F-HASH: Securing Hash Functions Using Feistel Chaining

Duo Lei

Department of Science, National University of Defense Technology,
Changsha, China
Duoduolei@gmail.com

Abstract. The Feistel structure is well-known as a good structure for building block ciphers, due to its property of invertibility. It can be made non-invertible by fixing the left half of the input to 0, and by discarding the left half of the output bits. It then becomes suitable as a hash compression function. This paper uses the compression function to build a hash called F-Hash, based on a new structure, which is immune to recent attack styles. In this paper, a more precise evaluation method, based upon conditional probability, is given.

Keywords: Hash Function, Conditional probability, Feistel structure

1 Introduction

Most hash functions, including MD5[15] and SHA-1[14] are based upon compression functions iterated using the Merkle-Damgård structure [10, 22] with fixed IVs [36]. Since MD5 and SHA-1 were attacked by Wang et. al [6, 42, 43], more attention has been paid to the area of hash functions, with the intention of strengthening these hash functions or finding new, resilient hash functions.

The Wang attacks are based upon differential cryptanalysis, first known to academics in the early nineties [5]. Attacks against block ciphers and hash functions are very similar, as witnessed by the use of block cipher cryptanalysis techniques against hash functions [4, 26]. Increasing attention has been paid to designing hash functions using the same technology as block ciphers [4], as that technology is well-established.

One such technology is the Feistel structure, which is frequently used to build block ciphers using a $2n \rightarrow 2n$ invertible transformation. By fixing the leftmost n bits of the input to 0, and by outputting the rightmost n bits, the Feistel structure becomes a non-invertible n to n transformation that is suitable for building hash functions. We call this the FL-structure (Feistel-like structure). If the round function of the FL structure is chosen in the same way as the round function of the Feistel-based block cipher, then the new construction, when using more than two rounds, inherits many of the properties of the corresponding block cipher, with the exception of invertibility. The benefit of the FL structure is that it leverages the extensive analysis conducted on the Feistel structure.

Therefore we think that the non-invertibility of the FL structure qualifies it as a good component for building hash compression functions to resist pre-image and collision attacks.

This paper discusses building dedicated hash functions using Feistel structure, including a complete hash function based on the FL structure, which we name F-Hash. We give a proof of security of this hash function, and show that for robust underlying block ciphers, the hash function is immune against all known attacks.

In this paper, we give a new evaluation model to quantify the security of structures based upon conditional probability. We discuss the influence of the compression function's conditional probability on the entire construction's conditional probability. This informs us that the outputs of F-Hash are uniformly distributed for any fixed IV , or for any selected message, if only the compression function has the following properties: that the function is immune against adaptive chosen plaintext and adaptive chosen-key attacks and that the distribution of the function's inputs and outputs are uniform.

It is basic requirement upon the block cipher that the distribution of input and output are independent. Patarin [27–31] provided security proofs for the Feistel Structure, and Piret [32, 33] gave proofs of round functions with random permutations. Similar conclusions were also given by Vaudenay [40, 41]. Luby and Rackoff [21] introduced a method that permitted the assessment of security of some block cipher constructions. The limitations of these proofs are that although they discuss immunity against adaptive chosen plaintext attacks, they neglect adaptive chosen-key attacks. For F-Hash, we should consider security against chosen-key attacks. In this paper we give a proof that if the Feistel cipher is immune against adaptive chosen plaintext attacks and the key schedule algorithm is sufficiently non-linear, then we can build a secure F-hash.

In Section 2, we give Feistel constructions and define the new hash function. In Section 3, we provide security proofs. In Section 4, we provide a discussion and conclusion.

2 The Feistel Constructions

A Feistel structure is a general way of constructing block ciphers from simple functions. The original idea was invented by Horst Feistel [11] for use with block ciphers. The security of the Feistel structure is not obvious but comprehensive analysis of DES [13] has shown that it is a good way to construct ciphers. No weakness has been found in the Feistel structure itself. In this section, we present some constructions derived from the Feistel structure, but first introduce the structure itself.

2.1 The Feistel Compression Function

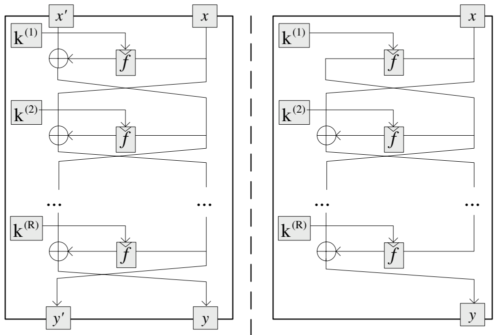
Let I_n be the set $\{0,1\}^n$ where $a, b \in I_n$, and $a\|b \in I_{2n}$ and $|a| = n$. The key schedule algorithm is $\psi(k)$ and $k^{(i)}$are the round keys, such that $k^{(0)} = k$,

$k^{(i)} = \psi(k^{(i-1)})$, $i \geq 1$. The round function of the block cipher is $f : I_n \times I_n \rightarrow I_n$ [1]. The round output $y = f(k^{(i)}, x)$, also denoted by $f_{k^{(i)}}$. Then with $\circ$ denoting composition of a function, a Feistel block cipher is denoted by $E^{Fe} : I_n \times I_{2n} \rightarrow I_{2n}$, where:

$E^{Fe}(k, x'\|x) \overset{def}{=} \Psi^R(f)(x'\|x) = \Psi(f_{k^{(R)}}) \circ \Psi(f_{k^{(R-1)}}) \circ \ldots \circ \Psi(f_{k^{(1)}}(x'\|x))$,
$y'\|y \overset{def}{=} \Psi(f_k)(x'\|x) = \begin{cases} y' = x \\ y = x' \oplus f_k(x) \end{cases}$ such that $x', x, y', y \in I_n$; $(x'\|x)^L$ and $(x'\|x)^R$: and the leftmost and rightmost n bits of binary sequence $x'\|x$, respectively. Also $\tilde{0}$ is an n-bit string with all bits equal to 0.

Also define an SPN block cipher $E^{Sp} : I_n \times I_n \rightarrow I_n$, such that $E^{Sp}(k, x) = f_{k^{(R')}} \circ f_{k^{(R'-1)}} \circ \ldots \circ f_{k^{(1)}}(x)$. Let $F_c : I_n \times I_n \rightarrow I_n$ be Feistel-Like Structured Function with round function, $F_c(k, x) = (E^{Fe}(k, \tilde{0}\|x))^R$; F_{c-1} : Feistel-Structured function with one round fewer than F_c, and E^{-1} be the inverse of E, where E is a permutation.

Remark 1. Unless otherwise noted, E^{Fe}, E^{Sp}, F_c and F_{c-1} have same key schedule algorithm $\psi(k)$ and round function f. The round function f is a permutation.



Feistel Block Cipher FL Function

Fig. 1. Contrast Between Feistel Structure and FL-Structure

Definition 1 (FL Structure). *Let function $F_c : I_n \times I_n \rightarrow I_n$, if $F_c(k, x) = \Psi^R(f)(\tilde{0}\|x) = \Psi(f_{k^{(R)}}) \circ \Psi(f_{k^{(R-1)}}) \circ \ldots \circ \Psi(f_{k^{(1)}}(\tilde{0}\|x))^R$, then we call function F_c the FL-Function (Feistel Like function) and the entire structure FL-Structure(Feistel Like Structure). This is shown in Figure 1.*

2.2 F-HASH Function

Let $z = H(m, x)$: the hash function, where m be message and x be initial value; message[2] $m \in I_{n \cdot t}$, $\overline{m}_i$ be message block with $\overline{m}_i \in I_n$, $m = \overline{m}_t\|\overline{m}_{t-1}\| \ldots \|\overline{m}_1$,

[1] We assume that the key length $\kappa = n$, with padding if required.

