

## KEYED TILlich - ZEMOR HASH FUNCTION

\*Joju K.T.

*Department of Mathematics, Prajyoti Niketan College, Pudukad, Kerala, India, Pin-680301*

Lilly P.L

*Department of Mathematics, St. Joseph's College, Irinjalakuda, Kerala, India, Pin- 680121**(Received on: 28-12-12; Revised & Accepted on: 22-01-13)*

## ABSTRACT

At CRYPTO, 94 Tillich and Zemor proposed a family of hash function based on computing a suitable matrix product in groups of the form  $SL_2(F_{2^n})$ . But Markus Grassl, Ivana Illich, Spyros Magliveras and Rainer Steinwadt found collision for the same between palindrome bit strings of length  $2n+2$ . We construct a keyed hash function by using the same generators of Tillich-Zemor hash function, which resists the palindrome collision.

**Keywords:** Collision, Group, Hash function, Irreducible polynomial, Preimage, Second preimage.

## 1. INTRODUCTION

## 1.1 Cryptographic Hash Functions and MACs

Hash functions ([1], [5], [11]) are functions that compress an input of arbitrary length into fixed number of output bits, the hash result. If such a function satisfies additional requirements it can be used for cryptographic applications, for example to protect the authenticity of messages sent over an insecure channel. The basic idea is that the hash result provides a unique imprint of a message, and that the protection of a short imprint is easier than the protection of message itself. Related to hash functions are message authentication codes (MACs). These are also functions that compress an input of arbitrary length into a fixed number of output bits, but the computation depends on a secondary input of fixed length, the key. Therefore MACs are also referred to as keyed hash functions. In practical applications the key on which the computation of a MAC depends is kept secret between two communicating parties. For an (unkeyed) hash function, the requirement that the hash result serves as a unique imprint of a message input implies that it should be infeasible to find colliding pairs of messages. In some applications however it may be sufficient that for any given hash result it is infeasible to find another message hashing to same result. Depending on these requirements Praneel [10] provides the following informal definitions for two different types of hash functions.

A one-way hash function is a function  $h$  that satisfies the following conditions:

1. The input  $x$  can be of arbitrary length and the result  $h(x)$  has a fixed length of  $n$  bits.
2. Given  $h$  and an input  $x$ , the computation of  $h(x)$  must be easy.
3. The function must be one-way in the sense that given a  $y$  in the image of  $h$ , it is hard to find a message  $x$  such that  $h(x) = y$  (preimage-resistance), and given  $x$  and  $h(x)$  it is hard to find a message  $x' \neq x$  such that  $h(x') = h(x)$  (second preimage-resistance).

A collision-resistant hash function is a function  $h$  that satisfies the following conditions:

1. The input  $x$  can be of arbitrary length and the result  $h(x)$  has a fixed length of  $n$  bits.
2. Given  $h$  and an input  $x$ , the computation of  $h(x)$  must be easy.
3. The function must be collision-resistant: this means that it is hard to find two distinct messages that hash to the same result (i.e., find  $x$  and  $x'$  with  $x \neq x'$  such that  $h(x) = h(x')$ ).

For a message authentication code, the computation depends on a secondary input, the secret key. The main idea is that an adversary without knowledge of this key should be unable to forge the MAC result for any new message, even when many previous messages and their corresponding MAC results are known. The following informal definition was given by Praneel [10]. A message authentication code or MAC is a function  $h$  satisfies the following conditions:

**\*Corresponding author: Joju K.T.**

*Department of Mathematics, Prajyoti Niketan College, Pudukad, Kerala, India pin-680301*

The input  $x$  can be of arbitrary length and the result  $h(K, x)$  has a fixed length of  $n$  bits. The function has a secondary input the key  $K$ , with a fixed length of  $k$  bits.

