Al-Mustansiriyah Journal of Science
ISSN: 1814-635X (print), ISSN:2521-3520 (online) Volume 34, Issue 3, 2023 DOI: http://doi.org/10.23851/mjs.v34i3.1278

Research Article **Open Access**

Issuing Digital Signatures for Integrity and Authentication of Digital Documents

Hassan Kassim Albahadily[1*], Ismael Abdulsatar Jabbar[1], Alaa Abdulhussaien Altaay[1], Xunhuan Ren[2]

[1]Department of Computer Science, College of Science, Mustansiriyah University, 10052 Baghdad, IRAQ.
[2]Belarusian State University of Informatics and Electronics, Minsk, Belarus.

*Correspondent contact: hassan@uomustansiriyah.edu.iq

Article Info

Received
16/01/2023

Revised
24/04/2023

Accepted
29/04/2023

Published
30/09/2023

ABSTRACT

In this paper, we developed a secure system for issuing digital signature for digital documents and certificates using message digest algorithm (MD5). The developed system is providing the integrity and authentication for certificates by combining the information of participant and supply them to MD5 algorithm to produce a unique hash key of 32 digits which is hard to guess or attack. The process provides a certificate which allows for the authentication of a document at any time. The proposed system used and tested to produce digital certificate for participant of electronic seminars in Mustansiriyah university which were about 1000 certificate. The results were good and very fair to authenticate certificates and preventing forgery.

KEYWORDS: MD5, digital signature, hash function, security.

الخلاصة

في هذه الورقة البحثية طورنا نظام امني لاصدار التواقيع الرقمية للوثائق والشهادات الرقمية باستعمال الخوارزمية الهاضمة MD5 . النظام المطور يوفر السلامة والموثوقية للوثائق الرقمية بواسطة جمع معلومات المشاركين وتغذيتها للخوارزمية الهاضمة MD5 التي بدورها تنتج شفرة هاش ذات طول ٣٢ رقم والتي تكون صعبة التوقع والتخمين من قبل المهاجم. المعالجة المقترحة في النظام توفر الموثوقية للشهادات الرقمية في اي وقت. النظام المقترح تم استخدمه في اختبار الشهادات الرقمية الصادرة من الجامعة المستنصرية والمقدرة ١٠٠٠ شهادة. كانت النتائج جيدة ونجحت الطريقة امام الاختبارات القياسية وقدمت مستوى عالي من الموثوقية والحماية من التزوير.

INTRODUCTION

The digital life is growing every day causing increase in the amounts of electronic documents and bringing new technologies and methods for life steeping us forward to the E-life in the near future. E-learning and distance learning and the huge growth of E-commerce and E-government requiring the treatment of digital document which handling sensitive information and must be protected from being tampered by malicious. One important type of digital document is the certificate for participants in the seminars and workshops which is granted to people who attending the electronic seminars. These certificates need to be checked and validated.
The main hash function is shown in Figure 1.

One of the most power of such functions is taken

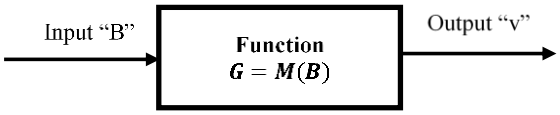

Figure 1: Hash Function

variable length of "B" as input as well as given a fixed "V" as fixed length of hash code output [1] as shown with Eq. (1)

$$G = M(B) \quad (1)$$

The main purposes of a hash function are scramble, accept an input of arbitrary length and output a fixed length result, manipulate data irreversibly. Because of this property of getting flexible value of length and giving a limited

value of length there are a common use of hash function like widely demanded in the cryptography as well as steganography because of most common features: one-way function and without collusions [2][3][4].

Hash functions utilized with several points of security fields to achieve integrity purpose. Hash function can be used in many fields [5][6], like used alone, file integrity verification, public key fingerprint, password storage, combined with encryption functions, information hiding and digital signature.

