情報保護學會論文誌
第14券 第1號, 2004. 2

유니버설 일방향 해쉬 함수에 대한 블록 암호 기반 구성 방법

이 원 일†

고려대학교 정보보호기술연구센터

Construction of UOWHF based on Block Cipher

Wonil Lee†

CIST, Korea University

요 약

Preneel, Govaerts, Vandewalle은 1994년 [13]에서 블록 암호를 이용한 충돌 저항 해쉬 함수의 구성에 관한 64 가지 방법을 고려하였다. 그들은 64 가지 방법 중에서 12 가지가 안전하다고 주장하였으나 이에 대한 엄밀한 증명은 제시하지 않았다. Black, Rogaway, Shrimpton은 2002년 [2]에서 위의 64 가지 방법에 대하여 수학적 정의를 토대로 엄밀하게 분석한 결과를 제시하였다. 그들은 블랙 박스 안전성 모델 하에서 64 가지의 압축 함수들 중 12 가지가 충돌 저항 해쉬 함수가 되며 이에 대응하는 64 가지의 확장된 해쉬 함수들 중에서 20 가지가 충돌 저항 해쉬 함수가 됨을 증명하였다. 본 논문에서는 Preneel, Govaerts, Vandewalle이 제시한 64 가지 방법을 기초로 하여 블록 암호를 이용한 UOWHF의 구성 방법에 대한 최초의 결과를 제시한다. 본 논문에서는 Black, Rogaway, Shrimpton이 사용한 블랙 박스 안전성 모델을 통하여 64 가지의 압축 함수 집합들 중에서 30 가지의 압축 함수 집합들이 UOWHF가 되며 이에 대응하는 64 가지의 확장된 해쉬 함수 집합들 중에서 42 가지가 UOWHF가 됨을 보일 것이다. 이러한 결과는 또한 CRHF를 구성하는 것보다 UOWHF를 구성하는 것이 용이하다는 것을 간접적으로 시사한다. 또한 본 논문의 중요 결과들 중 하나는 블랙 박스 모델 하에서는 UOWHF의 안전성을 위하여 일반적으로 널리 알려진 안전성 모델 하에서는 필수적인 마스크 키들이 필요하지 않다는 것이다. 이는 UOWHF의 효율성을 매우 향상시킴을 의미한다.

ABSTRACT

Preneel, Govaerts, and Vandewalle considered the 64 basic ways to construct a collision resistant hash function from a block cipher.[13] They regarded 12 of these 64 schemes as secure, though no proofs or formal claims were given. Black, Rogaway, and Shrimpton presented a more proof-centric look at the schemes from PGV.[2] They proved that, in the black box model of block cipher, 12 of 64 compression functions are CRHFs and 20 of 64 extended hash functions are CRHFs. In this paper, we present 64 schemes of block-cipher-based universal one way hash functions using the main idea of PGV and analyze these schemes in the black box model. We will show that 30 of 64 compression function families are UOWHF and 42 of 64 extended hash function families are UOWHF. One of the important results is that, in this black box model, we don't need the mask keys for the security of UOWHF in contrast with the results in general security model of UOWHF. Our results also support the assertion that building an efficient and secure UOWHF is easier than building an efficient and secure CRHF.

Keywords: *Block cipher, Cryptographic hash function, CRHF, UOWHF, Proving security*

접수일: 2003년 12월 26일; 채택일: 2004년 2월 3일

† wonil@cist.korea.ac.kr

Ⅰ. 서 론

1.1 배 경

본 논문에서 우리는 블록 암호를 이용하여 유니버설 일방향 해쉬 함수(Universal One Way Hash Function, 이하 UOWHF라 표기함) 를 구성하는 방법에 관하여 알아본다.

UOWHF는 일반적으로 널리 알려진 충돌 저항 해쉬 함수(Collision Resistant Hash Function, 이하 CRHF라 표기함) 보다 더 약한 암호학적 기본 요소이다. CRHF의 경우, 어떤 고정된 해쉬 함수에 대한 충돌 쌍을 찾아야 하는 것이 공격자의 목표이다. 반면에 UOWHF의 경우, 공격자는 먼저 어떤 메시지(입력 값)를 자신이 선택하여 내놓아야만 한다. 이 과정이 끝난 후에 특정한 해쉬 함수가 공격자에게 주어지게 된다. 그러면 이 해쉬 함수에 대하여, 먼저 내놓은 메시지와 충돌을 발생시키는 다른 메시지를 찾아야 하는 것이 공격자의 목표이다. 만일 공격자가 이 해쉬 함수에 대한 충돌 쌍을 찾아낸다면 자신의 목표를 달성하는 것이다. 이와 같이, UOWHF에 대한 공격자는 자신이 공격할 구체적인 대상이 되는 해쉬 함수가 정해지기 전에 메시지를 선택하여 내놓아야만 하므로 공격자의 공격 과정이 CRHF에 대한 공격자보다 더욱 어렵게 된다. 따라서 UOWHF가 CRHF보다 더욱 약한 암호학적 기본 요소가 되는 것이다.

UOWHF가 CRHF의 대안으로써 주목받고 있는 이유는 다음과 같다. 첫째, UOWHF를 구성하는 것이 CRHF를 구성하는 것보다 훨씬 쉬워 보인다는 것이다. 이것은 MD4와 SHA-0, PKC'98에서 제안된 해쉬 함수 등이 (CRHF 관점에서) 공격당한 예를 통하여 알 수 있다.[3,4,7] 둘째는 대부분의 증명 가능한 전자 서명을 구성하는 데 있어서 가장 중요한 역할을 하는 해쉬 함수가 UOWHF이기만 해도 전자 서명의 안전성을 보장할 수 있다는 사실이다.[1]

이제까지 UOWHF에 대한 연구는 UOWHF인 압축 함수 집합이 존재한다는 가정 하에 이 압축 함수 집합을 이용하여 UOWHF인 확장된 해쉬 함수 집합을 구성하는 다양한 방법들의 개발에 관심이 집중되어 왔다.[1,11,14,16] 이와는 대조적으로 본 논문에서는 어떤 압축 함수가 UOWHF라고 가정하지 않으며 대신 블록 암호를 이용하여 압축 함수 집합을 구성한 후 어떠한 경우에 이 압축 함수 집합이 UOWHF가 되는가를 알아 볼 것이다. 또한 이 압축 함수 집합을 Merkle-Damgard 방법에 의하여 확장시켜 확장된 해쉬 함수 집합을 얻었을 경우 UOWHF 관점에서 어떤 경우들이 안전한 경우인가를 분석할 것이다. 이 과정에서 우리는 블록 암호를 랜덤한 전단사 함수로 가정하는 블랙 박스 안전성 모델을 이용하여 안전성을 분석한다.