1. Given  $h, K$  and an input  $x$ , the computation of  $h(K, x)$  must be easy.
2. Given a message  $x$  (with unknown  $K$ ), it must be hard to determine  $h(K, x)$ .
3. Even when a large set of pairs  $\{x_i, h(K, x_i)\}$  is known, it is hard to determine the key  $K$  or to compute  $h(K, x')$  for any new message  $x' \neq x_i$

**Definition 2.1** A hash function  $h: \mathcal{D} \rightarrow \mathcal{R}$  where the domain  $\mathcal{D} = \{0,1\}^*$ , and the range  $\mathcal{R} = \{0,1\}^n$  for some  $n \geq 1$ .

**Definition 2.2** A MAC is a function  $h: \mathcal{K} \times \mathcal{M} \rightarrow \mathcal{R}$  where the key space  $\mathcal{K} = \{0,1\}^k$ , the message space  $\mathcal{M} = \{0,1\}^*$ , and the range  $\mathcal{R} = \{0,1\}^n$  for  $k, n \geq 1$

Since its introduction at CRYPTO'94 the Tillich-Zemor hash function has kept on appealing cryptographers by its originality, its elegance, its simplicity and its security[4],[13]. The function computation can be parallelized and even the serial version is quite efficient as it only requires XOR, SHIFT and TEST operations. Uniform distribution of the outputs follows from a graph theoretical interpretation of the hash computation.

Markus Grassl, Ivana Illich, Spyros Magliveras and Rainer Steinwadt found collision for the same between palindrome bit strings of length  $2n+2$ [8]. In this paper we present the reinforced version of the Tillich-Zemor hash function by adding key to the same. We claim that it will resist palindrome collision. This paper is organized as follows:

The Tillich-Zemor hash function and its palindrome collision is recalled in section 2. In section.3 we present the new keyed hash function and verify that the keyed hash function resists the palindrome collision.

## 2. PRELIMINARIES

### 2.1 Tillich-Zemor Hash function

Let  $n$  be a positive integer and let  $p(x)$  be an irreducible polynomial of degree  $n$  over the field  $F_2$  ([3], [12]). Let  $A_0$  and  $A_1$  be the following two matrices:

$A_0 = \begin{pmatrix} x & 1 \\ 1 & 0 \end{pmatrix}$ ,  $A_1 = \begin{pmatrix} x & x+1 \\ 1 & 1 \end{pmatrix}$ , that have determinant 1. We call these matrices the generators of the Tillich-Zemor hash function. Let  $v = b_1 \dots b_m \in \{0, 1\}^*$ , be the bitstring representation of a message. The Tillich-Zemor hash value of  $v$  is defined as:

$$H(b_1 \dots b_m) = A_{b_1} \dots A_{b_m} \text{ mod } p(x)$$

Let  $K = F_2[x] / (p(x)) \approx F_2^n$  ([6], [7]). The image of the Tillich-Zemor hash function are the matrices of the group  $SL_2(K)$ , that is the group of matrices with elements in  $K$  and determinant one.

Let  $h(b_1 \dots b_m) = A_{b_1} \dots A_{b_m}$  be the Tillich-Zemor hash function without modular reduction.

i. e,  $h: \{0,1\}^* \rightarrow SL_2(F_2[x])$ .

### 2.2. Palindrome Collision

If  $v = b_1 \dots b_m \in \mathcal{M}$  is a bitstring of length  $m$ ; we denote  $v^r = b_m \dots b_1$ , the reversal of  $v$ , ie the reflection of  $v$  which interchanges  $b_1$  with  $b_m$ ,  $b_2$  with  $b_{m-1}$ , etc. Bitstring  $v \in \mathcal{M}$  satisfying  $v = v^r$  are known as palindrome. In order to have the palindrome collision we will make the following change in the generators.

Let  $B_0 = A_0^{-1} A_0 A_0 = A_0 = \begin{pmatrix} x & 1 \\ 1 & 0 \end{pmatrix}$  and  $B_1 = A_0^{-1} A_1 A_0 = \begin{pmatrix} x+1 & 1 \\ 1 & 0 \end{pmatrix}$ . Define the hash functions  $H'$  and  $h'$  with new generators as follows:

$$H'(b_1 \dots b_m) = B_{b_1} \dots B_{b_m} \text{ mod } p(x) \text{ and } h'(b_1 \dots b_m) = B_{b_1} \dots B_{b_m}. \text{ Then we have the following proposition [8].}$$

**Proposition 1.** Let  $v, v' \in \mathcal{M}$ . Then  $H(v) = H(v')$  iff  $H'(v) = H'(v')$ .

That is, collision for  $H$  and  $H'$  are equivalent.