MD5 (Message Digest Method 5) is a cryptographic hashing algorithm which was created for authenticating messages and verify the integrity of any message as well as content verification and digital signatures. It is converted standard data into an unrecognizable format by generating a 128-bit digest from a string of any length.

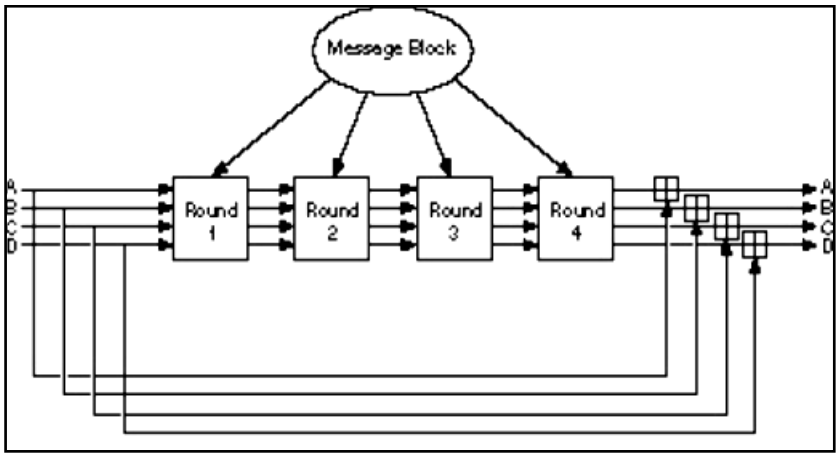

Figure 2: MD5 main loop

The hash algorithm MD5 has been chosen to generate the digital hash code because of resistance to collision aattacks, making it not possible in a situation for an attacker to get the same hash value and verifying changes could be happened to a document result in the generation of unique hash codes, thus exposing with any step to change potentially relevant evidence. Message digest used in the proposed system with the following jobs: Joined with the encryption process, Achieving the integrity and authentication process.

The "MD5" algorithm used as enhanced version of "MD4" algorithm which ginning more complexity compared with the "MD4" algorithm, both of the designs are similar and gives a 128-bit as output [7][8]. MD5 main loop algorithm can be shown with in Figure 2.

For constructing the work start with expanding code process by making code of 64-bit with multiple of 512 bit. The expanding process through adding one-bit at the end of messages while set of zeros added also as needed. The four variables of 32 bits work for the initialization such variables called "chaining variables" [9].

A = 0x01234567
B = 0x89abcdef
C = 0xfedcba98
D = 0x76543210

Four of rounds forming the main loop of the Md5 algorithm which runs 512-bit length of the message as required. Within each round coping the four variables values into other variables. The nonlinear operation achieved because at each round different operation work in 16 times on the copied variables. At each operation the functions can be applied as follow [10][11]:

$$F(X, Y, Z) = (X \wedge Y) \vee (\neg X \wedge Z) \quad (2)$$

$$G(X, Y, Z) = (X \wedge Z) \vee (Y \wedge \neg Z) \quad (3)$$

$$H(X, Y, Z) = X \oplus Y \oplus Z) \quad (4)$$

$$I(X, Y, Z) = Y \oplus (X \vee \neg Z) \quad (5)$$

where $\oplus,\wedge,\vee,\neg$ denote the "XOR", "AND", "OR" and "NOT" operations respectively.

These process well designed in order to condition the equivalent bits of "X, Y, and Z "are autonomous and impartial, such that, the bit of the output will even have considered autonomous and impartial.

The function F is the bitwise restricted: "If X Then Y Else Z". The function H is the bit-wise

parity operator. One MD5 operation can illustrated using Figure 3. [12].