[1]에서도 언급한 바와 같이 UOWHF라는 명칭은 이 암호학적 기본 요소가 가지고 있는 안전성과 관련된 성질을 효율적으로 설명하지 못하고 있다. 따라서 본 논문에서는 UOWHF가 가져야 하는 성질을 '표적 충돌 저항성 (Target Collision Resistance)' 라고 부를 것이다.

1. 2 본 논문의 결과 및 공헌 내용

우선 블록 암호 $\{E_a\}_{a\in\{0,1\}^n}$가 주어졌다고 가정하자. 여기서 키 $a\in\{0,1\}^n$에 대한 함수 E_a는 $E_a:\{0,1\}^n \to \{0,1\}^n$ 이다. 이제 $0<l<n$(Ex. $l=n/2$) 이라 놓자. 여기서 우리는 블록 암호 $\{E_a\}_{a\in\{0,1\}^n}$를 이용하여 압축 함수 집합(Compression function family) $F=\{f_k\}_{k\in\{0,1\}^l}$, $f_k:\{0,1\}^{2n-l}\to\{0,1\}^n$을 구성하고자 한다.

우리는 다음와 같이 압축 함수 집합 $F=\{f_k\}_{k\in\{0,1\}^l}$를 구성하는 64가지 방법을 정의한다. 즉, 임의의 $k\in\{0,1\}^l$ 에 대하여

$$f_k(x)=E_a(b)\oplus c$$

이다. 이 때 $x=(h,m)$, $|h|=n$, $|m|=n-l$이고 $a,b,c\in\{h,(m\|k),h\oplus(m\|k),v\}$ 이다. 여기서 v는 고정된 상수이고 $|v|=n$이다.

위에서 보다시피 각각의 a,b,c는 4가지 경우의 수를 가지므로 총 64가지 방법이 정의된다. 즉, n 과 l 이 고정되면 $F=\{f_k\}_{k\in\{0,1\}^l}$ 를 구성하는 64 가지 방법이 존재하게 된다. 그러면 이제 어떤 $k\in\{0,1\}^l$ 에 대한 압축 함수(compression function) f_k의 확장된 해쉬 함수(extended hash function) H_k를 다음과 같이 정의하자.

Function $H_k(m_1\ldots m_t)$
  for $i\leftarrow 1$ **to** t **do**
    $h_i\leftarrow f_k(h_{i-1},m_i)$

return h_t

여기서 h_0는 고정된 상수이다 (편의상 0^n이라 간주하자). 또한 각각의 $1 \le i \le t$에 대하여 $|m_i| = n - l$이다. 우리는 $\Phi = \{H_k\}_{k \in \{0,1\}^l}$를 압축 함수 집합 $F = \{f_k\}_{k \in \{0,1\}^l}$의 '확장된 해쉬 집합(Extended hash family)'이라 부르기로 하자. 여기서 확장된 해쉬 집합의 키 k가 압축 함수 집합의 키 k와 같다는 점을 주시하기 바란다. 이에 대한 부연 설명은 후반부에 자세히 하기로 한다.

여기서 우리는 위의 분류표를 이용하여 본 논문의 결과를 설명할 것이다. 위의 분류표는 64 가지의 압축 함수 집합들과 그에 대응되는 64 가지의 확장된 해쉬 집합들을 각각 세 개의 그룹으로 분류하여 놓은 것이다. 표 1을 볼 때 주의할 것은 구체적인 집합을 지시하는 인덱스로써 그림 1 안의 첫 번째 열에 표기된 인덱스와 동일한 것을 사용했다는 점이다. 우리는 Group- C3과 Group-E3 안의 원소들에 대해서는 인덱스를 부여하지 않았다. 왜냐면 Group-C3과 Group-E3 안의 원소들은 UOWHF 관점에서 모두 쉽게 공격당하기 때문이다. 다음은 표 1에 대한 자세한 설명과 더불어 본 논문의 결과 및 공헌 내용에 대한 설명이다.

우선 64 가지 압축 함수 집합들(Compression function families) 에 관한 설명을 먼저 하기로 하자. 대략적으로 설명하면 Group -C1과 Group-C2 안의 모든 압축 함수 집합들은 UOWHF가 된다.(즉, 총 64개의 압축 함수 집합들 중 30개가 UOWHF이다.) 그러나 엄밀히 이야기하면 Group-C1과 Group-C2는 다음과 같은 차이점을 가지고 있다. Group-C1의 경우, 이 그룹 안의 각각의 압축 함수 집합에 대한 임의의 공격자가 취할 수 있는 성공 확률의 상한값(upper bound)은 $(q-1)/2^{n-1}$이다.(정리 1 참조). 반면에, Group-C2의 경우, 임의의 공격자가 취할 수 있는 성공 확률의 상한값은 $(q-1)/2^{l-1}$ 이다.(정리 2 참조). 여기서 $0 < l < n$인 점을 주시하면 Group-C1의 표적 충돌 저항성(Target Collision Resistance)이 Group-C2의 표적 충돌 저항성보다 더 좋음을 알 수 있다. 다시 말하면 UOWHF 관점에서 Group-C1 이 Group -C2 보다 더 좋음을 알 수 있다. 게다가 Group-C1은 [2]에서 정의된 Group-1의 UOWHF 버전임을 주시하자. 또한 [2]를 참조하면 Group-C1의 표적 충돌 저항성은 [2]에서 정의된 Group-1의 충돌 저항성(Collision Resistance)과 거의 동일함을 알 수 있다. 그러나 Group-C2에 대한 CRHF 버전[2]은 CRHF가 아니다. 반면에 UOWHF 관점에서 Group-C2의 원소들은 (Group-C1 보다 약하지만) UOWHF가 된다. 게다가 여기서 주목할 점은 생일 공격(Birthday Attack)이 UOWHF에는 적용되지 못한다는 것이다. 그러므로 해쉬값의 크기가 많이 줄어들 수 있다. 이는 Group-C2의 안전성도(l에 의존하여) 합리적일 수 있음을 의미한다. Group-C3의 모든 원소들은 UOWHF가 아니다. 왜냐하면 단지 한 두개의 질문을 가지고 공격에 성공할 수 있기 때문이다. 이러한 공격 방법들은 [2,13]의 공격 방법들과 매우 유사하므로 본 논문에서는 이러한 공격들에 대한 묘사를 생략한다.

표 1. 64 가지 방법들의 분류.(Group-Ci 에서 'C'는 Compression function family Group-Ei 에서 'E'는 Extended hash family를 의미함.)

