for modified SHA1 produced different hash values for the two files. Shown in Table-1 is the summary of the test results for the collision attack for the five pairs of files that were used during the experimentation.

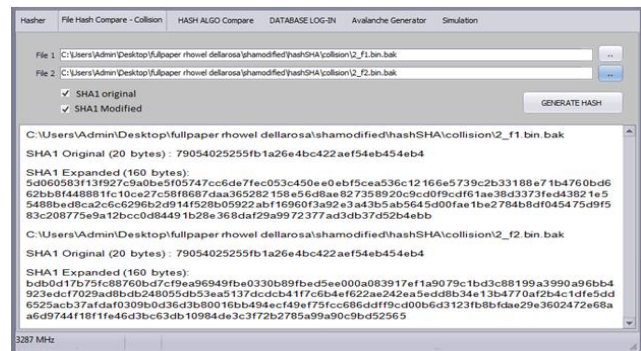


Figure-7. Result in testing collision attack.

Table-1. Experimentation result for collision attack.

Pair of Files	SHA1 Algorithm	
	Original	Modified
man.jpg and woman.jpg	Successful	Failed
text1.bin and text2.bin	Successful	Failed
set1.bin and set2.bin	Successful	Failed
try.exe and hi.exe	Successful	Failed
f1.bak and f2.bak	Successful	Failed

It should be interpreted from Table-1 that when testing each pair of files for collision resistance, the same hash values were produced by the original SHA1 while this was not the case for the enhanced SHA1 algorithm. The enhanced SHA1 has thus been able to resolve the collision attack.

Resistance to a Brute-force Attack

Resistance to brute force attack was done using the password checker online tool. The UI for this tool is shown in Figure-8 which also gives details of the program.

Brute-force attack cracking time estimate

Machine	Time
Standard Desktop PC	About 62 sexagintillion years
Fast Desktop PC	About 16 sexagintillion years
GPU	About 6 sexagintillion years
Fast GPU	About 3 sexagintillion years
Parallel GPUs	About 3.104953534412724e+29 quinquagintillion years
Medium size botnet	About 6.209907068825447e+25 quinquagintillion years

Figure-8. UI for password checker online to crack the hash code.

The output hash value of the modified SHA1 algorithm was tested; it was revealed that it would take about 62 sexagintillion years to crack. This length of time renders the modified SHA1 virtually impossible to crack using currently-available cracking tools, unlike the original SHA1.



Resistance to Dictionary Attack

To assess how robust, the enhanced SHA1 is against a dictionary attack, CrackStation was used. The UI for the system, which also shows a result for the dictionary attack, is shown in Figure-9.

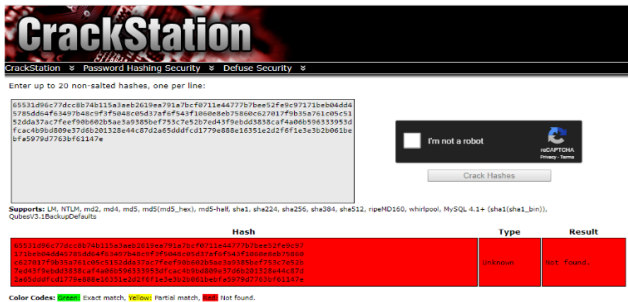


Figure-9. Crack Station UI.

For the input of the plaintext “admin”, results showed that the tool was unable to generate a plaintext hash code generated by the enhanced SHA1 in its database. The tool was unable to recognize the hash format, and its type was either unknown or unidentified in its database. This is due to the fact that the method only works with well-known hashing algorithms like MD5, SHA256, SHA512, and MD160.

Resistance to Rainbow Table Attack

As shown in Figure-10, another attack was used to check the reliability of the modified SHA1 algorithm using a rainbow table attack.

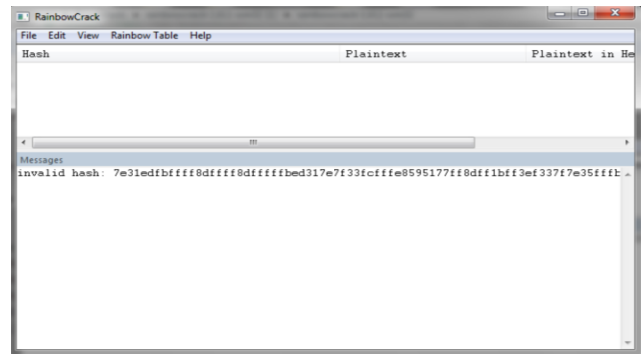


Figure-10. Cracking the hash code using a rainbow table attack.

