تحسين خوارزمية MD5 بالاعتماد على عدة طرق

أ.م. د. علاء فرحان
dralaa_cs@yahoo.com

الباحث محمد عبد الجابر علي
mo_force@yahoo.com

الجامعة التكنولوجية – قسم علوم الحاسبات

المستخلص

خوارزميات Hash هي الجزء المهم في العديد من تطبيقات التشفير والبروتوكولات الأمنية . هذا البحث يعمل على تعزيز خوارزمية MD5 ضد بعض الهجمات. التطوير الذي تم اعتماده في تحسين خوارزمية MD5 هو استخدام DNA coding ، والمعادلات الغير خطية (NLFSR)، Logistic function . بالاضافة إلى توسيع مدخلات خوارزمية MD5 إلى 1024 بت بدلا من 512 بت، والمخرجات إلى 160 بت بدلا من 128 بت. عمل الخوارزمية التي تم تحسينها تعمل على تعقيد البيانات المدخلة قبل الدخول في MD5 rounds باستخدام التقنيات السابقة ذكرها . إن كفاءة طريقة تعديل ثبتت من خلال العديد من الأمثلة المقارنة.

الكلمات المفتاحية: **MD5 ,DNA CODING ,NLFSR, LOGISTIC FUNCTION**

Enhancement MD5 depend on multi techniques

Assist .prof .Dr. Alaa K. Farhan **Mohammed Abed Al-jabber Ali**
University of Technology - Computer Science Department

Abstract

Hash algorithms are critical part in many cryptography applications and security protocol suites. This research works on enhancement MD5 algorithm. The enhancement method depended on multi techniques such as DNA coding, non-Linear Feedback Shift Register (NLFSR), and Logistic function of Chaos theory. It will expand of input MD5 algorithm to 1024 bits instead of 512 bits, and change output to 160 bits instead of 128 bit. The complexity try's increase of data input by preprocessing before entrance MD5 operation, where, will use NLFSR and DNA coding in addition the use Logistic function of Chaos theory with each rounds of MD5. The efficiency of the modified method had proved by several comparative examples.

Keyword :MD5 ,DNA CODING ,NLFSR, LOGISTIC FUNCTION

1. Introduction:

Message Digest Algorithm, Encryption Hash Function or simply the Hash Function, takes a series of variable length as inputs in returns it generates a series of fixed length as output. Commonly known as hash value or message digest. The most interested key function of hash function is the process of digitally signing documents (digital signature), while hash function is used for a lot of other purposes, the Digital signature hash function gathers lot of developers to develop yet enhance the outcome of such a function. [1]

MD5 algorithm is one of the hash algorithms used in many important applications, therefore, make many researchers and hackers looking for security lapses to weaken this algorithm.

MD5 is one of the family MD, MD2, MD3 and MD4 after development. Entrance to it any data and outputs a fixed length of information to digest of 128 bits along while this part of the message digest often claims that digital data footprint. MD5 employs a series of NLR algorithm to do circular process, to prevent hackers from restoring the original data. Since this algorithm is irreversible, this feature prevents the resulting inverse process data leakage. Security has been proven to be good in term of theory and practice, speed of execution, ease of implementation and the integrity are some of the required key features of MD5 algorithm, which makes use of large-scale applications in the non-confidential to the public. MD5 is also an excellent intermediate process in any security technology. [2]

2. DNA coding

“DNA computing” may be An manifestation from claiming registering that replaces universal workstation innovation In view of silicon by DNA, natural chemistry What's more atomic science.

DNA computing, or all the more for the most part, organic and sub-atomic processing, is a multi-disciplinary quickly advancing field. With the quick progression of DNA computing. Contrasted with Binary processing 0's and 1's, a solitary DNA strand arrangement is a synthesis of four bases, they are (A, C, G) and (T), where (A) and (T) are supplement to each other, so are (C) and (G).

In the present day hypothesis of electronic PC, all data is communicated by 0's and 1's (paired framework). In any case, in DNA coding hypothesis data is spoken to by DNA successions.

To make this conceivable in electronic PC we utilize double numbers to express the four bases in DNA arrangement and two-bit parallel number to speak to a base.