In [8] Markus Grassl, Ivana Illich, Spyros Magliveras and Rainer Steinwadt observed the following property of palindrome messages.

**Proposition 2.** Let  $v$  be a palindrome of even length say  $v = b_m \dots b_1 b_1 \dots b_m$ . Let  $a_0, \dots, a_m$  be the following polynomials

$$a_i = \begin{cases} 1, & \text{if } i = 0 \\ x + b_1 + 1, & \text{if } i = 1 \\ (x + b_i)a_{i-1} + a_{i-2} & \text{if } 1 < i \leq m \end{cases}$$

Then  $h'(v) = \begin{pmatrix} a^2 & b \\ b & d^2 \end{pmatrix}$  for  $a = a_m$ ,  $d = a_{m-1}$  and for some  $b \in F_2[x]$

$$\text{Moreover, } h'(0v0) + h'(1v1) = \begin{pmatrix} a^2 & a^2 \\ a^2 & 0 \end{pmatrix}.$$

From proposition 2 we see that the square root of the upper left entries of  $h'(b_1 b_1); h'(b_2 b_1 b_1 b_2); h'(b_3 b_2 b_1 b_1 b_2 b_3);$  etc , satisfying a Euclidean algorithm sequence (in reverse order) where each quotient is either  $x$  or  $x+1$ . These sequences are often called maximal length sequences for the Euclidean algorithm or maximal length Euclidean sequences. Mesirov and Sweet[9] showed that, when  $a \in F_2[x]$  is an irreducible, there exists exactly two polynomials  $d$  such that  $a, d$  are the first terms of a maximal length Euclidean sequences. In their collision algorithm [8] they apply Mesirov and Sweet's algorithm to the irreducible polynomial  $a = p(x)$ .

**Proposition 3.** (Mesirov and Sweet) Given any irreducible polynomial  $p$  of degree  $n$  over  $F_2$ , there is a sequence of polynomials  $p_n, p_{n-1}, \dots, p_1, p_0$  with  $p_n = p$ , and  $p_0 = 1$  and additionally the degree of  $p_i$  is equal to  $i$  and  $p_i \equiv p_{i-2} \pmod{p_{i-1}}$ .

Note that once we know a polynomial  $q = p_{n-1}$  as mentioned in proposition 3 which matches our given polynomial  $p_n = p$ , the Euclidean algorithm will uniquely compute the sequence  $p_n, p_{n-1}, \dots, p_1, p_0 = 1$ .

The quotients  $x + \beta_i$  ( $i = 1 \dots n$ ) occurring in Euclid's algorithm allow us to derive the bits  $b_i$  of the palindrome in proposition 2. We have  $p_1 = x + b_1 + 1$  and therefore  $b_1 = \beta_1 + 1$ , while  $b_i = \beta_i$  for some  $i > 1$ . That is the bit  $\beta_1$  has to be inverted. Thus the desired collision will be

$$H(0\beta_n \dots \beta_1^{-1} \beta_1^{-1} \dots \beta_n 0) = H(1\beta_n \dots \beta_1^{-1} \beta_1^{-1} \dots \beta_n 1) \text{ where } \beta_i^{-1} \text{ indicates the inversion of } \beta_i.$$

### 2.3. To find the maximal length Euclidean sequence:

1. Construct a matrix  $A \in F_2^{(n+1) \times n}$  from the  $n+1$  polynomials

$$\begin{aligned} g_0 &\equiv x^0 \pmod{p(x)}, \\ g_i &\equiv x^{i-1} + x^{2i-1} + x^{2i} \pmod{p(x)} \text{ for } i = 1, 2, \dots, n \end{aligned}$$