Results indicated that the tool failed to crack the "admin" hash value that is 1280-bit created from the enhanced SHA1. This can be due to the fact that only hashes for known hash algorithms such as MD5, SHA2, SHA256, and SHA512 are in the database. Moreover, the same hash codes are used during a dictionary attack to verify that the modified SHA1 is resilient to a rainbow table attack.

Comparison of the Results of SHA Family on Collision Attack

A comparison of the different SHA family along with the proposed enhancement is significant because it allows other readers and researches to appreciate the features of the proposed study that is not present in the existing SHA family. Table-2 compares improved SHA1 to SHA1, SHA256, SHA384, and SHA512 in terms of computing pre-images attack, second pre-image attack, and heavy collision attack.

Table-2. Comparison Results of different SHA algorithms and the enhanced SHA1 on Collision Attack.

Name of the attack	SHA1	SHA256	SHA384	SHA 512	This study
Pre-image attack	2 ¹⁶⁰	2 ²⁵⁶	2 ³⁸⁴	2 ⁵¹²	2 ¹²⁸⁰
Second Pre-image attack	2 ¹⁶⁰	2 ²⁵⁶	2 ³⁸⁴	2 ⁵¹²	2 ¹²⁸⁰
Strong Collision attack	2 ⁸⁰	2 ¹²⁸	2 ¹⁹²	2 ²⁵⁶	2 ⁶⁴⁰

The pre-image for an n-bit hash code is determined by formula 2n, where n is the number of bits generated, according to [36]. Strong collision is also measured by the formula 2n/2, where n is the number of bits generated by a hash algorithm. This implies that the longer the hash value, the more operations and permutations the attacker would have to run in order to carry out such attacks.

CONCLUSIONS

Based on the results of the study, the following conclusions are drawn, 1. The modified SHA1 algorithm was evaluated to be 100% resistant to the collision attack. 2. The modified SHA1 algorithm exhibits a very high level of resilience against brute-force attack based on the

approximately 1.6907x10⁴⁷¹years it would take to crack generated hash values. 3 The modified SHA1 algorithm exhibits 100% resilience against a dictionary attack and 100% resilience against a rainbow table attack. The conclusions above show that modification of SHA1 through the expansion of its hash value from 160 bits to 1280 bits is effective in addressing the collision problem of SHA1. Furthermore, the modification was able to generate hash values that were resilient against the common forms of security attacks. Having produced a more secured and modified SHA1, it can be said that this study has successfully achieved its objectives.



ACKNOWLEDGEMENT

The author acknowledges the support of the Research and Development office of Surigao del Sur State University.

REFERENCES

- [1] G. Raj, R. Kesireddi and S. Gupta. 2015. Enhancement of Security Mechanism for Confidential Data using AES-128, 192 and 256bit Encryption in Cloud, 1st International Conference on Next Generation Computing Technologies (NGCT-2015), 374-378.
- [2] A. Arora, A. Rastogi, A. Khanna and A. Agarwal. 2017. Cloud Security Ecosystem for Data Security and Privacy, 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence. 288-292.
- [3] M. Ramesha, S.B. Sridhara, A.B. Gowda, N. Anughna and G. Bharathi. 2020. Design and Development of Low Power BTED Cryptography Algorithm on FPGA, International Journal of Advanced Trends in Computer Science and Engineering. 9(4): 4359-4362.
- [4] D. Morres. 2015. SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions, Federal Information Processing Standards Publication.
- [5] L. Zhong, W. Wan and D. Kong. 2016. Java web login authentication based on improved md5 algorithm. 131-135.
- [6] X. Zheng and J. Jin. 2012. Research for the Application and Safety of MD5 Algorithm in Password Authentication, 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2012). 2216-2219.
- [7] P. Walia and V. Thapar. 2015. Implementation of New Modified MD5-512 bit Algorithm for Cryptography, International Journal of Innovative Research in Advanced Engineering (IJIRAE). 1(6): 87-97.
- [8] NIST. FIPS. 1995. Secure Hash Standard, FIPS PUB. 180-1, 17.
- [9] Q. Dang. 2015. Secure Hash Standard, Gaithersburg. 1-36.
- [10] N. Kishore and B. Kapoor. 2016. Attacks on and advances in secure hash algorithms, IAENG Int. J. Comput. Sci. 43(3): 326-335.
- [11] S. Rao. 2015. Advanced SHA-1 Algorithm Ensuring Stronger Data Integrity, Int. J. Comput. Appl. 130(8): 25-27.
- [12] R. Karthik and A.K. Parvathy. 2017. Non-convex Economic Load Dispatch using Cuckoo Search Algorithm, IndonesJ Electr. Eng. ComSci. 5(1): 48-57.
- [13] K. Saravanan and A. Senthilkumar. 2013. Theoretical Survey on Secure Hash Functions and issues, Int. J. Eng. Res. Technol. 2(10): 1150-1153.
- [14] P. Garg and N. Tiwari. 2012. Performance Analysis of SHA Algorithms (SHA-1 and SHA-192): A Review 2(3): 130-132.
- [15] S. Song. 2012. A Study on Area-Efficient Design of Unified MD5 and HAS-160 Hash Algorithms, The Journal of the Korean Institute of Information and Communication Engineering. 5(16): 1015-1022.
- [16] M. Chouhan and V. Kapoor. 2015. A new security system using ECC AND MD5, International Journal of Engineering Research & Technology. 1798-1801.
- [17] X. Wang, Y. Yin and H. Yu. 2005. Finding Collisions in the Full SHA-1, Adv. Cryptol. - CRYPTO 2005. 17-36.
- [18] A. Kumarkasgar, J. Agrawal and S. Shahu. 2012. New modified 256-bit MD5 Algorithm with SHA Compression Function, Int. J. Comput. Appl. 42(12): 47-51.
- [19] P. Karpman, T. Peyrin and M. Stevens. 2015. Practical Free-Start Collision Attacks on 76-step SHA-1, 2: 1-22.
- [20] G. Gupta and S. Sharma. 2013. Enhanced SHA-192 algorithm with a larger bit difference, Proc. - 2013 Int. Conf. Commun. Syst. Netw. Technol. CSNT. 152-156.
- [21] X. Xu, Q. Zhao and C. Li. 2012. Advanced framework for iterative hash functions, Proc. - 2012 Int. Conf. Comput. Sci. Electron. Eng. ICCSEE. 2: 599-602.
- [22] M. Stevens, P. Karpman and T. Peyrin. 2016. Freestart Collision for Full SHA-1, Annual International Conference on the Theory and