[2] When message block is used as key, the message block length be κ.

$\overline{m}_t \subseteq m$; A selected m be denoted $m_i \in I_{n \cdot t}, t \geq 1$; $y = F(x_m, x_h)$: the hash compression function, where x_h be chaining value, x_m be message block; $z = H^M(m, x)$: iterated hash with M-D construction, if $m = \overline{m}_t \| \ldots \| \overline{m}_1$, then $z = F(\overline{m}_t, F(\overline{m}_{t-1}, \ldots F(\overline{m}_1, x) \ldots))$ with compression function $y = F(x_m, x_h)$; Message Padding: adding 1 and zero at the end of Message.[3].

Definition 2 (Feistel Compression Function). *If the function $F_c : I_n \times I_n \rightarrow I_n$ is used as compression function of iterated hash with format $y = F_c(x_m, x_h)$, where x_h is the chaining value, we call this function the Feistel Compression Function.*

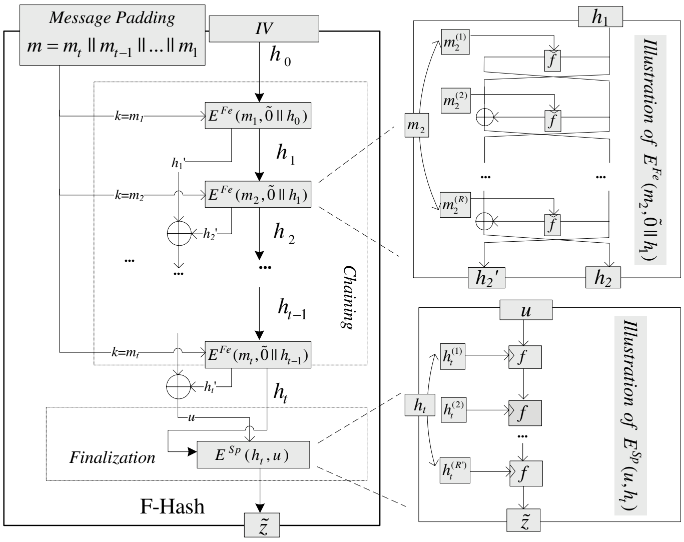


Fig. 2. F-HASH Function

Definition 3 (F-HASH). *Let F-HASH $H^F : I_{n \cdot t} \times I_n \rightarrow I_n$, $\tilde{z} = H^F(m, x)$, $m \in I_{n \cdot t}$ is defined as*

$$\begin{aligned} h_0 &= x \\ h_i &= F_c(\overline{m}_i, h_{i-1}), \qquad (i = 1, \ldots, t) \\ h_i' &= F_{c-1}(\overline{m}_i, h_{i-1}), \qquad (i = 1, \ldots, t) \\ \tilde{z} &= E^{Sp}(\bigoplus_{i=1}^{t} h_i', h_t). \end{aligned}$$

[3] That is to make thing simple. In realized design, we should select a length related padding.

Lemma 1. $h'_i \| h_i = E^{Fe}(\overline{m}_i, \tilde{0} \| h_{i-1}) \Leftrightarrow h'_i = F_{c-1}(\overline{m}_i, h_{i-1}), y = F_c(\overline{m}_i, h_{i-1})$

The figure illustration of F-HASH is given in Fig 2.

F-HASH has two way of understanding. The definition itself mean to tell us the F-HASH has two Feistel compression function. If we can give security proof of Feistel compression function and the structure, then security of F-HASH can be given, more discussion is in next section. Motivation of that kind of definition is to make the proof become easy. In fact, two simple hash structure also given. The $H^M(m,x) \overset{def}{=} h_t$ is a hash function with Merkle-Damagråd construction and Feistel compression function. And $H^O(m,x) \overset{def}{=} \bigoplus_{i=1}^{t} h'_i$ can also be seen as a new hash function. Contrast among those three kind of hash is given in following section.

The Lemma 1 implies another way of understanding. The F-HASH is build from a block cipher of $2n$ to $2n$, at last iteration, we use E^{SP} to make it un-invertible. That is an instance of transformed Double-pipe hash[25].

Lemma 2. *FL-function $y = F_c(k,x)$, with R round, is sequent to following structure. Let $x_0 = \tilde{0}$, $x_1 = x$ and $x_{i+1} = x_{i-1} \oplus f(x_i), i = 1, \ldots, R-1$ then $y = x_R$.*

The Lemma 2 tell us FL-structure itself is a iterated structure. Since, for any block cipher $E : I_n \times I_{2n} \rightarrow I_{2n}$, if we fix the left-most n input bits with zero, and output just right-most n bits of last round, then the structure become an un-invertible structure. But the structure looks strange.

3 Security Proof of F-HASH

The securities of F-HASH is based on security of Feistel Block cipher and the security of the structure.

In this section, firstly, we give a security proof of FL-Function is good hash compression function. In other word, if there exist a secure E^{FE}, then there exist a secure F_c, where secure E^{FE} is in view point of block cipher design and the secure F_c is in view point of hash function design.

The maximum gap between the secure E^{FE} and secure F_c is as follows. Most security discussions on block cipher are immune against chosen plaintext(CPA) and adaptive chosen plaintext attack(ACPA). The immunity against adaptive chosen key attack(ACKA) is not required and not very certain. For hash compression function, the security of immune against adaptive chosen message attack(ACMA)[4] and adaptive chosen chaining value attack(ACCA) are both important. We can assume the block cipher is immune against ACPA, but we can not assume the cipher is immune against ACKA. Our aim is build a secure hash compression function based on secure block cipher. And also we have to consider

[4] No weakness means can not find the inner relation between output and chaining value or output and message block.

the block cipher is invertible, what is that properties of Feistel compression function.

Secondly, we should give the proof of the structure is secure. To give the proof of this, our motivation is given the proof output distribution of the F-HASH is near to uniformly distribution, for fixed message and fixed initial value. Since, we have give the proof of the compression function is immune against ACMA and ACCA. If output distribution is near to uniform distribution, then that implies the best way of attack is exhaustive attack and the bound is near to bound of random search. Then, we give the proof of the hash is immune against collision attack and preimage attack.

Thirdly, we will discuses the hash structure being immune against some known attack or not. Since that part related with the contrast with other structure, that is given in next section.

3.1 Security of Feistel Compression Function

Since the proof of immune against ACPA on Feistel block cipher is known and ACKA on Feistel block cipher is uncertain. And the ACKA on Feistel block cipher can be converted to a ACMA on Feistel compression function. The motivation of this subsection is that, if there exist a secure Feistel block cipher, immune against ACPA, then there exist a secure Feistel compression function, immune against ACMA and ACCA. We think that is main advantage of building hash compression function based on Feistel compression function than building hash compression function from PGV[35, 7]

We give following assumption for E^{Fe}, which is based upon the known results of Feistel structured block cipher.

The motivation of Assumption 1 is to make an assumption of existing a secure Feistel block cipher.

Assumption 1 *For E^{Fe} and E^{Sp}*

1. *The ciphers are immune against CPA. CPA and ACPA have same complexity;*
2. *The distributions of plaintext and ciphertext of those ciphers are independent, for each fixed key,*
3. *The best way to find weak keys ($E^{Fe}(k, x'\|x) = E^{Fe}(k', x'\|x)$) of E^{Fe} and E^{Sp} is exhaustive key search attack based on birthday paradox;*
4. *No weakness are found in E^{Fe} and E^{Sp}.*

In the proof of security of Feistel structure[27–30], the compression function is assumed as pseudo random function. If the $\psi(k)$ is pseudo random function, then $\psi(k) \oplus k$ is still a pseudo random function.