In the binary system theory of "0" and "1", interconnected, it can generate "00" , "11","01", and "10" are also interconnected. Addition and subtraction of the DNA sequence since the development of DNA computing, researchers have proposed utilizing mathematical operation of DNA grouping to supplant the conventional PC logarithmic operation. Based on this, we use the DNA process, in addition to achieving the computing matrix DNA sequence of the message and the key to a polynomial. The algorithm of this paper discovers the rules of adding DNA using the Department of Defense two operations binary number at (01 – A), (10 – T), (00 – C), (11 – G), and you can find rules in Figure 1 and Figure 2. [3]

1	2	3	4	5	6	7	8
00 – A	00– A	00 – C	00 – C	00 – G	00 – G	00 – T	00 – T
01 – C	01 – G	01 – A	01 – T	01 – A	01 – T	01 – C	01 – G
10 – G	10 – C	10 – T	10 – A	10 – T	10 – A	10 – G	10 – C
11 – T	11 – T	11 – G	11 – G	11 – C	11 – C	11 – A	11 – A

Fig 1 - DNA coding. [3]

+	A	T	C	G
A	T	G	A	C
T	G	C	T	A
C	A	T	C	G
G	C	A	G	T

Fig 2 - DNA addition operation. [3]

3. Logistic map

Chaotic systems need complicated formula, the degree of chaos is linked with the degree of complication, but sometimes a very simple functions can lead not only to chaos system, but how this chaos generated from an ordered behavior.

The logistic function, utilized as a part of population dynamics, it is the following equation 1: [4]

$$\boldsymbol{Xn + 1 = \mu Xn * (1 - Xn)} \dots\dots\dots\dots\dots\dots\dots\dots \boldsymbol{(1)}$$

Where ($0 \leq \mu \geq 4$) and X0 = 0.1 as initial condition.

Utilizing the arbitrary numbers encoded the information and messages to create a mystery key for encryption. [4]

4. Stream Ciphers

One-Time Pad (OTP) is a symmetric cryptosystem invented in 1882 is considered the only proven unbreakable cryptosystem, tests were made by Claude Shannon [Sha49] in 1949 using information theoretical arguments.

Given a plaintext message (p) and a secret random key (z) of the same length, the cipher text c is given by equation 2:

$$\boldsymbol{c = p \oplus z} \dots\dots\dots\dots\dots\dots\dots\dots \boldsymbol{(2)}$$

Where $\oplus$ denotes bitwise XOR. While plaintext, cipher text and key entities are string of bits (0's,1's).

Stream ciphers uses OTP main key feature (Secret random key as long as the plain text). A stream cipher generates a long arbitrary key stream using n-bit

secret key as input. The generated key is then use OTP to mask and unmask the message.

Stream ciphers have a fixed key length, ordinarily in the scope of 80–256 bits. In any case, since using the same key twice will be the same key stream output, one must be cautious with the use. So it is common for the cipher stream to take a second parameter, an initialization value (IV). The secret key can then be reused if one chooses a new confinement randomly selected for each new encryption task, such confinement does not need to be secret. [5]

4.1 Linear Feedback Shift Registers (LFSR)

A LFSR comprises of clocked stockpiling components (flip-flops). Also an input way. The amount from claiming capacity components provides for us those degrees of the LFSR. Clinched alongside different words, a LFSR with m flip-flops is said on make of degree m. Those input system computes the enter to the most recent flip-flop as XOR-sum of specific flip-flops in the movement register. Since the XOR may be a straight operation, such circuits are known as straight sentiment movement registers. [6]