압축 함수 집합들 (Compression function families)
Group-C1 : $F_{1\ldots12}$ (12 schemes)
Group-C2 : $F_{\{13\ldots34\}-\{15,17,19,20\}}$ (18 schemes)
Group-C3 : (34 schemes)
확장된 해쉬 집합들 (Extended hash families)
Group-E1 : $\Phi_{1\ldots20}$ (20 schemes)
Group-E2 : $\Phi_{21\ldots42}$ (22 schemes)
Group-E3 : (22 schemes)

이제 64 가지의 확장된 해쉬 함수 집합들 (Extended hash function families)에 관하여 설명하기로 한다. 먼저 대략적으로 설명하면, Group-E1과 Group-E2 안의 모든 확장된 해쉬 함수 집합들은 UOWHF가 된다.(즉, 총 64개의 확장된 해쉬 함수 집합들 중 42개가 UOWHF이다.) 그러나 Group -E1과 Group-E2는 다음과 같은 차이점을 가지고 있다. Group-E1의 경우 공격자의 성공 확률의 상한값은 질문의 개수 q와 블록의 크기 n에 의하여 정해진다 (그림 1 참조). 반면에, Group-E2의 경우 공격자의 성공 확률의 상한값은 질문의 개수 q와 키의 크기 l에 의하여 정해진다(그림 1 참조). 여기서 또한 $0 < l < n$인 점을 주시하면 Group -E1의 표적 충돌 저항성이 Group-E2의 표적 충돌 저항성보다 더 좋음을 알 수 있다 (여기서 주의할 점은 l이 n에 아주 가까운 수일 때에는 몇 몇 경우에 Group

-E1의 표적 충돌 저항성이 Group-E2의 표적 충돌 저항성보다 더 좋지 않을 수도 있다. 그러나 l이 n에서 아주 가깝지 않으면서 작은 수인 경우에는 모두 Group-E1의 표적 충돌 저항성이 Group -E2의 표적 충돌 저항성보다 더 좋다고 말할 수 있다. 또한 우리는 본 논문의 결과에 따라 UOWHF의 안전성을 위하여 n에서 아주 가깝지 않은 작은 수 l을 선택할 것이므로 일반적으로 성립한다고 말할 수 있다). 예를 들어, 그림 1에서 $t=2$(Group -E1의 원소)와 $t=39$(Group-E2의 원소)인 경우를 비교해 보자. $t=2$ 인 경우는 공격 성공 확률의 상한값이 $q(q+1)/2^{n-l}$이고, $t=39$ 인 경우는 공격 성공 확률의 상한값이 $q(q+1)/2^{l-l}$이 된다. 따라서 $t=2$ 인 경우의 표적 충돌 저항성이 $t=39$인 경우 보다 더 좋다는 것을 알 수 있다.

게다가 Group-E1은 [2]에서 정의된 Group-1과 Group-2의 UOWHF 버전임을 주시하자. 또한 [2]을 참조하면 Group-E1의 표적 충돌 저항성은 [2]에서 정의된 Group-1과 Group-2의 충돌 저항성(collision resistance)과 거의 동일함을 알 수 있다. 그러나 Group -E2에 대한 CRHF 버전 [2]은 CRHF가 아니다. 반면에 UOWHF 관점에서 Group-E2의 원소들은(Group-E1 보다 약하지만) UOWHF가 된다. 게다가 여기서도 주목할 점은 생일 공격이 UOWHF에는 적용되지 못한다는 것이다. 그러므로 해쉬 값의 크기가 많이 줄어들 수 있으며 이는 Group-E2의 안전성도(l에 의존하여) 합리적일 수 있음을 의미한다. Group-E3의 모든 원소들은 UOWHF가 아니다. 왜냐하면 우리는 단지 한두개의 질문을 가지고서 공격에 성공할 수 있기 때문이다. 본 논문에서는 이러한 간단한 공격들에 대한 묘사는 생략하기로 한다.

1. 3 블랙 박스 안전성 모델

우리가 사용하려는 안전성 모델은 Shannon이 처음으로 사용했던 블랙 박스 안전성 모델 (Black-Box Security Model) 이다.[15] 이 안전성 모델은 [8,9,17]에서도 사용되었다. 우선 키 길이 τ와 블록 크기 n을 고정시키자. 블랙 박스 모델에서는 공격자 A에게는 오라클 E 와 E^{-1}가 주어지는데 E는 랜덤 블록 암호 $E:\{0,1\}^{\tau}\times\{0,1\}^{n}\rightarrow\{0,1\}^{n}$를 나타내고 E^{-1}는 E 의 역함수를 의미한다. 즉, 각각의 키 $a\in\{0,1\}^{\tau}$는 무작위로 선택된 $\{0,1\}^{n}$ 위에서의 전단사 함수 $E_a(\cdot)=E(a,\cdot)$를 가리키고 공격자에게는 이러한 암호화 및 복호화 오라클 E 와 E^{-1}가 주어지는 것이다. E^{-1}의 경우, 입력 (a,y)에 대하여 $E_a(x)=y$를 만족하는 x를 출력하는 오라클이다. 블랙 박스 모델에 관한 더욱 자세한 설명과 이 안전성 모델의 의미에 대한 논의는 [2,15]를 참조하기 바란다.

1. 4 증명 방법

[2]에서 Black, Rogaway, Shrimpton은 널리 알려진 증명 방식인 귀류법 (암호 이론 분야에서는 Reduction이라는 용어로 많이 알려져 있음)을 이용하여 그들이 정의한 Group-1을 분석하였다. 그러나 그들은 그들이 정의한 Group-2를 분석하는 데에는 귀류법을 사용하지 않았다. 왜냐면 Group-2 안에 있는 압축 함수들은 모두 CRHF가 되지 않기 때문이다.([2] 참조)

본 논문에서는 Group-E1과 GroupE2 안에 있는 모든 확장된 해쉬 집합을 분석할 때 귀류법을 사용하지 않을 것이다. [2]에서 Group-2를 분석할 때와 마찬가지로 모든 확장된 해쉬 집합을 증명할 때 기초 확률론을 이용하여 직접적으로 증명할 것이다. 이는 우리가 이용하려는 안전성 모델이 블랙 박스 모델이기 때문에 가능한 것이다.

여기서 언급하고 싶은 것은 이러한 블랙 박스 모델 하에서는 UOWHF의 안전성을 위하여(일반적으로 널리 알려진 안전성 모델 하에서는 꼭 필요한) 마스크 키들이 필요하지 않다는 것이다.[1,11,14,16] 이것은 본 논문의 중요 결과들 중 하나이다.