Assumption 2 *If E^{Fe} satisfies Assumption 1, then replacing the key schedule algorithm with $k^{(i)} \overset{def}{=} \psi(k^{i-1}) \oplus k, k^{(0)} = k$, the new E^{Fe} still satisfies Assumption 1.*

The Assumption 2 implies the key schedule algorithm $\psi(k)$ has property of that does not exit i, j with $\psi(k^{(i)}_{\{j\}}) \equiv k_{\{j\}}$.

Theorem 1. *Let the round function f be the format of $f(k^{(i)}, x) = f(x \oplus k^{(i)})$, if E^{Fe} satisfies Assumption 1 and Assumption 2, then there exist FL-function $\tilde{y} = \tilde{F}_c(k, x)$ and $\tilde{y}' = \tilde{F}_{c-1}(k, x)$ satisfy following properties*

1. *$\tilde{F}_c$ is immune against ACCA and ACMA;*
2. *the distributions of $\tilde{Y}$ and X_m are independent for each fixed x_h and the distributions of $\tilde{Y}$ and X_h are independent for each fixed x_m, that are also hold for $\tilde{Y}'$;*
3. *The best way to find collision $\tilde{F}(m, x) = \tilde{F}(m', x))$ is exhaustive search attack based on birthday paradox;*
4. *There are no weakness in $\tilde{F}_c$ and $\tilde{F}_{c-1}$.*

Proof. Firstly, we give a conclusion of that. If E^{Fe} has rounds r, key schedule algorithm $\psi(k)$ and round function $f(k^{(i)}, x) = f(x) \oplus k^{(i)}$, then we have

$$E^{Fe}(k_0, \tilde{k} \| x \oplus \tilde{k})^R \oplus \tilde{k} = \tilde{F}_c(\tilde{k}, x) \tag{1}$$

where the key schedule algorithm of $\tilde{F}_c$ is $\tilde{k}^{(i)} = \psi(k_0^{(i)}) \oplus \tilde{k}$.

The proof of Equation (1) is follows

When $r = 1$

$$\begin{aligned} &(\Psi(f_{k_0^{(1)}})(\tilde{k} \| x \oplus \tilde{k}))^R = \tilde{k} \oplus f(x \oplus \tilde{k} \oplus \psi(k_0^{(1)})) \\ = \ &\tilde{0} \oplus f(x \oplus (\psi(k_0^{(0)}) \oplus \tilde{k})) \oplus \tilde{k} = (\Psi(f_{k_0^{(1)} \oplus \tilde{k}})(\tilde{0} \| x \oplus))^R \oplus \tilde{k} \end{aligned}$$

Assume that for $r < k$, the equation is true then

$$\begin{aligned} &(\Psi(f_{k_0^{(r)}} \circ \ldots \circ f_{k_0^{(1)}})(\tilde{k} \| x \oplus \tilde{k}))^R \\ = &(\Psi(f_{k_0^{(r-2)} \oplus \tilde{k}} \circ \ldots \circ f_{k_0^{(1)} \oplus \tilde{k}})(\tilde{0} \| x))^R \oplus \tilde{k} \\ &\oplus f((\Psi(f_{k_0^{(r-1)} \oplus \tilde{k}} \circ \ldots \circ f_{k_0^{(1)} \oplus \tilde{k}})(\tilde{0} \| x))^R \oplus \tilde{k} \oplus k_0^{(r)}) \\ = &(\Psi(f_{k_0^{(r-2)} \oplus \tilde{k}} \circ \ldots \circ f_{k_0^{(1)} \oplus \tilde{k}})(\tilde{0} \| x))^R \\ &\oplus f((\Psi(f_{k_0^{(r-1)} \oplus \tilde{k}} \circ \ldots \circ f_{k_0^{(1)} \oplus \tilde{k}})(\tilde{0} \| x))^R \oplus (\tilde{k} \oplus k_0^{(r)})) \oplus \tilde{k} \\ = &(\Psi(f_{k_0^{(r)} \oplus \tilde{k}} \circ \ldots \circ f_{k_0^{(1)} \oplus \tilde{k}})(\tilde{0} \| x))^R \oplus \tilde{k}. \end{aligned}$$

Secondly, We give the proof of that, if E^{Fe} satisfy Assumption 1, then $\tilde{F}_c$ satisfy the item 1,2 and 3.

Since $\tilde{F}_c$'s message block and chaining value equivalence to the left and right most n bits plaintext of E^{Fe} with a fixed key, respectively, output of $\tilde{F}_c$ equivalence to the right-most n bits ciphertext of E^{Fe} with the fixed key, then we can get the conclusion directly. □

Theorem 2. *Let the round function f be the format of $f(k^{(i)}, x) = f(x \oplus k^{(i)})$, if F_c and F_{c-1} satisfy item 1,2,3 and 4 of Theorem 1, then there exist a Feistel block cipher satisfy item 1,2,3 and 4 of Assumption 1*

Proof. The proof is analog to the proof of Theorem 1. Since we have

$$F_c(k, x) = \tilde{E}^{Fe}(k, x'_0 \| x \oplus x'_0)^R \oplus x'_0 \tag{2}$$

where the key schedule algorithm of $\tilde{E}^{Fe}$ is $x_0'^{(i)} = \psi(k^{(i)}) \oplus x'_0$.

The proof can be given similar to equation (1), we omit it. □

3.2 The Output Distribution of F-HASH

This subsection gives the output distribution of F-hash for selected massage and selected initial value. The conclusions are that, the output distribution of F-Hash for fixed message or fixed initial value is near to that of compression function. So, if we can build a secure Feistel compression function, then we can guarantee the output distribution of the F-HASH. For any fixed message or fixed initial value, the output distribution being near to uniformly distributed means that, if we can not find some inner relation between input and output value of compression function, then the complexity of exhaustive search attack on collision or preimage attack gets the minimum value.

Firstly, we give some basic notation and definition, which will be used in the output probability distribution of F-HASH.

Let X be a random variable which takes on a finite set S of values $x \in S$ with probability $P_X(x) \overset{def}{=} P(X = x)$. Also, let X', Y, Y', Z and $\tilde{Z}$ be random variables which take on finite sets of values.

Let $G : I_\kappa \times I_m \to I_n$, $y = G(k, x)$ then:

$\{(y, k, x)\} \overset{def}{=} \{(y, k, x) | k \in I_\kappa, x \in I_m, y \in I_n\}$;

$\{(y, k, x)\}^G \overset{def}{=} \{(y, k, x) | (y, k, x) \in \{(y, k, x)\}, G(k, x) = y\}$;

$\{((y_0, k, x)\}^G \overset{def}{=} \{((y_0, k, x) | (y, k, x) \in \{(y, k, x)\}^G, y = y_0\}$;

$\{(y, k, x)\}^G_{x \in \Lambda} \overset{def}{=} \{(y, k, x) | (y, k, x) \in \{(y, k, x)\}^G, x \in \Lambda\}$;

$\{\{((y_0, k, x)\}^G\}_{y_0 \in \Lambda} \overset{def}{=} \bigcup_{y_0 \in \Lambda} \{\{((y_0, k, x)\}^G\}$.