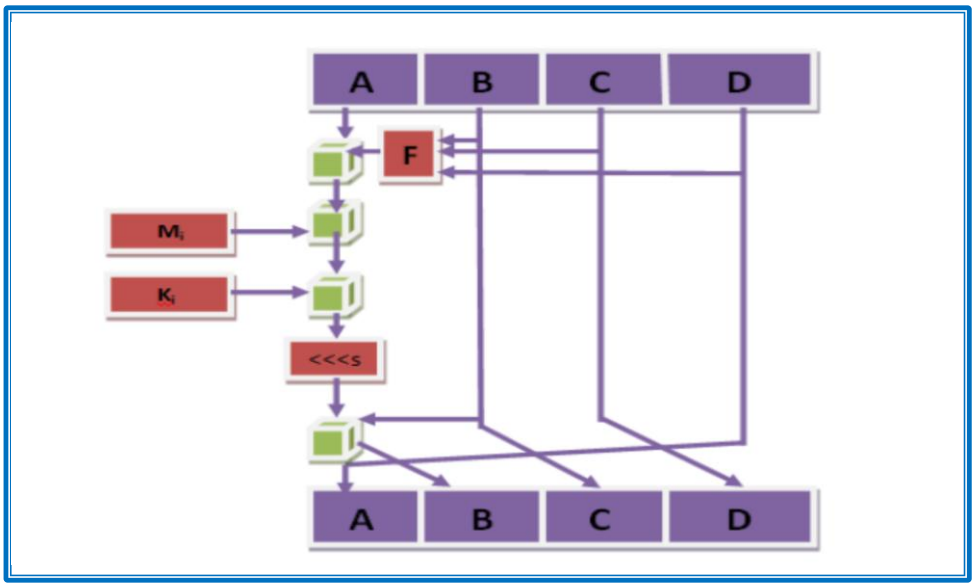
4. 2 Non-Linear Feedback Shift Register (NLFSR)

Non-Linear Feedback Shift Register is a as a relatable point part for advanced stream ciphers, particularly in RFID Also smartcard provisions. NLFSRs need aid referred to make all the more safe to cryptanalytic strike over Linear Feedback Shift Registers (LFSRs). It is known how to generate a *n*-bit NLFSR of maximal length 2^n, by extending a maximal-length LFSR with *n* stages, but the construction of Different vast NLFSRs with guaranteed in length periods stays an open issue. Using brute force methods, a rundown from claiming maximum-period n-bit NLFSRs to n < 25 need been made and also to n=25 Furthermore n=27. [7].

5. MD5 algorithm

The MD5 is a broadly utilized cryptographic hash capacity creating a 128-piece (16-byte) hash esteem, the yield is communicated as a 32-digit hexadecimal number. MD5 can be discovered installed in a wide assortment of cryptographic applications and normally used to check information trustworthiness.

MD5 algorithm planned by educator Ronald (Rivest, 1992). It was as successor to MD4 which was designated to be unstable, Rivest planned MD5 in 1991 as a substitution. Figure (3) Explain how work MD5 to one block [8].



MD5 will be An change work that's transforms its enter (a set for information of whatever length) of the yield (a hash esteem of 128-bit length) this conversion may be irreversible Furthermore it may be An sequential preparing technique.

1st it includes 64-bit double digits of the end for information stream, this 64-bit will be the first information extent on odds.

Then afterward that cushioning methodology will be used to settle on those bit length about information continuously transformed a different from claiming 512-bit period.

Those information are separated under obstructs for 512-bits each, computations need aid performed for each square done a requested way.

Those enter of the initial square operation is An 128-bit recognized Concerning illustration introductory value, same time those yield of this square operation will be those information of the following square operation.

The yield of the most recent piece operation may be the "MD5 hash value".

Those preparing workflow may be illustrated in the accompanying algorithm: -.

Algorithm (1): MD5 algorithm

Input: message of discretionary length

Output: 128 bit hash value

Begin

Step1 : M= (S0, S1,……., Sn-1), Message to hash , after padding// Each Si is a (32-bit) word and (N * 16)

MD5 (M)

Step2 : (A,B,C,D) = (0x67452301,0xefab89 , 0x98badcfe ,Ox10325476)// initialize (A,B,C,D) = IV

Step3 : for k=0 to N/16 -1

for l = 0 to 15 // Copy block S to R
Rl = S16k+l

for l = 0 to 63// Copy R to M
Mj = Rσ(j)

(Z-4 , Z-3 , Z-2 , Z-1) = (A , D , C , B) // initialize Z

For j= 0 to 63 // Rounds zero , one , two and three
If 0 < j < 15 // Round zero: Steps 0 through 15, uses are (F) function