II. 정 의

2. 1 표기법

τ 와 n을 1 보다 큰 정수라고 놓자. 이 때 블록 암호는 함수 $E:\{0,1\}^{\tau}\times\{0,1\}^{n}\rightarrow\{0,1\}^{n}$로 정의되며 각각의 키 $a\in\{0,1\}^{\tau}$에 대하여 $E_a(\cdot)=E(a,\cdot)$는 $\{0,1\}^{n}$ 위에서의 전단사 함수(permutation) 이다. 만일 E가 어떤 블록 암호일 때 E^{-1}는 이것의 역함수 - 즉 각각의 키 $a\in\{0,1\}^{\tau}$에 대하여 $E_a^{-1}(y)$는 $E_a(x)=y$를 만족하는 x를 의미한다 - 이다. $Bloc(\tau,n)$ 을 모든 블록 암호 $E:\{0,1\}^{\tau}\times\{0,1\}^{n}\rightarrow$

$\{0,1\}^n$ 들의 집합이라고 하자. 앞으로 $Bloc(\tau, n)$에서 무작위로 어떤 원소를 뽑는 행위가 의미하는 것은 각각의 $a \in \{0,1\}^\tau$ 에 대하여 랜덤한 전단사 함수 $E_a(\cdot)$ 를 뽑는 것을 의미하는 것으로 약속하자.

'(블록 암호 기반) 해쉬 함수 집합(Hash function family)'은 $\{\phi_k\}_{k\in\{0,1\}^l}$를 의미한다. 이때 ϕ_k : $Bloc(\tau, n)\times D \rightarrow R$, $\tau, n, c \geq 1$, $D \subseteq \{0,1\}^*$, $R = \{0,1\}^c$ 이다. 함수 ϕ_k는 어떤 원소 $\alpha \in D$ 이 주어졌을 때 E-오라클을 사용하여 $\phi_k^E(\alpha) = \phi_k(E, \alpha)$ 을 계산하는 어떤 프로그램이 주어져 있어야만 한다.

어떤 해쉬 함수 집합 $F = \{f_k\}_{k\in\{0,1\}^l}$, $f_k : Bloc(\tau, n) \times D \rightarrow R$를 고려하자. 이때 이 집합에서 $a + b \geq c$를 만족하는 어떤 $a, b \geq 1$에 대하여 $D = \{0,1\}^a \times \{0,1\}^b$ 라면 $F = \{f_k\}_{k\in\{0,1\}^l}$를 ***'압축 함수 집합 (Compression function family)'*** 이라고 한다. 이제 $h_0 \in \{0,1\}^a$를 고정시키자. 압축 함수 집합 $F = \{f_k\}_{k\in\{0,1\}^l}$ $f_k : Bloc(\tau, n) \times D \rightarrow R$ 의 ***'확장된 해쉬 함수 집합 (Extended hash function family)'*** 은 해쉬 함수 집합 $\Phi = \{H_k\}_{k\in\{0,1\}^l}$, $H_k : Bloc(\tau, n) \times (\{0,1\}^b)^* \rightarrow \{0,1\}^a$로써 다음과 같이 정의된다. 각각의 $m_1 \ldots m_t \in (\{0,1\}^b)^*$에 대하여 $H_k^E(m_1 \ldots m_t) = h_t$이고 이때 $h_1 = f_k^E(h_{i-1}, m_i) = f_k(E, (h_{i-1}, m_i))$이다. 우리는 앞으로 대부분의 경우에 f와 H에 대한 윗첨자 E를 생략할 것이다.

어떤 유한 집합 S로부터 무작위로(randomly) 어떤 원소를 선택하여 x라고 표기하는 과정(experiment)을 $x \xleftarrow{R} S$로 표기하기로 하자. 공격자는 하나 또는 그 이상의 오라클에 접근할 수 있는 어떤 알고리즘을 말한다. 앞으로 이러한 오라클들을 윗첨자로 표기할 것이다.

2.2 블록 암호의 키 길이

본 논문에서는 [2, 13]과 같이, 블록 암호의 키 길이 τ가 평문 및 암호문의 길이 n과 같다고 가정한다. 즉, 본 논문에서는 $E : \{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1\}^n$과 $Bloc(n, n)$인 경우만을 고려할 것이다.

2.3 UOWHF

(블록 암호 기반) 해쉬 함수 집합 $\{\phi_k\}_{k\in\{0,1\}^l}$, ϕ_k : $Bloc(n, n)\times D \rightarrow R$, $n, c \geq 1$ $D \subseteq \{0,1\}^*$, $R = \{0,1\}^c$, 의 표적 충돌 저항성(Target collision resistance)을 측정하기 위하여 우리는 먼저 $Bloc(n, n)$으로부터 블록 암호 알고리즘 E가 무작위로 선택되었다고 가정하자. 공격자 A는 $E(\cdot,\cdot)$와 $E^{-1}(\cdot,\cdot)$에 대한 오라클(oracle)을 부여받는다. 그리고 공격자 A는 다음과 같은 게임을 시작한다.

1. $A^{E,E^{-1}}$는 어떤 원소 $\alpha \in D$ 를 선택하여 내놓는다.(우리는 이 원소 α을 표적 메시지(Target message)라고 부를 것이다. 또한 이 과정을 'Stage 1' 이라고 부를 것이다.)
2. $A^{E,E^{-1}}$는 $\{0,1\}^l$으로부터 무작위로 선택된 키 k 를 부여받는다.
3. $A^{E,E^{-1}}$는 $\alpha \neq \alpha'$이면서 $\phi_k^E(\alpha) = \phi_k^E(\alpha')$를 만족시키는 α'을 찾아야 한다. (우리는 이 원소 α'을 형제 메시지(sibling message) 라고 부를 것이다. 또한 이 과정을 'Stage 2' 이라고 부를 것이다.)

여기서 형제 메시지는 키에 의존하여 정해질 수 있지만, 표적 메시지는 키와는 독립적으로 선택된다는 점을 주목하라. 우리는 공격자가 할 수 있는 질문(query)들의 총 개수와 표적 메시지와 충돌을 일으키는 형제 메시지를 찾는 데에 성공할 확률 간의 관계를 분석함으로써 안전성을 분석할 것이다.

정의 2. (압축 함수 집합의 표적 충돌 저항성)
$F = \{f_k\}_{k\in\{0,1\}^l}$, $f_k : Bloc(n, n) \times \{0,1\}^a \times \{0,1\}^b \rightarrow \{0,1\}^c$를 블록 암호 기반 압축 함수 집합이라고 하자. A를 어떤 공격자라고 하자. 그러면 A의 '이득(advantage)'은 다음과 같은 실수이다.

$$Adv_F^{comp}(A) =$$
$$\Pr[E \xleftarrow{R} Bloc(n, n); (h, m) \leftarrow A^{E,E^{-1}};$$
$$k \xleftarrow{R} \{0,1\}^l; (h', m') \leftarrow A^{E,E^{-1}}:$$
$$(h, m) \neq (h', m') \ \& \ f_k^E(h, m) = f_k^E(h', m')].$$

임의의 $q \geq 1$에 대하여 $Adv_F^{comp}(q) = \max_A\{Adv_F^{comp}(A)\}$ 라고 정의하자. 이때 이 값은 최대 q개(E질문의 개수 + E^{-1} 질문의 개수) 의 질문을 오라클에게 던질 수 있는 모든 공격자 A가 얻을 수 있는 이득의 최대값을 의미한다.