$T_{F_{c-1}} \overset{def}{=} \max\limits_{x_{m_0}, y_0} \#\{(y_0, x_{m_0}, x_h)\}^{F_{c-1}}$; $S_{F_{c-1}} \overset{def}{=} \max\limits_{x_{h_0}, y_0} \#\{(y_0, x_m, x_{h_0})\}^{F_{c-1}}$;

$T_{F_c} \overset{def}{=} \max\limits_{x_{m_0}, y_0} \#\{(y_0, x_{m_0}, x_h)\}^{F_c}$;

$S_{F_c} \overset{def}{=} \max\limits_{x_{h_0}, y_0} \#\{(y_0, x_m, x_{h_0})\}^{F_c}$.

Theorem 3 (Derived Probability). *Let function $y = G(m, x)$, $G : I_{n \cdot t} \times I_n \to I_n$, $t \in \mathbf{N}$, let the distributions of independent random variable M and X are $P_X(x)$ and $P_M(m)$, let function $\chi_{G(m,x)}(y)$ is defined as that*

$$\chi_{G(m,x)}(y) \overset{def}{=} \begin{cases} 1 & y = G(m, x) \\ 0 & y \neq G(m, x) \end{cases}$$

Then the distribution of random variable Y can be derived from X and M by

$$P_Y(y) \overset{def}{=} P_Y(y = G(M,X)) = \sum_{x \in I_n} \sum_{m \in I_{n \cdot t}} P_{XM}(x,m)\chi_{G(m,x)}(y)$$
$$= \sum_{x \in I_n} \sum_{m \in I_{n \cdot t}} P_X(x)P_M(m)\chi_{G(m,x)}(y)$$

we call the probability of Y, the derived probability of M and X.

For any $y \in I_n$, if does not exist $P_Y(y)$, then we have $P_Y(y) \overset{def}{=} 0$.

Definition 4 (Conditional probability). *Directed followed Theorem 3, the conditional probability is defined as follows*

1. $P_{Y|M=m_0}(y_0) = \sum_{x \in I_n} P_X(x)\chi_G(y_0, m_0, x)$;
2. $P_{Y|X=x_0}(y_0) = \sum_{m \in I_{n \cdot t}} P_M(m)\chi_G(y_0, m, x_0)$;
3. $P_{Y|X=x_0,M=m_0}(y_0) = \chi_G(y_0, m_0, x_0)$.

Remark 2. If X and M are uniformly distributed, which means $P_M(m) = \frac{1}{2^{n \cdot t}}$, $P_X(x) = \frac{1}{2^n}$, we use notation of $P_{\dot{Y}}(y)$, that is also holds in the conditional probability.

Let $\tilde{z} = E^{Sp}(u, z)$, $z = H^M(m, x) = h_t$, $h_0 = x$, $u = O_h(m, x) \overset{def}{=} \bigoplus_{i=1}^{t} h'_i$, $h_i = F_c(\overline{m}_i, h_{i-1})$, $h'_i = F_{c-1}(\overline{m}_i, h_{i-1})$, $m = \overline{m}_t \| \ldots \| \overline{m}_0$. x and m are independent and uniformly distributed in I_n and $\bigcup_{i=1}^{t} I_{n \cdot i}$, respectively.

Theorem 4. *If the compression function $F_c(x_m, x_h)$ satisfy item 2 of Theorem 1, then for $\tilde{z} = H^F(m, x)$, $z = H^M(m, x)$, and $u = H^O(m, x)$ we have*

1. $P_{\dot{Z}|M=m}(z) \leq 2^{-n} T_{F_c}^{\frac{|m|}{n}}$;
2. $P_{\dot{Z}|X=x}(z) \leq 2^{-n} S_{F_c}$;
3. $P_{\dot{\tilde{Z}}|M=m}(\tilde{z}) \leq 2^{-n} T_{F_{c-1}}$;
4. $P_{\dot{\tilde{Z}}|X=x}(\tilde{z}) \leq 2^{-n} S_{F_{c-1}}$;
5. $P_{\dot{U}|M=m}(\tilde{z}) \leq 2^{-n} T_{F_{c-1}}$;
6. $P_{\dot{U}|X=x}(\tilde{z}) \leq 2^{-n} S_{F_{c-1}}$.

The proof is given by deduction theory in Appendix A.

In iterated hash function $z = H(m, x)$, we consider the x can be all value in I_n, because we can redefine a hash function $H(m\|x, IV) = H(m, F(x, IV)) \overset{def}{=} H'(m, x')$. So, the distribution of H with fixed message means the distribution of H' with fixed message. Moreover, if we have $P_{\dot{Z}|M=m_0}(z_0) = p$, then we have $Pr[x \to I_n : H(m_0\|x, IV) = z_0] \approx p$.

From the Theorem 4, we have the output distribution of F-HASH and H^O is same as that of compression function. But the upper bound of $P_{\dot{Z}|M=m}(z)$ is unknown, which is relies on the realized compression function. More discussion will be given in section 4.

3.3 Immunity Against Collision Attack and Preimage Attack

Let us first give a definition about the immunity.

Definition 5. *The definitions about the advantage of A in finding Primage and Collision of function H are as follows, write $\widetilde{A}dv(A) \overset{def}{=} \max\{Adv(A)\}$ where the maximum is get the luckiest adversary's advantage, $Adv(q) \overset{def}{=} \max\{Adv(A)\}$, where the maximum is taken over adversaries that ask at most q queries. If F is invertible with F^{-1}, then A can ask queries of F and F^{-1}, the whole search space is whole space.*

– ***Fixed Start Preimage Attack***

$$\widetilde{A}dv_H^{FixP}(A) = \max_{y_0,x_0} Pr[y_0 \in I_n, x_0 \in I_n; \omega \leftarrow A^{F,H} : \omega \in \{(z_0, m, x_0)\}^H]$$

– ***Fixed Start Collision Attack***

$$\begin{aligned}\widetilde{A}dv_H^{FixC}(A) = \max_{y_0,x_0} Pr[x_0 \in I_n; \omega, \omega' \leftarrow A^{F,H} :\\ \omega, \omega' \in \sigma, \sigma \in \{\{(z_0, m, x_0)\}^H\}_{z_0 \in I_n}]\end{aligned}$$

This subsection gives security of F-HASH against collision attack and preimage attack. The conclusion of this part is based on Theorem 18 of [24], if the compression function is immune against ACMA and ACCA, and the output distribution of the iterated hash structure is given, then the security bound of collision attack and preimage attack is as follows:

- $\widetilde{A}dv_H^{FixP}(q) \leq \max\{2q\frac{S_C}{2^n}, q\frac{T_C}{2^n}\}$;
- $\widetilde{A}dv_H^{FixC}(q) \leq \max\{q(q-1)\frac{S_C}{2^n}, q\frac{T_C}{2^n}\}$

Where $T_C \overset{def}{=} \max\limits_{x_{m_0},y_0} \#\{(y_0, x_{m_0}, x_h)\}^F$, $S_C \overset{def}{=} \max\limits_{x_{h_0},y_0} \#\{(y_0, x_m, x_{h_0})\}^F$, and $y = F(x_m, x_h$ is hash compression function.

Theorem 5. *Let F_c and F_{c-1} satisfy Theorem 1, then F-HASH $\tilde{z} = H^F(m, x)$ has*

- $\widetilde{A}dv_{H^F}^{FixP}(q) \leq \max\{2q\frac{S_{F_{c-1}}}{2^n}, q\frac{T_{F_{c-1}}}{2^n}\}$;
- $\widetilde{A}dv_{H^F}^{FixC}(q) \leq \max\{q(q-1)\frac{S_{F_{c-1}}}{2^n}, q\frac{T_{F_{c-1}}}{2^n}\}$.