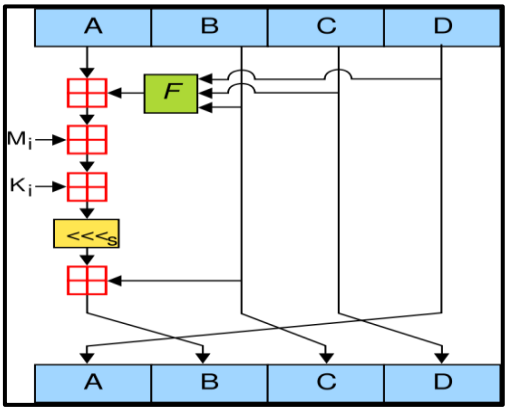

Figure 3: One MD5 operation

MATERIALS AND METHODS

Generating Digital certificate database need to have a list of available works that's is already proposed by presenters and registers list of people who they are interest in specific work. Let $x_1 \dots x_n$ set of the registers people, while $y_1 \dots y_k$ set of the available work, such that $\forall\, x_i$ interest in the work $\forall\, y_j$ then form a list of participants P, such that $P = \{p_1 \dots p_m\}$ where $m \leq n$. For all elements in set P collect all the information and issuing a digital certificates database. The flowcharts shown in the Figure 4. When the digital certificate produce need to be assigned for the organization database certificate as well as assigning a signature of 128 bit using MD5 algorithm which is owe way function such code will be based on the several information collected from the participants. The proposed field to feed MD5 algorithm will be the following:

1. Name of the participants.
2. Title of the work.
3. Data of the seminar or workshop.
4. Digital documentation number (Issuing by the organization).
5. Name of the organization that issuing the certificates.
6. Mobile phone number.

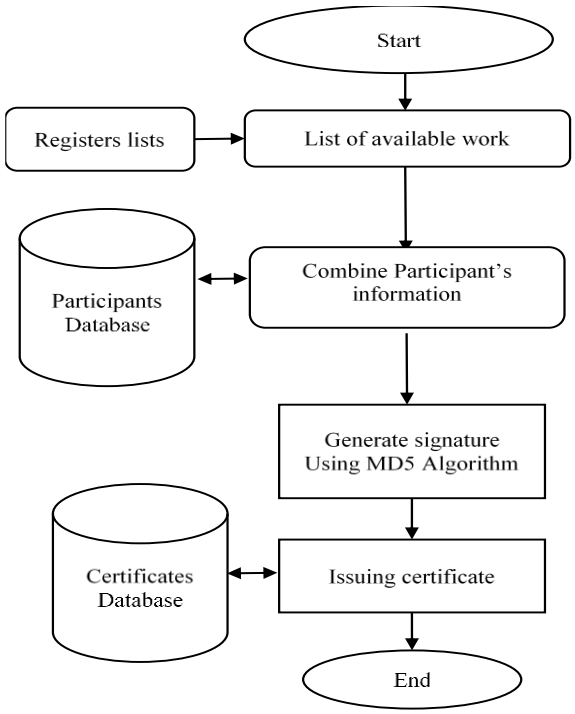

Figure 4: Issuing Digital certificate proposed algorithm

The name of the participant could be duplicate with other names as well as the title of the work with low possibility with date that's is why using digital document number (DDN) as well as mobile number there are unique numbers for every person to avoid generating the same hash code this process can show in the Figure 5.