Applications of Cryptographic Techniques. 9665: 459-483.

- [23] A. Bhandari. 2017. Enhancement of MD5 Algorithm for Secured Web Development, *Journal of Software*, 12(4): 240-252.
- [24] Y. Sasaki. 2015. Improved Single-Key Distinguisher on HMAC-MD5 and Key Recovery Attacks on Sandwich-MAC-MD5 and MD5-MAC, *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*.1: 26-38.
- [25] A. Kasgar, J. Agrawal and S. Sahu. 2012. New Modified 256-bit MD5 Algorithm with SHA Compression Function, *International Journal of Computer Applications*. 42(12): 47-51.
- [26] Asbduvaliyev, S. Lee and Y. Lee. 2016. Modified SHA1 hash function (mSHA1).
- [27] R. Rivest. 1991. The MD4 message-digest algorithm. *CRYPTO'90, LNCS*. 303-311.
- [28] H. Tiware and K. Asawa. 2013. Enhancing the Security Level of SHA-1 by Replacing the MD Paradigm, *Journal of Computing and Information Technology -CIT* 21, 4: 223-233.
- [29] M. Alam and S. Ray. 2013. Design of an Intelligent SHA-1 Based Cryptographic System: A CPSO Based Approach, *International Journal of Network Security*.15(6): 465-470.
- [30] San Jose, B. Gerardo and B. Tanguilig. 2015. Enhanced SHA-1 on Parsing Method and Message Digest Formula, *Proceedings of the Second International Conference on Electrical, Electronics, Computer Engineering and their Applications (EECEA2015)*, Manila, Philippines. 1-9.
- [31] R. Siddhartha. 2015. Advanced SHA-1 Algorithm Ensuring Stronger Data Integrity, *International Journal of Computer Applications*, 8(130).
- [32] Dobre. 2013. Hash Function XOR. *Information Security*.
- [33] Vandierndonck and K. De Bosschere. 2005. XOR-based Hash Function, *IEEE Transactions on Computers*, 54(7): 800-812.
- [34] K. Kiran and R. Shantharama. 2020. FPGA Implementation of Simple Encryption Scheme for Resource-Constrained Devices, *International Journal of Advanced Trends in Computer Science and Engineering*. 9(4): 5631-5639.
- [35] S. Nagesh, K. Anil Kumar, S. Suchitra and S. Abhishek. 2020. Data Security in Cloud Environment Based on Comparative Performance Evaluation of Cryptographic Algorithms. *International Journal of Advanced Trends in Computer Science and Engineering*. 9(4): 4989-4997.
- [36] Gordon. 2016. Properties of Hash Functions, Sirindhorn International Institute of Technology, Thailand: Thammasat University.