Proof. F_c and F_{c-1} satisfy Item 2 of Theorem 1, the best way of finding collision or preimage attack is exhaustive search. F_c and F_{c-1} satisfy Item 2 of Theorem 1, then output distribution satisfy Theorem 4, we have the conclusion from Theorem 18 of [24]. □

Theorem 5 give the upper bound of collision attack and preimage attack. The values of $T_{F_{c-1}}$ and $S_{F_{c-1}}$ are discussed in next section.

4 Discussions and Motivations

Our security discussion of F-HASH is based on the security of Feistel Block cipher. In section 3, we assume exist a secure Fesitel block cipher. In fact, there are many thing to be considered in design of FL-function and that will be given in this section.

In section 2, we mentioned three modes of hash function and give the F-HASH two ways of understanding. In this section the motivations and contrasts are given.

The properties of FL-Function influence the security of F-HASH. We will consider the two sides together.

And also, we give the motivation of E^{SP}.

4.1 The Value of T_{F_c},S_{F_c},$T_{F_{c-1}}$ and $S_{F_{c-1}}$

This subsection gives the upper bound of T_{F_c},S_{F_c},$T_{F_{c-1}}$ and $S_{F_{c-1}}$. but those values are only with maximum probability of equals 1, which is need more discussion. But the more discussion on those values are also required in block cipher design.

Let $g : I_{2n} \rightarrow I_{2n}, y'\|y = \mathtt{g}(x'\|x)$ be a random permutation, then we have[3]

$$P_{Y'\|Y|X\|X'=x_0'\|x_0}(y'\|y = \mathtt{g}(x_0'\|x_0)) = 2^{-2n}$$

then $y = (\mathtt{g}(x_0'\|x))^R$ is random function. Let $\mathtt{f}(x_0', x) \overset{def}{=} (\mathtt{g}(x_0'\|x))^R$

$$P_{Y|X=x_0}(y = \mathtt{f}(x_0', x_0)) = 2^{-n}$$

then we have[3]

$$P_{Y|X_1=x_1,X_2=x_2}(y = \mathtt{f}(x_0', x_1), y = \mathtt{f}(x_0', x_2)) = \begin{cases} 2^{-2n} & x_1 \neq x_2 \\ 2^{-n} & x_1 = x_2 \end{cases}$$

In block cipher E^{Fe}, for each fixed key, if we can not distinguish the E^{Fe} from a Pseudo-random permutation, then we have

$$P(T_{F_c} = k) = 2^{-k\cdot n}2^n, \qquad P(T_{F_{c-1}} = k) = 2^{-k\cdot n}2^n, \qquad k \in \mathbf{N}$$

If the F_c is selected as Equation 1, then we have

$$P(S_{F_c} = k) = 2^{-k\cdot n}2^n, \qquad P(S_{F_{c-1}} = k) = 2^{-k\cdot n}2^n, \qquad k \in \mathbf{N}$$

If for each $x_0'\|x_0$, we can not distinguish $E^{Fe}(k, x_0'\|x_0)$ from a random function then we have

$$P(S_{F_c} = k) = 2^{-k\cdot n}2^n, \qquad P(S_{F_{c-1}} = k) = 2^{-k\cdot n}2^n, \qquad k \in \mathbf{N}.$$

From about discussion we have the value $T_{F_{c-1}}, S_{F_{c-1}}, T_{F_c}$ and S_{F_c} have max probability of equals 1, which gives the F-HASH is immune aganist preimage attack and collision attack. $H^O(m, x)$ has same bound as $H^F(m, x)$. But the upper bound of $H^M(m, x)$ relies the FL-Function. That is one motivation of adding E^{SP} function.

4.2 Round function and Key Schedule Algorithm

In the proof of Theorem 1, we find that x' can be moved into the key schedule algorithm and for the whole discussion we assume the round function f is permutation. The most common design of round function with permutation is SPN structure. The SP structure is used in Feistel structure can result in the linear part can be moved into the previous rotund or the posterior round[23], so we prefer the round function with SPS(SBox-Linear part-Sbox) structure.

The key schedule algorithm $\psi(k)$ is assumed not to be a linear transformation. We prefer the key schedule algorithm itself is pseudo random function, which has been discussed in PhD paper of Rijmen[37].

4.3 Known Attacks on F-HASH

The other attacks on F-HASH needs more discussion, we give some known attack, let $\overline{m}$ be message block and $m_i \overset{def}{=} \overline{m}_{ii_t}\| \ldots \|\overline{m}_{i1}$.

Multi Collision[18] Suppose that multi collision is possible, for each inner collision $H^M(m_{i+1}, H^M(m_i\| \ldots \|m_1, IV)) = H^M(m'_{i+1}, H^M(m'_i\| \ldots \|m'_1, IV))$, $i \in [1, t]$, if the inner collision can make true collision which requires $O_h(m_i, IV) = O_h(m'_i, IV)$, that does not always hold when the inner collision occurs, expect each h'_i are same for m'_i and m_i. Multi collision occur with high probability when $|m_i| = n$. The complexity of finding such inner collision is $\mathcal{O}(2^n)$. But that of $H^M(m, IV)$ and $H^O(m, IV)$ are $\mathcal{O}(2^{n/2})$.

Extension Attack[36] If the extension collision is possible, when there exists an inner collision $H^M(m, x_0) = H^M(m', x_0)$, the extension should be with $O_h(m''\|m, IV) = O_h(m''\|m', IV)$, as the complexity of finding $O_h(m''\|m, IV) = O_h(m''\|m', IV)$ is $\mathcal{O}(2^{n/2})$. So such extension attack on F-HASH is $\mathcal{O}(2^n)$, but that of $H^M(m, IV)$ and $H^O(m, IV)$ are $\mathcal{O}(2^{n/2})$. When the collision is final collision $H^F(m, x) = H^F(m', x)$, not a inner collision, the extension attack is impossible.

Fixed Point Attack The requirement of success of the fixed point attack is similar to that of a multi collision attack, which requires $O_h(m, IV) = O_h(m', IV)$ and the fixed block length should be $|m_i| = n$, so we also have such attack on F-HASH is $\mathcal{O}(2^n)$, but that of $H^M(m, IV)$ and $H^O(m, IV)$ are $\mathcal{O}(2^{n/2})$.

4.4 The Round Number of F_c

In block cipher design, E^{FE} requires almost two times round number as that of E^{SP}. Since the block size of E^{FE} is $2n$ bits. For FL-function, it convert an n bits message to n bits, we think F_c require same round number as E^{SP}. However, the best round number of E^{SP} is not known, so the round number discussion is only based on avalanche effect[44] of message.

4.5 The Motivation of E^{FE}

In fact, the E^{FE} can be replace by a block cipher $E : I_n \times I_{2n} \rightarrow I_{2n}$ with fixed key, h'_i be output of left n bits and h_i be right n bits output. In that way, the block cipher E requires same round number E^{FE}. And also we should use new key schedule algorithm, for the round key of E has $2n$ bits, the round key E^{SP} has n bits. And such design not as reasonable as E^{FE}.

And also, if we random select a block cipher E of $I_n \times I_{2n} \rightarrow I_{2n}$, then we can not get the Theorem 1 or Theorem 2. In that way, we can not give the security proof of Theorem 4. Moreover, if compression function has same definition from E as F_c from E^{SP}, h'_i is last previous round output and h_i is last round output, then we can not give the proof of h_i is independent from h'_i, but that property is clear in E^{FE}.