우리는 아래에 비슷한 방법으로 확장된 해쉬 함수 집합 $F = \{f_k\}_{k\in\{0,1\}^l}$, $f_k : Bloc(n, n) \times \{0,1\}^a \times$

t	j	$h_i=$	Ext low-bnd	Ext up-bnd	t	j	$h_i=$	Ext low-bnd	Ext up-bnd
	1	$E_x(x_i)\oplus v$	1	1	19	33	$E_{x_i}(w_i)\oplus v$	$.632(q-1)/2^n$	$q(3q-1)/2^n$
22	2	$E_{h_{i-1}}(x_i)\oplus v$	$1/2^n$	$4q/2^{l-1}$	26	34	$E_{h_{i-1}}(w_i)\oplus v$	$1/2^n$	$q(3q-1)/2^l$
13	3	$E_w(x_i)\oplus v$	$.632(q-1)/2^n$	$q(3q-1)/2^n$	36	35	$E_{w_i}(w_i)\oplus v$	$.632(q-1)/2^n$	$q(q+1)/2^{l-1}$
	4	$E_v(x_i)\oplus v$	1	1	37	36	$E_v(w_i)\oplus v$	$1/2^n$	$q(3q-1)/2^l$
	5	$E_x(x_i)\oplus x_i$	1	1	20	37	$E_{x_i}(w_i)\oplus x_i$	$.632(q-1)/2^n$	$q(3q-1)/2^n$
1	6	$E_{h_{i-1}}(x_i)\oplus x_i$	$(q-1)/(2^{n+2}-4)$	$q(q+1)/2^{n-1}$	4	38	$E_{h_{i-1}}(w_i)\oplus x_i$	$(q-1)/(2^{n+2}-4)$	$q(q+1)/2^{n-1}$
9	7	$E_w(x_i)\oplus x_i$	$.632(q-1)/2^n$	$q(q+1)/2^{n-1}$	27	39	$E_{w_i}(w_i)\oplus x_i$	$.632(q-1)/2^n$	$q(q+1)/2^{l-1}$
	8	$E_v(x_i)\oplus x_i$	1	1	38	40	$E_v(w_i)\oplus x_i$	$(q-1)/(2^{n+2}-4)$	$q(3q-1)/2^l$
	9	$E_x(x_i)\oplus h_{i-1}$	1	1	8	41	$E_{x_i}(w_i)\oplus h_{i-1}$	$.632(q-1)/2^n$	$q(q+1)/2^{n-1}$
21	10	$E_{h_{i-1}}(x_i)\oplus h_{i-1}$	$1/2^n$	$4q/2^{l-1}$	28	42	$E_{h_{i-1}}(w_i)\oplus h_{i-1}$	$1/2^n$	$q(q+1)/2^{l-1}$
11	11	$E_{w_i}(x_i)\oplus h_{i-1}$	$.632(q-1)/2^n$	$q(q+1)/2^{n-1}$	29	43	$E_{w_i}(w_i)\oplus h_{i-1}$	$.632(q-1)/2^n$	$q(q+1)/2^{l-1}$
	12	$E_v(x_i)\oplus h_{i-1}$	1	1	30	44	$E_v(w_i)\oplus h_{i-1}$	$1/2^n$	$q(q+1)/2^{l-1}$
	13	$E_x(x_i)\oplus w_i$	1	1	6	45	$E_{x_i}(w_i)\oplus w_i$	$.632(q-1)/2^n$	$q(q+1)/2^{n-1}$
3	14	$E_{h_{i-1}}(x_i)\oplus w_i$	$(q-1)/(2^{n+2}-4)$	$q(q+1)/2^{n-1}$	2	46	$E_{h_{i-1}}(w_i)\oplus w_i$	$(q-1)/(2^{n+2}-4)$	$q(q+1)/2^{n-1}$
14	15	$E_w(x_i)\oplus w_i$	$.632(q-1)/2^n$	$q(3q-1)/2^n$	39	47	$E_{w_i}(w_i)\oplus w_i$	$.632(q-1)/2^n$	$q(q+1)/2^{l-1}$
	16	$E_v(x_i)\oplus w_i$	1	1	40	48	$E_v(w_i)\oplus w_i$	$(q-1)/(2^{n+2}-4)$	$q(q+1)/2^{l-1}$
15	17	$E_x(h_{i-1})\oplus v$	$.632(q-1)/2^n$	$q(3q-1)/2^n$		49	$E_x(v)\oplus v$	1	1
	18	$E_{h_{i-1}}(h_{i-1})\oplus v$	1	1		50	$E_{h_{i-1}}(v)\oplus v$	1	1
16	19	$E_w(h_{i-1})\oplus v$	$.632(q-1)/2^n$	$q(3q-1)/2^n$	41	51	$E_{w_i}(v)\oplus v$	$.632(q-1)/2^n$	$q(q+1)/2^{l-1}$
	20	$E_v(h_{i-1})\oplus v$	1	1		52	$E_v(v)\oplus v$	1	1
17	21	$E_x(h_{i-1})\oplus x_i$	$.632(q-1)/2^n$	$q(3q-1)/2^n$		53	$E_{x_i}(v)\oplus x_i$	1	1
23	22	$E_{h_{i-1}}(h_{i-1})\oplus x_i$	$1/2^l$	$q(3q-1)/2^l$	31	54	$E_{h_{i-1}}(v)\oplus x_i$	$1/2^l$	$q(q+1)/2^{l-1}$
12	23	$E_{w_i}(h_{i-1})\oplus x_i$	$.632(q-1)/2^n$	$q(q+1)/2^{n-1}$	32	55	$E_{w_i}(v)\oplus x_i$	$.632(q-1)/2^n$	$q(q+1)/2^{l-1}$
35	24	$E_v(h_{i-1})\oplus x_i$	$1/2^l$	$q(3q-1)/2^l$		56	$E_v(v)\oplus x_i$	1	1
5	25	$E_x(h_{i-1})\oplus h_{i-1}$	$.632(q-1)/2^n$	$q(q+1)/2^{n-1}$		57	$E_{x_i}(v)\oplus h_{i-1}$	1	1
	26	$E_{h_{i-1}}(h_{i-1})\oplus h_{i-1}$	1	1		58	$E_{h_{i-1}}(v)\oplus h_{i-1}$	1	1
10	27	$E_{w_i}(h_{i-1})\oplus h_{i-1}$	$.632(q-1)/2^n$	$q(q+1)/2^{n-1}$	33	59	$E_{w_i}(v)\oplus h_{i-1}$	$.632(q-1)/2^n$	$q(q+1)/2^{l-1}$
	28	$E_v(h_{i-1})\oplus h_{i-1}$	1	1		60	$E_v(v)\oplus h_{i-1}$	1	1
7	29	$E_x(h_{i-1})\oplus w_i$	$.632(q-1)/2^n$	$q(q+1)/2^{n-1}$		61	$E_{x_i}(v)\oplus w_i$	1	1
24	30	$E_{h_{i-1}}(h_{i-1})\oplus w_i$	$1/2^l$	$q(q+1)/2^{l-1}$	34	62	$E_{h_{i-1}}(v)\oplus w_i$	$1/2^l$	$q(q+1)/2^{l-1}$
18	31	$E_w(h_{i-1})\oplus w_i$	$.632(q-1)/2^n$	$q(3q-1)/2^n$	42	63	$E_{w_i}(v)\oplus w_i$	$.632(q-1)/2^n$	$q(q+1)/2^{l-1}$
25	32	$E_v(h_{i-1})\oplus w_i$	$1/2^l$	$q(q+1)/2^{l-1}$		64	$E_v(v)\oplus w_i$	1	1