Zi = Zi-1 + ((Zi-4 + F(Zi-1 , Zz-2 , Zz-3) + Wh+Kj) <<<sj)

If 16 < j < 31 // Round one: Steps 16 through 31, uses are (G) function

Zi= Zi-1 + ((Zi-4 +G(Zi-1 , Zi-2 , Zi-3) + Wj+Kj) <<<sj)

If 32< j < 47 // Round two: Steps 32 through 47, uses are (H) function

Zi= Zi-1 + ((Zi-4 +H(Zi-1 , Zi-2 , Zi-3) + Wj+Kj) <<<sj)

If 48 < j < 63 // Round three: Steps 48 through 63, uses are (I) function

Zi= Zi-1 + ((Zi-4 +I(Zi-1 , Zi-2 , Zi-3) + Wj+Kj) <<<sj)

// Each addition is modulo 2^{32}
(A , B , C , D)=(Z60 + Z-4 , Z63 + Z-1 , Z62 + Z-1 , Z61 + Z-3)
next i
return A , B , C , D
End MD5
end program

6. Enhancement MD5 algorithm

The enhancement, which was on the MD5 includes increased complexity against brute force attacks and increase the percentage of probability to know the explicit provision against attackers on the previous algorithm, in figure (4) we explain how work algorithm to one block, can explain the enhance in MD5 by following algorithm: -

Algorithm (2) enhancement MD5 algorithm

Input: message of discretionary length

Output:160 bit hash value

Begin

Step1 : M= (S0, S1,……., Sn-1), Message to hash , after padding// Each Si is a (32- bit) word and (N * 16) MD5 (M)

Step2 : (A,B,C,D,E) = (0x67452301,0xefab89 , 0x98badcfe ,Ox10325476,Ox10D25F7A)//initialize (A,B,C,D,E) = IV

Chos = 3.5

Mu= 0.99

Step3 : for k=0 to N/16 -1

for l = 0 to 15 // Copy block S to R

Rl = S16k+l

for l = 0 to 63// Copy R to M

Mj = Rσ(j)

Generated Polynomial key By (NLFSR) with used polynomial function ($x^{32} + x^{28} + x^{27} + x + 1$) Poly_key=(Poly0,poly1,…, poly(N/16 *64))

(Z-4 , Z-3 , Z-2 , Z-1) = (A , D , C , B) // initialize Z

For j= 0 to 63 // Rounds 0 , 1 , 2 and 3

(W,Poly_Key)=DNA decoding (W,Poly_Key) by fig (1)

W = Addition operation (W, Poly_Key) by fig (2)

If 0 < j < 15 // Round zero: Steps 0 through 15, uses are (F) function

Zi = Zi-1 + ((Zi-4 + F(Zi-1 , Zz-2 , Zz-3) + Wh+Kj) <<<sj)

Chosj= ((Chos [j - 1] * Mu) * (1 - Chos[j - 1]))

Zi = (Z-5 * Chos[j]) * Zi ;

End if

(W,Poly_Key)=DNA decoding (W,Poly_Key) by fig (1)

W = Addition operation (W, Poly_Key) by fig (2)

If 16 < j < 31 // Round one: Steps 16 through 31, uses are (G) function

Zi= Zi-1 + ((Zi-4 +G(Zi-1 , Zi-2 , Zi-3) + Wj+Kj) <<<sj)

Chosj= ((Chos [j - 1] * Mu) * (1 - Chos[j - 1]))

Zi = (Z-5 * Chos[j]) * Zi ;

End if

(W,Poly_Key)=DNA decoding (W,Poly_Key) by fig (1)

W = Addition operation (W, Poly_Key) by fig (2)

If 32< j < 47 // Round two: Steps 32 through 47, uses are (H) function

Zi= Zi-1 + ((Zi-4 +H(Zi-1 , Zi-2 , Zi-3) + Wj+Kj) <<<sj)

Chosj= ((Chos [j - 1] * Mu) * (1 - Chos[j - 1]))

Zi = (Z-5 * Chos[j]) * Zi ;