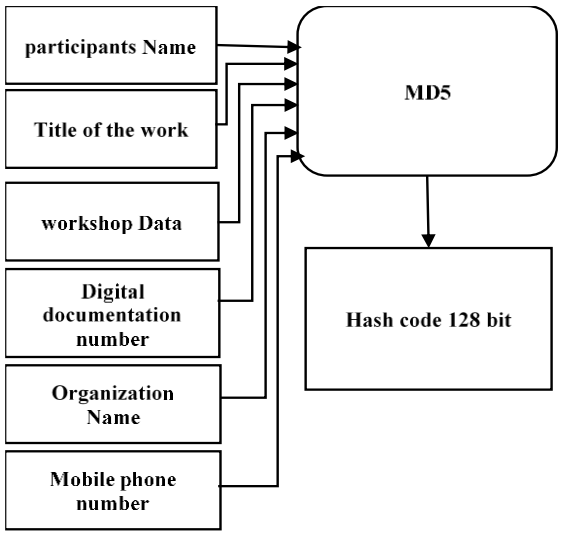

Figure 5: Proposed issuing hash code using MD5 algorithm

The visualization for the hash 128-bit MD5 code will be in the hexa-decimal number system each 4 bits will be form one digit to get at the final 32 digits for 128 bits. When the certificate received for validity, the first step is testing the signature generated database by reaching the participant as primary key to get the associated signature. If the certificate original, then confirm originality otherwise alert unoriginal certificate as shown in Figure 6.

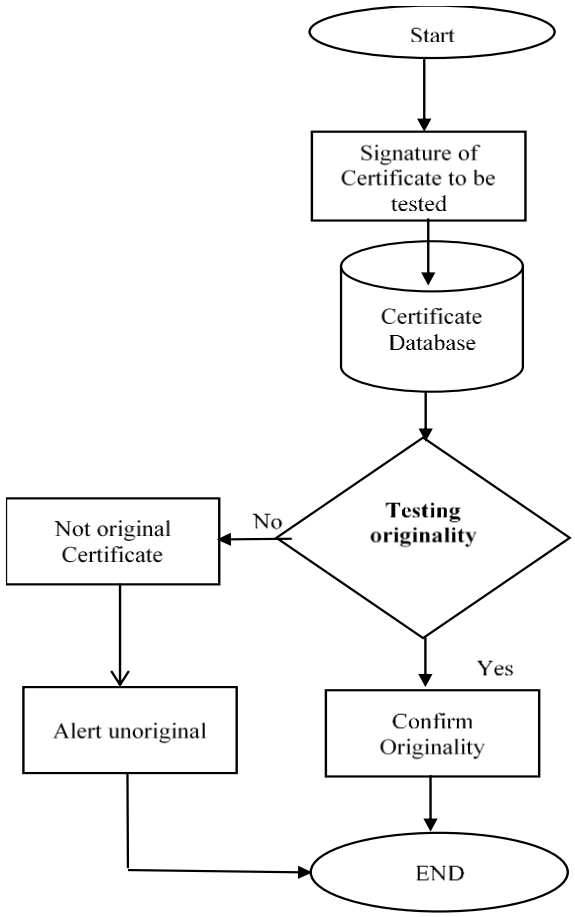

Figure 6: Digital certificate originality testing

The hash code generated using the MD5 algorithm will be sensitive to the input value, and any changes that occur will affect the output value of the generated hash code [1][13].

RESULTS AND DISCUSSION

The proposed system used to produce the digital hash code for the certificates of the electronic workshops. The input for the proposed system will contains set of fields combining the basic information of the workshop as shown in Table 1.

Table 1: The output hash for test information

Variable	Values
Name	Ismael Abdul Sattar Jabbar
Title	Deep learning and recommender system
Workshop Date	31/12/2021
Digital Document Number	217081
Name of the organization	University of Mustansiriyah
Phone number	07716830565

The output for the digital signature will be "c292d9dfaaad1663bd50cf260691fac0".
The proposed system generated several sets of digital certificates signed with proposed MD5 hash signature to form a database of generated signed certificates. After multiple runs recorded set of 10 signatures for the produced certificate which can shows in the Table 2.

Table 2: Digital certificate signature samples

Variable	Hash Function
1	bb62d2893a6ef0be73d7838173aa77f4
2	6bb83e9af6fa8fc1c55d5c01a1abefb2
3	31daf1ec6df9b55ab49163e92a899511
4	1a5b3a6e2d855d7fe91bc432922ac959
5	8a868059ce7e89b3ef1b765db45d7258
6	86b30204fdd04143b026b01de6ffe70b
7	3bb8beea470bdeeed7a9f4ab5ac14830
8	1de46a4451a6cc0aabd1259417fe4df5
9	905636111d955ab7d87bd9e1c3956cf8
10	6fd6f5b804105a789f15c4d899f05b54

The code will be printed out on the certificate as shown in Figure 7.