그림 1. 64 가지 확장된 해쉬 집합들에 관한 결과. 첫 번째 열은 본 논문 전반에 걸쳐 사용하는 인덱스를 의미한다. 두 번째 열의 인덱스는 [13]에서 사용된 인덱스이다. 세 번째 열은 어떤 $k\in\{0,1\}^l$에 대한 $f_k(h_{i-1},m_i)$ 를 나타낸다. 이때 x_i는 $(m_i\|k)$를 w_i는 $x_i\oplus h_{i-1}$를 나타낸다. 네 번째, 다섯 번째 열은 공격 성공 확률의 상한 및 하한 값을 나타낸다.

$\{0,1\}^b\to\{0,1\}^c$에 대한 공격자의 이득을 정의한다.

정의 1. (확장된 해쉬 함수 집합의 표적 충돌 저항성) $\Phi=\{H_k\}_{k\in\{0,1\}^l}$, $H_k: Bloc(n,n)\times D\to R$ 를 블록 암호 기반 확장된 해쉬 함수 집합이라고 하자. A를 어떤 공격자라고 하자. 그러면 A의 '이득'은 다음과 같은 실수이다.

$$Adv_{\Phi}^{coll}(A)=\Pr[E\xleftarrow{R} Bloc(n,n);\ M\leftarrow A^{E,E^{-1}};k\xleftarrow{R}\{0,1\}^l;M'\leftarrow A^{E,E^{-1}}:\ M\neq M'\ \&\ H_k^E(M)=H_k^E(M')].$$

임의의 $q \geq 1$에 대하여 $Adv_{\Phi}^{coll}(q) = \max_{A}\{Adv_{\Phi}^{coll}(A)\}$ 라고 정의하자. 이때 이 값은 최대 q개(E 질문의 개수 + E^{-1} 질문의 개수) 의 질문을 오라클에게 던질 수 있는 모든 공격자 A가 얻을 수 있는 이득의 최대값을 의미한다.

2.4 본 논문에서의 기본 가정들

본 논문에서는 이론 전개 및 증명의 효율성을 위하여 아래와 같은 기본 가정들이 성립한다고 약속한다.

- 첫째, 공격자는 자신이 이전에 던졌던 질문에 대한 답을 이미 알고 있을 경우, 이와 관련된 질문을 다시 던지지 않는다. 즉, 만일 공격자 A가 질문 $E_a(x)$를 던지고 이에 대한 답으로 y를 이미 받았을 경우 A는 질문 $E_a(x)$나 $E_a^{-1}(y)$을 다시 던지지 않는다.
- 둘째, 확장된 해쉬 집합 $\Phi = \{H_k\}_{k\in\{0,1\}^l}$ 에 대한 표적 메시지와 충돌하는 형제 메시지를 찾는 공격자 A가 어떤 표적 메시지를 내놓았다고 가정하자. 그러면, 공격자가 수행해야 하는 게임의 정의에 의하여, 이 후 공격자 A는 랜덤 키 k를 받게 된다. 여기서 우리는 이러한 랜덤 키 k를 받은 직후에 바로 공격자 A는 $H_k^E(M)$를 계산한다고 가정한다. 이 계산 과정에서 물론 A는 $H_k^E(M)$를 계산하기 위하여 필요한 E 또는 E^{-1} 질문을 던진다. 그리고 마지막으로 A가 표적 메시지와 충돌을 일으키기 위하여 생성한 M'을 출력하기 전에 A는 이미 $H_k^E(M')$를 계산했다고 가정하자. 이때에도 마찬가지로 A는 $H_k^E(M')$를 계산하기 위하여 필요한 E 또는 E^{-1} 질문을 던진다.
- 비슷하게, 압축 함수 집합 $F = \{f_k\}_{k\in\{0,1\}^l}$ 에 대해서도 위 두 번째 항과 동일한 가정을 한다.[1)]

Ⅲ. 압축 함수 집합의 표적 충돌 저항성

3. 1 Group-C1

우리는 우선 Group-C1 안에 있는 압축 함수 집합들에 관하여 분석한다. 즉, $t \in \{1,\ldots,12\}$에 대한 F_t들을 분석한다. 우리는 모든 $t \in \{1,\ldots,12\}$에 대하여 공격자의 성공 확률의 상한값은 질문의 개수 q와 블록 크기 n에 의하여 표현되며 이 값은 적절한 질문의 개수 q에 대하여 상당히 작은 값임을 보일 것이다. 이는 Group-C1 안의 모든 압축 함수 집합들이 UOWHF 임을 의미한다.

정리 1. (Group-C1의 표적 충돌 저항성)
$n \geq 1$ 이고 $t \in \{1,\ldots,12\}$라 하자. 그러면 임의의 $q \geq 1$에 대하여 $\mathrm{Adv}_{F_t}^{comp}(q) \leq \frac{q-1}{2^{n-1}}$ 이 성립한다.

(증명) 우리는 여기서 $F_1 = \{f_k\}_{k\in\{0,1\}^l}$, $f_k(h,m) = E_h(m\|k) \oplus (m\|k)$인 경우에 초점을 맞추기로 한다. $A^{?,?}$를 압축 함수 집합 F_1에 대한 공격자라고 하자. 그리고 총 q개의 질문을 E 또는 E^{-1}오라클에게 던질 수 있다고 가정하자. 우리는 좌측 오라클이 $E \xleftarrow{R} Bloc(n,n)$에 의하여 설정되고 우측 오라클은 E^{-1}에 의하여 설정된다. 이러한 실험(experiment)은 A의 관점에서 그림 2에서 정의된 알고리즘인 $SimulateOracles(A,n)$와 동일하다. $SimulateOracles(A,n)$을 실행시켜 얻은 출력을 $((x_1,s_1,y_1),\ldots,(x_q,s_q,y_q),\ out)$이라고 하자.