Placing in the  $i^{\text{th}}$  row of  $A$  the coefficients

$a_{i,0}, a_{i,1}, \dots, a_{i,n-1}$  of the polynomial

$$g_i = a_{i,0} + a_{i,1}x + \dots + a_{i,n-1}x^{n-1}.$$

2. Solve the linear system  $Au^t = (10 \dots 01)^t$  where  $u = (u_1 \dots u_n)$ .

3. Compute  $q(x)$  by multiplying  $p(x)$  by  $\sum_{i=1}^n u_i x^{-i}$  and taking only the non negative powers of  $x$ .

### 2.4. To find Collision for specified parameters

For each choice of  $F_2^n = F_2[x]/(p(x))$  we obtain two bitstrings  $v_1, v_2 \in \{0, 1\}^*$  with  $H'(0v_1 v_1^r 0) = H'(1v_1 v_1^r 1)$  for  $i = 1, 2$ . ie , we obtain two collisions of bitstrings of length  $2n+2$ .  $v_2$  can be obtained by reversing  $v_1$  followed by inverting the first and last bit. By proposition .1 we have  $H(0v_1 v_1^r 0) = H(1v_1 v_1^r 1)$ . From[8] we have

#### Collision for $SL_2(F_2[X]/x^{127} + x + 1)$

By collision algorithm we have  $H(0v_1 v_1^r 0) =$

$$\begin{pmatrix} x & 1 + x^2 + x^3 + x^{64} + x^{65} + x^{96} + x^{97} + x^{112} + x^{113} + x^{120} + x^{121} + x^{124} + x^{125} + x^{126} \\ 1 & x + x^2 + x^{63} + x^{64} + x^{95} + x^{96} + x^{111} + x^{112} + x^{119} + x^{120} + x^{123} + x^{124} + x^{125} \end{pmatrix} = H(1v_1 v_1^r 1)$$

and

$$\begin{aligned} H(0v_2 v_2^r 0) &= \begin{pmatrix} x & 1 + x + x^2 + x^{64} + x^{65} + x^{96} + x^{97} + x^{112} + x^{113} + x^{120} + x^{121} + x^{124} + x^{125} + x^{126} \\ 1 & 1 + x + x^{63} + x^{64} + x^{95} + x^{96} + x^{111} + x^{112} + x^{119} + x^{120} + x^{123} + x^{124} + x^{125} \end{pmatrix} \\ &= H(1v_2 v_2^r 1) \end{aligned}$$







**Colum 2:**

$$1+x^3+x^4+x^5+x^6+x^7+x^{11}+x^{13}+x^{15}+x^{16}+x^{19}+x^{20}+x^{21}+x^{22}+x^{23}+x^{24}+x^{31}+x^{39}+x^{43}+x^{45}+x^{47}+x^{63}+x^{64}+x^{67}+x^{68}+x^{69}+x^{70}+x^{79}+x^{80}+x^{83}+x^{84}+x^{85}+x^{86}+x^{87}+x^{88}+x^{95}+x^{96}+x^{99}+x^{100}+x^{101}+x^{102}$$

$$x+x^4+x^5+x^8+x^{12}+x^{16}+x^{17}+x^{20}+x^{21}+x^{24}+x^{25}+x^{32}+x^{40}+x^{44}+x^{48}+x^{64}+x^{65}+x^{68}+x^{69}+x^{80}+x^{81}+x^{84}+x^{85}+x^{88}+x^{89}+x^{96}+x^{97}+x^{100}+x^{101}$$

$$H_1(1 v_1 v_1^r 1) =$$

**Colum 1:**

$$x^2+x^3+x^5+x^6+x^{10}+x^{13}+x^{15}+x^{21}+x^{22}+x^{27}+x^{31}+x^{34}+x^{35}+x^{39}+x^{42}+x^{45}+x^{47}+x^{50}+x^{51}+x^{63}+x^{69}+x^{70}+x^{71}+x^{79}+x^{85}+x^{86}+x^{91}+x^{95}+x^{101}+x^{102}+x^{103}$$

$$1+x+x^5+x^6+x^8+x^{10}+x^{12}+x^{17}+x^{18}+x^{21}+x^{22}+x^{25}+x^{26}+x^{32}+x^{34}+x^{40}+x^{42}+x^{44}+x^{48}+x^{50}+x^{65}+x^{66}+x^{69}+x^{70}+x^{81}+x^{82}+x^{85}+x^{86}+x^{89}+x^{90}+x^{97}+x^{98}+x^{101}+x^{102}$$