4.6 The Motivation of E^{SP}

The Motivation of adding a new round function on last round of F-HASH is as follows:

1. The immediate motivation of E^{SP} is make the output distribution of F-HASH near to uniformly distribution. The Theorem 4 gives the clarification.
2. Some known attack on Merkle-Damagård construction is impossible for F-HASH. That is owe to the E^{SP}, Subsection 4.3 has given the detail.
3. Another motivation of E^{SP} is to make the chaining value and message block mixed completely in minimum speed. In that way, we want to design the E^{FE} with minimum round number, to make the hash fast, especially for long message.
4. Since E^{FE} is invertible, we can select a random n bit value assumed as y' to check $x' = \tilde{0}$ or not, then one try to build attack on the hash. The E^SP seems to make that kind of attack imposible.

We select a SPN function for last iteration is that, the security of SPN function is known and it can be designed easily as same round function and key schedule algorithm of E^{FE}. SPN structure not inherits all property of Feistel structure, the E^{SP} may stop some weakness of E^{FE} in iteration procedure, so we select a E^{SP} instead of F_c.

We select h_t and $\bigoplus h'_i$ for input of E^{SP} based on following consideration: oplus is a easiest way of getting a provable scheme of near uniformly distribution; h'_i is independent from h_i, $\bigoplus h'_i$ is also independent from h_i.

4.7 Conclusion

This paper presents a new way to construct hash function. Security of FL-Function relies on the security of Feistel structured block cipher. Design of FL-function requires higher design criteria than that of block cipher, all design principle of block cipher can be used in design of hash function. If we can design a secure F-HASH then we get a secure block cipher, too.

Acknowledgement The first two section of this version are rewritten by Dr.Matt Henricksen. He also gives many suggestions on reordering the remaining sections. And also, more comments will be gotten from him. Much thanks to Dr. Matt Henricksen for such editing and structural suggestions.

References

1. K. Aoki, T. Ichikawa, M. Kanda, M. Matsui, S. Moriai, J. Nakajima and T. Tokita.: Camellia: A 128-Bit Block Cipher Suitable for Multiple Platforms - Design and Analysis. SAC2000. LNCS 1281. Springer-Verlag, Berlin Heidelberg New York (2000) pp.39-56.
2. M. Bellare, K. Pietrzak, and P. Rogaway, Improved Security Analyses for CBC MACs, In Advances in Cryptology Crypto 2005, LNCS 3621, pp.527-545, 2005.
3. M. Bellare and P. Rogaway, Introduction to Modern Cryptography.
4. E.Biham. Recent advances in hash functions-the way to go. Presented at ECRYPT Conference on Hash Functions (Cracow, June 2005), see http://www.ecrypt.eu.org/stvl/hfw/Biham.ps.
5. E. Biham and A. Shamir, Diffrential cryptanalysis of DES-like cryptosystems, Journal of Cryptology, Vol.4, No.1, p..3-72, 1991.
6. E.Biham and R.Chen. Near-Collisions of SHA-0,In Advances in Cryptology CRYPTO'2004, LNCS 3152,p..290-305, 2004.
7. J.Black, P.Rogaway, and T.Shrimpton, "Black-box analysis of the block-cipher-based hashfunction constructions from PGV". In Advances in Cryptology - CRYPTO'02, LNCS 2442, Springer-Verlag, pp.320-335, 2002.
8. C.Chchin. Entropy Measures and Uncoditional Security in Cryptography, PHD thesis.
9. J.Daemen and V. Rijmen, "A new MAC Construction Alred and a Specific Instance Alpha-MAC,", Fast Software Encryption 2005, LNCS , Springer-Verlag.
10. I.Damgård. A design principle for hash functions. In G. Brassard, editor, Advances in Cryptology-CRYPTO'89, LNCS 435. Springer-Verlag, 1990.
11. H. Feistel. Cryptography and Computer Privacy. Scientific American.
12. D. Feng, W. Wu :Block Cipher Analysis and Design.
13. FIPS 46-3: Data Encryption Standard. In National Institute of Standards and Technology, Oct. 1999.
14. Helena Handschuh and David Naccache. SHACAL, 2001. Available at https://www.cosic.esat.kuleuven.ac.be/nessie/tweaks.html/shacal tweak.pdf.
15. Carlo Harpes, Gerhard Kramer, and James Massey. A generalization of linear cryptanalysis and the applicability of Matsui's Piling-up lemma. In Louis Guillou and Jean-Jacques Quisquater, editors, Advances in Cryp- tology - Proceedings of EUROCRYPT 95, LNCS 921 pp. 24-38, 1995.
16. P.Gauravaram, W.Millan, J. Gonzalez Neito and E. Dawson: 3C-A Provably Secure Pseudorandom Function and Message Authentication Code. A New mode of operation for Cryptographic Hash Function. The preliminary draft version of this work is available at eprint-2005/390 .
17. P.Junod and S. Vaudenay, FOX : a New Family of Block Ciphers, Selected Areas in Cryptography-SAC 2004,LNCS 2595, pp.131-146
18. A. Joux, Multicollisions in iterated Hash functions. Application to cascaded constructions. Proceedings Crypto 2004, Springer-Verlag LNCS 3152, pp.306-316, 2004.

19. A. Joux, P.Carribault, W. Jalby and C. Lemuet. Collisions in SHA-0. Presented at the rump session of CRYPTO 2004, 2004.
20. X. Lai and J. L. Massey: Hash functions based on block ciphers. In Advances in Cryptology Eurocrypt'92, LNCS 658. Springer-Verlag, Berlin Hei-delberg New York (1993) pp.55-70.
21. M. Luby and C. Rackoff, How to construct pseudorandom permutations from pseudorandom functions. SIAM Journal on Computing, Vol. 17, No. 2 (1988) pp.373-386.
22. R.C.Merkle, One Way Hash Functions and DES, In G. Brassard, editor, Advances in Cryptology-CRYPTO' 89, LNCS 435 Springer-Verlag, pp.428-446, 1990.
23. D. Lei, L. Chao, F. Keqin. New Observation On Camellia. Selected Area in Cryptography, SAC 2005, LNCS 3897, pp.51-64, 2006.
24. D. Lei. New Integrated proof method On Iterated Hash Structure. http://eprint.iacr.org/2006/147.
25. Stefan Lucks: A Failure-Friendly Design Principle for Hash Functions, ASIACRYPT 2005, LNCS 3788, pp.474-494, 2005.
26. Ongoing Research Areas in Symmetric Cryptography, January 2005. Avalaible at URL https://www.cosic.esat.kuleuven.ac.be/nessie/deliverables/ D.STVL.3-2.1.pdf.
27. J. Patarin About Feistel Schemes with Six (or More) Rounds, in Fast Software Encryption 1998, pp.103-121.
28. J. Patarin. Luby-Rackoff 7 Rounds are Enough for $2n^{(1-\varepsilon)}$ Security. CRYPTO'03, Springer, LNCS 2729, pp.513-529,
29. J. Patarin, Security of Random Feistel Scemes with 5 or more rounds. CRYPTO'04, LNCS 3152, pp.106-122, Springer.
30. J.Patarin, Generic Attacks on Feistel Schemes, Available from the author.
31. J.Patarin, Security of Random Feistel Schemes with 5 or more rounds, Available from the author.
32. G.Piret, Luby-Rackoff Revisited: On the Use of Permutations as Inner Functions of a Feistel Scheme,Designs, Codes and Cryptography, 39, pp.233-245, 2006
33. G.Piret, Block Ciphers: Security Proofs, Cryptanalysis, Design, and Fault Attacks, PHD, 2005.
34. B.Preneel: The State of Cryptographic Hash Functions. In Lectures on Data Security, LNCS 1561. Springer-Verlag, Berlin Heidelberg New York (1999) pp.158-182.
35. B. Preneel, R. Govaerts, and J. Vandewalle, " Hash functions based on block ciphers,", In Advances in Cryptology -CRYPTO'93, Lecture Notes in Computer Science,pages 368-378. Springer-Verlag, 1994.
36. B.Preneel, V. Rijmen, A.Bosselaers: Recent Developments in the Design of Conventional Cryptographic Algorithms. In State of the Art and Evolution of Computer Security and Industrial Cryptography. LNCS 1528. Springer-Verlag, Berlin Heidelberg New York(1998) pp.106-131.
37. V. Rijmen, Cryptanalysis and design of iterated block ciphers, Katholieke Universiteit Leuven, Belgium, 9 October 1997
38. C.E.Shannon, A Mathematical Theory of Communication, The Bell System Technical Journal, Vol.27, pp.379-423,1948.
39. C.E. Shannon, Communication Theory of Secrecy Systems, Bell System Technical Journal, Vol 28: pp.656-715, 1949.
40. S. Vaudenay. On the Lai-Massey scheme. In K. Lam, T. Okamoto, and C. Xing, editors, Advances in Cryptology - ASIACRYPT'99, LNCS 1716 pp.8-19, 2000.
41. S. Vaudenay. Decorrelation: A Theory For Block Cipher security. Journal of Cryptology, 16(4):pp.249-286, 2003.