End if

(W,Poly_Key)=DNA decoding (W,Poly_Key) by fig (1)

W = Addition operation (W, Poly_Key) by fig (2)

If 48 < j < 63 // Round three: Steps 48 through 63, uses are (I) function

$$Zi= Zi\text{-}1 + ((Zi\text{-}4 +I(Zi\text{-}1 , Zi\text{-}2 , Zi\text{-}3) + Wj+Kj) <<<sj)$$

Chos i= ((Chos [i - 1] * Mu) * (1 - Chos[i - 1]))

Zi = (Z-5 * Chos[i]) * Zi ;

End if

// Each addition is modulo 2^{32}

(A , B , C , D, E)=(Z60 + Z-5 , Z63 + Z-1 , Z62 + Z-1 , Z61 + Z-3,Z60 +Z-2)

next i

return A , B , C , D, E

end MD5

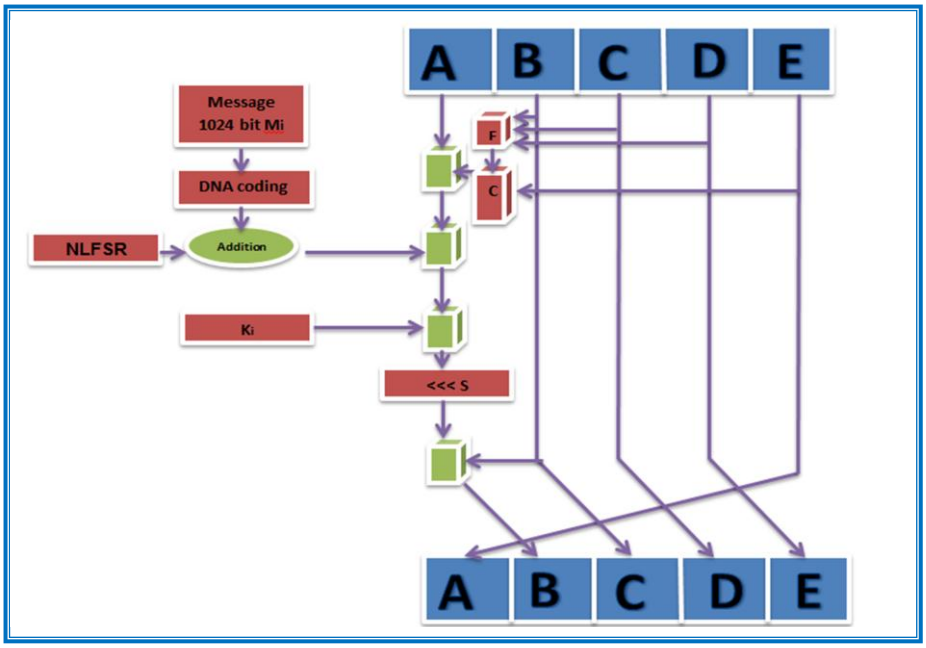


Fig 4: Enhancement MD5 algorithm

7. COMPARATIVE ANALYSIS

In this section, we have compared md5 algorithms with enhancement md5.

7.1 MD5 algorithm

- Input is a message of variable length.
- Output is a 128 bits hashed data.
- Divide message into 512 bits of block.
- Divide each block into 32 bits word multiple to 16.
- Padding "0" to 448 of block.
- Four register initial value (A, B, C, and D).
- Four rounds each round 16 steps.
- Attacks reported some extend and web broken MD5 [9,10,11,12,13]
- No. Of Attacks Needed To Find Original Message 2^{64} bit operations Required
- Speed Faster, 60 iterations
- Low complexity

7.2 Enhancement MD5 algorithms

- Input message of arbitrary length.
- Output 160 bits.

- Divide message into 1024 bits of block.
- Divide each block into 64 bits word multiple to 16.
- Padding "0" to 896 of block.
- $x^{32} + x^{28} + x^{27} + x + 1$Generated polynomial key by nonlinear feedback shift register use
- DNA encoding both key and each block message into.
- Addition oration between key and block message
- Five register initial value (A, B, C, D and E).
- Four round each round 17 steps.
- Use logistic function each round
- No Attacks reported broken MD5 in some web online. [10][11][12][13][14]
- No. Of Attacks Needed To Find Original Message2^{128} bit operations Required
- Slower from MD5 because perprossing to block message in each round
- More complexity from MD5, where DNA coding used with addition operation to block message permeation before entered MD5 operations (four rounds), and logistic function is use complex increased because sensitivity to the initial conditions, and used five register given longer output from MD5.