**Colum 2:**

$$1+x+x^2+x^3+x^5+x^7+x^8+x^{11}+x^{12}+x^{17}+x^{18}+x^{21}+x^{23}+x^{25}+x^{26}+x^{32}+x^{35}+x^{40}+x^{43}+x^{44}+x^{48}+x^{51}+x^{65}+x^{66}+x^{69}+x^{71}+x^{81}+x^{82}+x^{85}+x^{87}+x^{89}+x^{90}+x^{97}+x^{98}+x^{101}+x^{103}$$

$$x^2+x^3+x^6+x^7+x^9+x^{10}+x^{11}+x^{15}+x^{17}+x^{19}+x^{22}+x^{23}+x^{25}+x^{31}+x^{33}+x^{34}+x^{39}+x^{41}+x^{42}+x^{43}+x^{47}+x^{49}+x^{50}+x^{63}+x^{65}+x^{67}+x^{70}+x^{79}+x^{81}+x^{83}+x^{86}+x^{87}+x^{89}+x^{95}+x^{97}+x^{99}+x^{102}$$

It is evident that  $H_1(0 v_1 v_1^r 0) \neq H_1(1 v_1 v_1^r 1)$ . Similarly we can prove that

$H_1(0 v_2 v_2^r 0) \neq H_1(1 v_2 v_2^r 1)$ . Hence the keyed hash function resists the palindrome collision. Calculations are done using "SCILAB". The program is the following

**PROGRAM**

```

x=poly(0,'x');A0=[x,1;1,0] //matrix 1
x=poly(0,'x');A1=[x,x+1;1,1] //matrix 2
I=[1,0;0,1] //Identity Matrix
H1=[0v1v1^r0 ,or,1v1v1^r1 ,or,0v2v2^r0 ,or,1v2v2^r1]
k0=[1,0,0,0,1]
j=1
m=1
k=k0
hs=size(H1, '*')
ks=size(k0, '*')
i=ks+1
if(ks<hs) then
    while ( m<= hs-ks)
        k(1,i)=k0(1,j)
        i=i+1
        j=j+1
        if (j>ks) then
            j=1
        end
        m=m+1
    end
end

if (k(1,1)==0) then
    E=I
else

    if (H1(1,1)==0) then
        E=A0

    else
        E=A1
    end
end
    
```

```

end
if (k(1,2)==0)
    D=I
else
    if (H1(1,2)==0) then
        D=A0
    else
        D=A1
    end
end
R=E*D
for i=3:size(H1, '*')
    if (k(1,i)==0) then
        R=R*I
    else
        if (H1(1,i)==0) then
            R=R*A0
        else
            R=R*A1
        end
    end
    d11=degree(R(1,1))
    c11=coeff(R(1,1))
    for i=1:size(c11, '*')
        c11(1,i)=modulo(c11(1,i),2)
    end
    d12=degree(R(1,2))
    c12=coeff(R(1,2))
    for i=1:size(c12, '*')
        c12(1,i)=modulo(c12(1,i),2)
    end
    d21=degree(R(2,1))
    c21=coeff(R(2,1))
    for i=1:size(c21, '*')
        c21(1,i)=modulo(c21(1,i),2)
    end
    d22=degree(R(2,2))
    c22=coeff(R(2,2))
    for i=1:size(c22, '*')
        c22(1,i)=modulo(c22(1,i),2)
    end
    R(1,1)=inv_coeff(c11,d11)
    R(1,2)=inv_coeff(c12,d12)
    R(2,1)=inv_coeff(c21,d21)
    R(2,2)=inv_coeff(c22,d22)
    p= x^127+x+1           //polynomial
    [res,quo]=pdiv(R,p)
    R(1,1)=res(1,1)
    R(1,2)=res(1,2)
    R(2,1)=res(2,1)
    R(2,2)=res(2,2)
    R=abs(R)
    d11=degree(R(1,1))
    c11=coeff(R(1,1))
    for i=1:size(c11, '*')
        c11(1,i)=modulo(c11(1,i),2)
    end
    d12=degree(R(1,2))
    c12=coeff(R(1,2))
    for i=1:size(c12, '*')
        c12(1,i)=modulo(c12(1,i),2)
    end
    d21=degree(R(2,1))
    c21=coeff(R(2,1))