42. X. Wang, X.Lai, D.Feng and H.Yu., Cryptanalysis of the Hash Functions MD4 and RIPEMD, EUROCRYPT 2005, LNCS 3494, pp.1-18, Springer-Verlag, 2005.
43. X. Wang, H. Yu, How to Break MD5 and Other Hash Functions, EUROCRYPT'2005, Springer-Verlag, LNCS 3494, pp.19-35, 2005.
44. A.F.Webster and S. E. Tavares. On the design of S-boxes. Advances in Cryptology-CRYPTO'85, LNCS 218, pp.523-534.

A Proof of Theorem 4

To give the proof of theorem4, we give the Lemma3.

Lemma 3. *If the compression function $F_c(x_m, x_h)$ satisfy item 2 of Theorem 1 and $i \geq 2$, then U, Z, H_i and H'_i are independent from each other, where $i \in [1, t-1]$.*

Proof. That is direct conclusion of Theorem 1. □

Let us restate the following theorem.

Theorem 6. *If the compression function $F_c(x_m, x_h)$ satisfy item 2 of Theorem 1, then F-HASH $\tilde{z} = H^F(m, x)$, $z = H^M(m, x)$ have*

1. $P_{\dot{Z}|M=m}(z) \leq 2^{-n} T_{F_c}^{\frac{|m|}{n}}$;
2. $P_{\dot{Z}|X=x}(z) \leq 2^{-n} S_{F_c}$;
3. $P_{\dot{\widetilde{Z}}|M=m}(\tilde{z}) \leq 2^{-n} T_{F_{c-1}}$;
4. $P_{\dot{\widetilde{Z}}|X=x}(\tilde{z}) \leq 2^{-n} S_{F_{c-1}}$.

Proof. The proof of Theorem 5 is given by deduction theory.

1. When $t = 1$

$$\begin{aligned} P_{\dot{Z}|M=m}(z) &\leq \max_{m_0, z_0} \sum_{x \in I_n} P_X(x) \chi_{F_c(m_0, x)}(z_0) \\ &= \max_{m_0, z_0} \sum_{i \in [1, 2^n]} 2^{-n} \#\{(z_0, m_0, x_i)\}^{F_c} \leq 2^{-n} T_{F_c} \end{aligned}$$

Suppose $t < l$, the inequality is true. When $t = l$

$$\begin{aligned} P_{\dot{Z}|M=m}(z) &= P_{\dot{Z}|M=\overline{m}_l \| m'}(z) \\ &= \sum_{x \in I_n} P_X(x) \chi_{F_c(\overline{m}_l, H^M(m', x))}(z) \\ &= \sum_{x \in I_n} \sum_{u \in I_n} \frac{1}{2^n} \cdot \chi_{F_c(\overline{m}_l, u)}(z) \cdot \chi_{H^M(m', x)}(u) \\ &= \sum_{u \in I_n} \sum_{x \in I_n} \frac{1}{2^n} \cdot \chi_{F_c(\overline{m}_l, u)}(z) \cdot \chi_{H^M(m', x)}(u) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{u \in I_n} (\chi_{F_c(\overline{m}_l,u)}(z) \cdot \sum_{x \in I_n} \frac{1}{2^n} \chi_{H^M(m',x)}(u)) \\
&= \sum_{u \in I_n} \chi_{F_c(\overline{m}_l,u)}(z) \cdot P_{\dot{U}|M'=m'}(u) \\
&\leq 2^{-n} T_{F_c}^{l-1} \sum_{u \in I_n} \chi_{F_c(\overline{m}_l,u)}(z) \\
&\leq 2^{-n} T_{F_c}^{l-1} T_{F_c} = 2^{-n} T_{F_c}^{l}
\end{aligned}
$$

2. When $t = 1$

$$
\begin{aligned}
P_{\dot{Z}|X=x}(z) &\leq \max_{x_0, z_0} \sum_{m} P_M(m) \chi_{F_{c(m,x_0)}}(z_0) \\
&= \max_{x_0, z_0} \sum_{i} 2^{-n} \#\{(z_0, \overline{m}_i, x_0)\}^{F_c} \leq 2^{-n} S_{F_c}
\end{aligned}
$$

When $t > 1$

$$
\begin{aligned}
P_{\dot{Z}|X=x}(z) &= \sum_{m \in \cup_{i=1}^{l} I_{n \cdot i}} P_M(m) P_{\dot{Z}|M=m,X=x}(z) \\
&= \sum_{\overline{m}_l \in I_n} \sum_{m' \in \cup_{i=1}^{l-1} I_{n \cdot i}} P_{M'}(m') P_{M_l}(\overline{m}_l) \\
&\quad P_{\dot{Z}|M=m,X=x} P(z = F_c(\overline{m}_l, H^M(m', x))) \\
&= \sum_{\overline{m}_l \in I_n} \sum_{m' \in \cup_{i=1}^{l-1} I_{n \cdot i}} \sum_{u \in I_n} \\
&\quad P_{M'}(m') P_{M_l}(\overline{m}_l) \cdot \chi_{F_c(\overline{m}_l,u)}(z) \cdot \chi_{H^M(m',x)}(u) \\
&= \sum_{u \in I_n} \sum_{\overline{m}_l \in I_n} P_{M_l}(\overline{m}_l) \cdot \chi_{F_c(\overline{m}_l,u)}(z) \\
&\quad \sum_{m' \in \cup_{i=1}^{l-1} I_{n \cdot i}} P_{M'}(m') \cdot \chi_{H^M(m',x)}(u) \\
&= \sum_{u \in I_n} P_{\dot{Z}|U=u}(z) P_{\dot{U}|X=x}(u) \\
&\leq 2^{-n} S_{F_c} \sum_{u \in I_n} P_{\dot{U}|X=x}(z) = 2^{-n} S_{F_c}
\end{aligned}
$$