> Algorithm $SimulateOracles(A,n)$
> Initially, $i \leftarrow 0$ and $E_s(x) = undefined$ for all $(s,x) \in \{0,1\}^n \times \{0,1\}^n$
> Run $A^{?,?}$, answering oracle queries as follows:
> When A asks a query (s,x) to its left oracle:
> $i \leftarrow i+1$; $s_i \leftarrow s$; $x_i \leftarrow x$; $y_i \xleftarrow{R} \overline{Range}(E_s)$;
> $E_s(x) \leftarrow y_i$; return y_i to A
> When A asks a query (s,y) to its right oracle:
> $i \leftarrow i+1$; $s_i \leftarrow s$; $y_i \leftarrow y$; $x_i \xleftarrow{R} \overline{Domain}(E_s)$;
> $E_s(x_i) \leftarrow y$; return x_i to A
> When A halts, outputting a string out:
> return $((x_1,s_1,y_1),\ldots,(x_i,s_i,y_i),out)$

그림 2. 블록 암호 오라클에 대한 시뮬레이션. $Domain(E_s)$은 $E_s(x)$가 더 이상 $undefined$가 아닌 x들의 집합이고 $\overline{Domain}(E_s) = \{0,1\}^n - Domain(E_s)$ 이다. $Range(E_s)$은 $y = E_s(x)$가 더 이상 $undefined$가 아닌 y들의 집합이고 $\overline{Range}(E_s) = \{0,1\}^n - Range(E_s)$ 이다.

1) 위의 기본 가정들을 만족하지 않는 공격자 A^*는 이와 비슷한 계산 복잡도와 동일한 성공확률을 가지면서 위의 기본 가정들을 만족시키는 다른 공격자 A로 쉽게 변환될 수 있다. 따라서 이는 위의 기본 가정들이 일반성을 잃지 않는 타당한 가정임을 의미한다.

(h, m)을 A의 표적 메시지라고 하자. k를 A가 표적 메시지 (h, m)를 내보내고 나서 A에게 주어진 랜덤한 키라고 하자. 이 때 A가 공격에 성공했다고 가정해보자. 이것이 의미하는 것은 $(h', m') \neq (h, m)$이고 $f_k(h', m') = f_k(h, m)$를 만족하는 (h', m')을 A가 출력으로 내보냈다는 것이다. 우리가 지금 초점을 맞추고 있는 f_k 의 정의에 의하여 이것이 의미하는 것은 $E_k(m\|k) \oplus (m\|k) = E_{h'}(m'\|k) \oplus (m'\|k)$이다. 게다가 2. 4 절에서 제시한 본 논문의 기본 가정에 의하여 이것은 다음을 의미한다. 즉, 아래 식들을 만족하는 서로 다른 $1 \le i \neq j \le q$가 존재한다.

$$(x_i, s_i, y_i) = (m\|k, h, E_k(m\|k))$$
$$(x_j, s_j, y_j) = (m'\|k, h', E_{h'}(m'\|k))$$
$$y_i \oplus x_i = y_j \oplus x_j$$

우리는 위의 사건이 일어날 확률이 아주 작음을 보일 것이다. 여기서 다시 주시할 점은 다음과 같다. 즉, 본 논문의 기본 가정들에 의하여 공격자 A는 랜덤 키 k를 부여받은 즉시 표적 메시지 (h, m)의 해쉬값을 계산해야 한다. 여기서 w를 Stage 1과 표적 메시지 (h, m)의 해쉬값을 계산하는 데에 이용한 총 질문의 개수라고 놓자. 따라서 w는 다음 부등식 $1 \le w \le q$를 만족한다.

SimulateOracles(A, n)의 실행과정에서 만일 $w = 1$이면 C_w를 공사건(null event)이라 하고 만일 $2 \le w \le q$이면 C_w를 $y_w \oplus x_w = y_j \oplus x_j$를 만족하는 j가 $\{1, \ldots, w-1\}$ 안에 존재하는 사건이다. 임의의 $w+1 \le i \le q$ 에 대하여 C_i를 $y_i \oplus x_i = y_w \oplus x_w$인 사건이라고 하자. *Simulat eOracles*(A, n)의 실행 과정에서 임의의 $w \le i \le q$에 대하여 x_i 또는 y_i가 적어도 $2^n - (i-1)$의 개수를 가지는 집합으로부터 무작위로 선택된다는 점을 주시하기 바란다.

우선 w인 경우부터 고려해보자. 이 경우, 사건 C_w의 정의에 의하여 $\Pr[C_w] \le \frac{w-1}{2^n-(w-1)}$가 성립한다. 반면에 $w+1 \le i \le q$인 경우 우리는 $\Pr[C_i] \le \frac{1}{2^n-(i-1)}$과 같이 쓸 수 있다. 그러므로 만일 $q \le 2^{n-1}$이면 아래 부등식이 성립한다.

$$\begin{aligned}\mathrm{Adv}_{F_1}^{\mathrm{comp}}(A) &\le \Pr[C_w \vee \cdots \vee C_q] \\ &\le \Pr[C_w] + \sum_{i=w+1}^{q} \Pr[C_i] \\ &\le \frac{w-1}{2^n-(w-1)} + \sum_{i=w+1}^{q} \frac{1}{2^n-(i-1)} \\ &\le \frac{w-1}{2^n-2^{n-1}} + \sum_{i=w+1}^{q} \frac{1}{2^n-2^{n-1}} \\ &= \frac{q-1}{2^{n-1}}\end{aligned}$$

위의 부등식은 $q > 2^{n-1}$이면 무의미해지므로 우리는 $q \le 2^{n-1}$라는 조건을 생략해도 될 것이다. 따라서 우리는 $\mathrm{Adv}_{F_1}^{\mathrm{comp}}(q) \le \frac{q-1}{2^{n-1}}$이라고 결론내릴 수 있다. 지금까지의 증명은 F_1에 관한 것이었다. 그러나 $F_{2,\ldots,12}$에 대한 증명도 거의 동일하므로 생략하도록 한다.

3.2 Group-C2

본 절에서는 Group-C2 안에 있는 압축 함수 집합들에 관하여 분석한다. 즉, $t \in B$, $B = \{13, \ldots, 34\} - \{15, 17, 19, 20\}$에 대한 F_t들을 분석한다. 우리는 모든 $t \in B$에 대하여 공격자의 성공 확률의 상한값은 질문의 개수 q와 키의 크기 l에 의하여 표현되며 이 값은 적절한 질문의 개수 q에 대하여 상당히 작은 값임을 보일 것이다. 대략적으로 설명하면, 이는 Group-C1 안의 모든 압축 함수 집합들이 UOWHF 임을 의미한다.

정리 2. (Group-C2의 표적 충돌 저항성)
$n \ge 1$ 이고 $t \in B$, $B = \{13, \ldots, 34\} - \{15, 17, 19, 20\}$라 하자. 그러면 임의의 $q \ge 1$에 대하여 $\mathrm{Adv}_{F_t}^{\mathrm{comp}}(q) \le \frac{q-1}{2^{l-1}}$이 성립한다.