```



```
for i=1:size(c21, '*')
    c21(1,i)=modulo(c21(1,i),2)
end
d22=degree(R(2,2))
c22=coeff(R(2,2))
for i=1:size(c22, '*')
    c22(1,i)=modulo(c22(1,i),2)
end
R(1,1)=inv_coeff(c11,d11)
R(1,2)=inv_H1=[]
R=[]
coeff(c12,d12)
R(2,1)=inv_coeff(c21,d21)
R(2,2)=inv_coeff(c22,d22)
end
disp(R)
```

## REFERENCES

- [1] Bart Van Rompay, Analysis and Design of Cryptographic Hash Functions, MAC algorithms and Block Ciphers, Doctoral Dissertation, KU Leuven 2004D/2004/7515 ISBN 90-5682-527-5
- [2] Wieb Bosma, John Cannon, and Catherine Playoust. The Magma Algebra System I: The User Language. Journal of Symbolic Computation, 24 (1997), pp.235-265.
- [3] Christophe Petit, Jean-Jacques Quisquater, Preimages for the Tillich-Zemor hash function, Proceedings of the 17 th International Conference on Selected Areas in Cryptography pp 282-301, Springer-Verlag Berlin, Heidelberg 2011 ISBN:978-3-64[1]
- [4] Christophe Petit, Jean-Jacques Quisquater, Jean-Pierre Tillich and Gilles Zemor, Hard and easy Components of Collision Search in the Zemor-Tillich Hash Function: new Attacks and Reduced Variants with Equivalent Security. In M. Fischlin, editor, CT-RSA, volume 5473 Lecture Notes in Computer Science, pages 182-194, Springer-Verlag, 2009.
- [5] Douglas R Stinson, *Cryptography theory and practice*, Second Edition, Chapman & Hall/CRC.
- [6] John R Durbin, *Modern Algebra*, John Wiley & Sons.
- [7] Joju K.T and Lilly P.L Alternate form of Hashing with Polynomials. Proceedings of the International workshop on Cyber Security (IWCS 2k11), ISBN: 978-8206-25-7, St. Joseph's College, Irinjalakuda pp. 43-45, 2011
- [8] Markus Grassl, Ivana Ilic, Spyros Magliveras, and Rainer Steinwandt, Cryptanalysis of the Tillich-Zemor hash function, Journal of Cryptology, Volume 24 Number-1. pp 148-156.
- [9] Jill P. Mesirov and Melvin M. Sweet. Continued Fraction Expansions of Rational Expressions with Irreducible Denominators in Characteristic 2. Journal of Number Theory, 27 (1987), pp.144-148.
- [10] B. Praneel: Analysis and Design of Cryptographic Hash Functions. Doctoral Dissertation K.U. Leuven Jan. 1993
- [11] Stefan Lucks, Design principles of Iterated Hash function, ePrint Archive: Report (2004), pp.1-22.
- [12] J.P.Tillich and G. Zemor, *Hashing with SL2*, Advances in Cryptology Lecture Notes in Computer Science, vol. 839(1994), Springer-Verlag, pp. 40-49.
- [13] Vladimir Shpilrain, *Hashing with polynomials*, Proceedings of ICISC 2006, Springer (2006).

**Source of support: Nil, Conflict of interest: None Declared**