3. $\forall\ t \geq 1$

$$
\begin{aligned}
P_{\dot{\widetilde{Z}}|M=m}(\tilde{z}) &= P_{\dot{\widetilde{Z}}|M=m}(\tilde{z} = E^{Sp}(u, z), u = O_h(m, x), z = H^M(m, x)) \\
&= \sum_{x,u,z \in I_n} P_X(x) \chi_{E^{Sp}(z,u)}(\tilde{z}) \chi_{H^M}(z, m, x) \chi_{O_h(m,x)}(u) \\
&= \sum_{u,z \in I_n} \chi_{E^{Sp}(z,u)}(\tilde{z}) \sum_{x \in I_n} P_X(x) \chi_{H^M}(z, m, x)
\end{aligned}
$$

$$\begin{aligned}
&\sum_{x\in I_n} P_X(x)\chi_{O_h(m,x)}(u)\\
&= \sum_{u,z\in I_n} \chi_{E^{Sp}(z,u)}(\tilde{z})P_{\dot{U}|M=m}(u)P_{\dot{Z}|M=m}(z)\\
&\leq \max_{u_0} P_{\dot{U}|M=m}(u_0)2^n \sum_z P_{\dot{Z}|M=m}(z) \sum_u 2^{-n}\chi_{E^{Sp}(z,u)}(\tilde{z})\\
&= \max_{u_0} P_{\dot{U}|M=m}(u_0)2^n \sum_z P_{\dot{Z}|M=m}(z)P_{\tilde{Z}|Z=z}(\tilde{z})\\
&\leq \max_{u_0} P_{\dot{U}|M=m}(u_0) \max_{z_0,\tilde{z_0}} 2^n P_{\tilde{Z}|Z=z_0}(\tilde{z_0}) \sum_z P_{\dot{Z}|M=m}(z)\\
&= \max_{u_0} P_{\dot{U}|M=m}(u_0)
\end{aligned}$$

And also

$$\begin{aligned}
P_{\dot{U}|M=m}(u) &= \sum_{x\in I_n} P_X(x)P_{U|M=m,X=x}(u = h'_1 \oplus \bigoplus_{i=2}^{t} h'_i)\\
&= \sum_{x\in I_n} P_X(x)P_{U|M=m'\|\overline{m}_1,X=x}(u = v \oplus h'_1, v = \bigoplus_{i=2}^{t} h'_i)\\
&= \sum_{x\in I_n}\sum_{v\in I_n} P_X(x)P_{UV|M=m'\|\overline{m}_1,X=x}(u = v \oplus h'_1, v = \bigoplus_{i=2}^{t} h'_i)\\
&= \sum_{x\in I_n}\sum_{v\in I_n} P_X(x)P_{U|M_1=\overline{m}_1,X=x}(u = h'_1 \oplus v)\\
&\quad P_{V|M'=m',X=x}(v = \bigoplus_{i=2}^{t} h'_i)\\
&= \sum_{v\in I_n} P_{U|M_1=\overline{m}_1}(u = h'_t \oplus v)P_{V|M'=m'}(v = \bigoplus_{i=2}^{t} h'_i)\\
&= \max P_{U|M_1=\overline{m}_1}(u)\sum_v P_{V|M'=m'}(v = \bigoplus_{i=2}^{t} h'_i) \leq \frac{T_{F_{c-1}}}{2^n}
\end{aligned}$$

4. $\forall\ t \geq 1$

$$\begin{aligned}
P_{\dot{\tilde{Z}}|X=x}(\tilde{z}) &= P_{\dot{\tilde{Z}}|X=x}(\tilde{z} = E^{Sp}(u,z), u = O_h(m,x), z = H^M(m,x))\\
&= \sum_{u,z\in I_n}\ \sum_{m\in\cup_{i=1}^{t} I_{n\cdot i}} P_M(m)P_{\tilde{Z}|\dot{U}=u,Z=z}(\tilde{z})\\
&\quad P_{\dot{U},\dot{Z}|M=m,X=x}(u = O_h(m,x), z = H^M(m,x))\\
&\quad \text{Since } P_M(x) = 2^{-\sum_i^t i\cdot n} \text{ and } u,z \text{ are independent}\\
&= \sum_{u,z\in I_n} \chi_{E^{Sp}(z,u)}(\tilde{z})
\end{aligned}$$

$$
\begin{aligned}
&\sum_{m\in\cup_{i=1}^{t} I_{n\cdot i}} P_M(m)P_{\dot{U}|M=m,X=x}(u=O_h(m,x))\\
&\sum_{m\in\cup_{i=1}^{t} I_{n\cdot i}} P_M(m)P_{\dot{Z}|M=m,X=x}(z=H^M(m,x))\\
&=\sum_{u,z\in I_n}\chi_{E^{Sp}(z,u)}(\tilde{z})P_{\dot{U}|X=x}(u)P_{\dot{Z}|X=x}(z)\\
&\le \max_{u_0} P_{\dot{U}|X=x}(u_0)2^n\sum_{z}P_{\dot{Z}|X=x}(z)\sum_{u}2^{-n}P_{\dot{\tilde{Z}}|U=u,Z=z}(\tilde{z})\\
&=\max_{u_0} P_{\dot{U}|X=x}(u_0)2^n\sum_{z}P_{\dot{Z}|X=x}(z)P_{\dot{\tilde{Z}}|Z=z}(\tilde{z})\\
&\le \max_{u_0} P_{\dot{U}|X=x}(u_0)\max_{z_0,\tilde{z_0}}2^n P_{\dot{\tilde{Z}}|Z=z_0}(\tilde{z_0})\sum_{z}P_{\dot{Z}|X=x}(z)\\
&=\max_{u_0} P_{\dot{U}|X=x}(u_0)
\end{aligned}
$$

And also

$$
\begin{aligned}
P_{\dot{U}|X=x}(u) &= \sum_{m\in\cup_{i=1}^{t} I_{n\cdot i}} P_M(m)P_{U|M=m,X=x}(u=h_1'\oplus\bigoplus_{i=2}^{t}h_i')\\
&=\sum_{m\in\cup_{i=1}^{t} I_{n\cdot i}} P_M(m)P_{U|M=m'\|\overline{m}_1,X=x}(u=v\oplus h_1', v=\bigoplus_{i=2}^{t}h_i')\\
&=\sum_{m\in\cup_{i=1}^{t} I_{n\cdot i}} P_M(m)P_{UV|M=m'\|\overline{m}_1,X=x}(u=v\oplus h_1', V=\bigoplus_{i=2}^{t}h_i')\\
&=\sum_{m\in\cup_{i=1}^{t} I_{n\cdot i}}\sum_{v\in I_n} P_M(m)P_{U|M_1=\overline{m}_1,X=x}(u=h_1'\oplus v)\\
&\quad P_{V|M'=m',X=x}(v=\bigoplus_{i=2}^{t}h_i')\\
&=\sum_{v\in I_n}\sum_{\overline{m}_t\in I_n} P_{M_t}(\overline{m}_t)P_{U|M_1=\overline{m}_1,X=x}(u=h_1'\oplus v)\\
&\quad \sum_{m'\in\cup_{i=1}^{t-1} I_{n\cdot i}} P_{V|M'=m',X=x}(v=\bigoplus_{i=2}^{t}h_i')\\
&=\sum_{v\in I_n} P_{U|X=x}(u=h_1'\oplus v)P_{V|X=x}(v=\bigoplus_{i=2}^{t}h_i')\\
&=\max_{u_0} P_{U|X=x}(u_0)\sum_{v}P_{V|X=x}(v=\bigoplus_{i=2}^{t}h_i')\le\frac{S_{F_{c-1}}}{2^n}
\end{aligned}
$$

The item 5 and item 6 have been given in item 3 and item 4. □