7.3 Execution comparative

The following table explains the difference between output from MD5 and MD5 enhancement output, where, the difference is clear at every input, any change in every input even the simple will the difference in output with a note of complexity quotient. The use of different inputs result was different. :-

Table 1: MD5 and MD5 enhancement execution [8]

Test String	MD5	MD5 enhancement
“”	D41D8CD98F00B204E 9800998ECF8427E	28218E31024708B8B29 EB890E30503F712 C24AAD
abcdefghijkl mnopqrstuvw xyz	9FC9D606912030DCA 86582ED62595CF7	95244E262D0943F57E7 BA6F6E793EA68F0 691B8A
123456789	25F9E794323B45388 5F5181F1B624D0B	D9CCBAF98649E673FAE C5BB932CB50B5F3 3A44BC
aimit	C884202C3C2CCF128 CD315AC632593FB	C137C29C1DA9715CD79 B5D51E62A0DBD04 ED0135
AIMIT	0E616B0B6FC9242E8 EA9824C470C1AA4	85C0257C291AD7BADD D497FF574556B7F8 29CE01
123456789aim it	2B8651BA54951B977 A46ACE9BC4702AD	D8CE4F9DA46B3EA21E 15BDD5C6995F5643 7C2D9A

8. CONCLUSION

The modified MD5 algorithm can predict explicit enumeration through brute force Attack, by using reference [10][11][12][13][14] use the complexity of MD5 proposed makes it difficult to know the explicit text. DNA coding has been use in cryptography this adds complexity to MD5 proposed, you have used the chaos function is to develop mathematical models nonlinear and attracted many mathematical because of the high sensitivity of primary value, this is characterize use in complexity the enhancement MD5. Nonlinear feedback shift register (NLFSR) use to generate of key that is works to increase the complexity of the text before input to prossing of MD5 round. The purpose of the expansion algorithm to 1024 bit is to make more speed and avoid the development slow because of who conducted the algorithm.

REFERENCES

[1] Robshaw," **On Recent Results for MD2, MD4 and MD5**", Bulletin, NO.4, N O V E M B E R12, 1 9 9 6.

[2] Vivek T, Priyanka W," Implementation of New Modified MD5-512 bit Algorithm for Cryptography" , **IJIRAE, Volume 1 Issue 6 (July 2014).**

[3] Lili L, Qiang Z, Xiaopeng W, "**A RGB image encryption algorithm based on DNA encoding and chaos map**", Computers and Electrical Engineering (2012) .

[4] Stokes, Distinguished, "**Cryptography using Chaos** ", Lectures co-financed by the European Union in scope of the European Social Fund March, 2010.

[5] Paul Stankovski," **Cryptanalysis of Selected Stream Ciphers**", Doctoral Dissertation Cryptology Lund, June 2013.

[6] C. Paar, J. Pelzl, Understanding Cryptography, Springer-Verlag Berlin Heidelberg, 2010

[7] Retrieval from web site https://en.wikipedia.org/wiki/Nonlinear_feedback_shift_register [2015].

[8] Retrieval from web site https://en.wikipedia.org/wiki/MD5 [2016] .

[9] C.G Thomas, Robin Thomas Jose," **A Comparative Study on Different Hashing Algorithms**", International Journal of Innovative Research in Computer and Communication Engineering,Vol. 3, Special Issue 7, October 2015.

Enhancement MD5 Test from website:-

[10] Retrieval from web site https://crackstation.net/.

[11] Retrieval from web site http://md5decryption.com/.

[12] Retrieval from web site http://md5.web-max.ca/.

[13] Retrieval from web site http://www.md5online.org/.

[14] Retrieval from web site https://hashkiller.co.uk/md5-decrypter.aspx.