Figure 7: The certificate with hash code from MD5

CONCLUSIONS

The proposed system not only generates digital certificates for participants but also creates a digital signature associated with each certificate to ensure the integrity of the digital certificate. The length of the generated hash code used to represent a digital signature will be sufficiently complex (12 bits) to cover a wide range of productions. Additionally, the generated signature will be unique due to the one-way MD5 algorithm, making this mechanism secure in case of attempts to analyze the hash code to identify the fields that have already produced such a hash code.However, there are some limitations to consider, including the vulnerability of the MD5 hash function to attacks that could result in the generation of identical codes. To enhance security, it is advisable to use more robust hash functions such as SHA. Furthermore, it would be beneficial to embed the hash code of the certificate as metadata within the certificate's image file, along with all relevant information and references to the issuing organization. Additionally, we recommend converting the hash code into a QR code for added convenience.

Disclosure and Conflicts of Interest: The authors advertise that they have no conflicts of interest.

REFERENCES

[1] A. Rawat, D. Agrawal, An Enhanced Message Digest Hash Algorithm for Information Security, 2015.
[2] R. Kundu, A. Dutta, Cryptographic Hash Functions and Attacks-A Detailed Study. International Journal of Advanced Research in Computer Science, 2020, vol. 11, No. 2.
[3] W. Easttom, Cryptographic Hashes. In Modern Cryptography, Springer, Cham, 2021, pp. 205-224.
[4] J. Ismael Abdul Sattar, S.Hameed Shaker, and Mohammed Najm Abdullah. "Database meet Link generator based on linear feedback shift register and message digest algorithm." Webology (2022): 2236-2243.
[5] M. Al-Awawdeh, Strengthening the MD5 File Integrity Algorithm with User Fingerprint (Doctoral dissertation, Middle East University), 2019.
[6] Sattar Jabbar, Ismael Abdul, Hassan Kassim Albahadilyr, and Alaa A. Jabbar Altaay. "Design and Implementation Digital Invitation System Based on Secure Hash Algorithm 3." International Journal of Online & Biomedical Engineering 19.5 (2023).
[7] N. Kishore, P. Raina, Parallel cryptographic hashing: Developments in the last 25 years. Cryptologia, vol. 43, No. 6, 2019, pp. 504-535.
[8] Sattar, Ismael Abdul, and Jamal Nasir Hasoon. "Hiding System Based on Double MD5 Hashes and LFSR Generators." (2013).

[9] S. Long, A Comparative Analysis of the Application of Hashing Encryption Algorithms for MD5, SHA-1, and SHA-512. In Journal of Physics: Conference Series, vol. 1314, No. 1, 2019, p. 12210).

[10] M. Gillela, V. Prenosil, and V. Ginjala, Parallelization of brute-force attack on MD5 hash algorithm on FPGA. In 2019 32nd International Conference on VLSI Design and 2019 18th International Conference on Embedded Systems (VLSID), 2019, pp. 88-93.

[11] F. Sagar, Cryptographic Hashing Functions—MD5. no. September, 2016, pp.1-9.

[12] P. Gupta, S Kumar, A comparative analysis of SHA and MD5 algorithm. architecture, 1, 2014, p. 5.

[13] J. Ismael AbdulSattar, S. Hameed Shaker. "Adaptive Hiding Algorithm Based on Mapping Database." iJIM 17.01 (2023): 97.

How to Cite

H. K. Albahadily, I. A. Jabbar, A. A. Altaay, and X. . Ren, “Issuing Digital Signatures for Integrity and Authentication of Digital Documents”, *Al-Mustansiriyah Journal of Science*, vol. 34, no. 3, pp. 50–55, Sep. 2023.