정리 2에 대한 증명은 정리 1의 증명과 유사하므로 생략한다.

Ⅳ. 확장된 해쉬 함수 집합의 표적 충돌 저항성

4.1 Group-E1

본 절에서는 Group-E1 안에 있는 확장된 해쉬 함수 집합들을 분석한다. 즉, $t \in \{1, \ldots, 20\}$에 대한 Φ_t들을 분석한다. 우리는 모든 $t \in \{1, \ldots, 20\}$에 대하여 공격자의 성공 확률의 상한값은 질문의 개수 q와 블록 크기 n에 의하여 표현되며 이 값은 적절

한 질문의 개수 q에 대하여 상당히 작은 값임을 보일 것이다. 이는 Group-E1 안의 모든 확장된 해쉬 함수 집합들이 UOWHF 임을 의미한다.

여기서 우리는 $t=1$인 경우, 즉 Φ_1 만을 분석한 결과를 제시할 것이다. 왜냐면 다른 경우들은 아래의 증명과 비슷한 방법으로 모두 증명되기 때문이다. 모든 결과들은 그림 1에 제시해 놓았다.

정리 3. $n\geq 1$ 이라 하자. 그러면 임의의 $q\geq 1$에 대하여 $\mathrm{Adv}^{\mathrm{uowhf}}_{\Phi_1}(q)\leq \frac{q(q+1)}{2^{n-1}}$ 이 성립한다.

(증명) $h_0, v\in\{0,1\}^n$을 고정시키자. Φ_1은 압축 함수 집합 F_1에 대응되는 확장된 해쉬 함수 집합 임을 주시하라. 이 때 $F_1=\{f_k(h,m)\}_{k\in\{0,1\}^l}$, $f_k(h,m)=E_k(m\|k)\oplus(m\|k)$이다.

우리는 이제 방향 그래프(directed graph) $G=(V_G,E_G)$를 다음과 같이 정의하자. $V_G=\{0,1\}^n\times\{0,1\}^n\times\{0,1\}^n$이고 어떤 유향 간선(arc) $(x,s,y)\rightarrow(x',s',y')$가 E_G안의 원소가 되기 위해서는 오직 $s'=x\oplus y$일 때에만 가능하다.

Φ_1에 대한 공격자를 $A^{?,?}$라 하자. 우리는 좌측 오라클이 $E\cdot^R Bloc(n,n)$에 의하여 설정되고 우측 오라클은 E^{-1}에 의하여 설정된다고 가정한다. 그리고 총 q개의 질문을 E 또는 E^{-1} 오라클에게 던질 수 있다고 가정하자. 그러면 우리는 $\mathrm{Adv}^{\mathrm{uowhf}}_{\Phi_1}(q)\leq \frac{q(q+1)}{2^{n-1}}$ 이 됨을 보여야만 한다.

우선 알고리즘 *SimulateOracles*(A,n)을 실행한다. A가 이 알고리즘을 실행하는 과정에서 방향 그래프 G의 정점들을 다음과 같은 방법으로 색칠 (coloring)한다.

- 초기에 모든 G의 정점들은 색칠이 되지 않은 상태이다.
- A가 E-질문 (s,x)를 던지고 이에 대한 답으로 y를 받거나 A가 E^{-1}-질문 (s,y)를 던지고 이에 대한 답으로 x를 받을 때, 만일 $s=h_0$가 성립하면 (x,s,y)에 빨강 색칠을 하고 그렇지 않을 경우 (x,s,y)에 검정 색칠을 한다.

본 논문의 기본 가정에 의하여 공격자가 던지는 모든 질문은 이전에 아무런 색칠이 되지 않은 (오직) 하나의 정점에 빨강색 또는 검정색 칠을 하게 만든다.

우리는 논의의 진행을 위해 필요한 몇가지 정의를 또한 내리고자 한다. 만일 G의 어떤 정점이 빨강색이나 검정색으로 칠해졌다면 우리는 이 정점에 색이 칠해졌다고 말하기로 하자. 만일 G안의 어떤 경로 (path) 가 있어서 이 경로를 구성하는 모든 정점들에 색이 칠해졌다면 이 경로에 색이 칠해졌다고 말하기로 하자. 만일 G안의 어떤 경로가 색이 칠해졌고 이것이 공격자의 표적 메시지에 대응하는 경로라면 이 경로를 표적 경로(target path)라고 부르기로 하자. 만일 두 정점 (x,s,y)와 (x',s',y')가 $x\oplus y=x'\oplus y'$를 만족하면 두 정점 (x,s,y)와 (x',s',y')가 충돌한다고 말하기로 하자. 이제 G 안의 어떤 두 개의 서로 다른 경로 P 와 P'가 있어서 모두 색이 칠해져있고 두 개의 경로 모두 빨간색이 칠해진 정점으로 시작하고 경로의 마지막 정점 두 개가 서로 충돌하면 이 경로 P 와 P'가 충돌한다고 말하기로 하자.

C를 최대 q개의 질문을 던질 수 있는 공격자의 공격 결과로써 방향 그래프 G안에 어떤 표적 경로와 충돌을 발생시키는 경로가 생겨날 사건으로 정의하자. 그러면 다음 Claim 1과 2가 성립한다.

Claim 1. $\mathrm{Adv}^{\mathrm{uowhf}}_{\Phi_1}(A)\leq \Pr[C]$

Claim 2. $\Pr[C]\leq \frac{q(q+1)}{2^{n-1}}$

Claim 1에 대한 증명은 [2]에 있는 증명과 매우 유사하므로 본 논문에서는 생략한다. Claim 2에 대한 증명은 [18]을 참조하기 바란다. 따라서 Φ_1에 대한 정리는 Claim 1과 2를 조합하여 얻어진다.

$\Phi_{2,\ldots,20}$에 대한 결과는 위의 증명 방식과 비슷하게 얻어지므로 생략한다.

4.2 Group-E2

본 절에서는 Group-E2 안에 있는 확장된 해쉬 함수 집합들을 분석한다. 즉, $t\in\{21,\ldots,42\}$에 대한 Φ_t들을 분석한다. 우리는 모든 $t\in\{21,\ldots,42\}$에 대하여 공격자의 성공 확률의 상한값은 질문의 개수 q와 키의 크기 l에 의하여 표현되며 이 값은 적절한 질문의 개수 q에 대하여 상당히 작은 값임을

보일 것이다. 이는 Group-E2 안의 모든 확장된 해쉬 함수 집합들이 UOWHF 임을 의미한다.

여기서 우리는 $t=21$인 경우, 즉 Φ_{21} 만을 분석한 결과를 제시할 것이다. 왜냐면 다른 경우들은 아래의 증명과 비슷한 방법으로 모두 증명되기 때문이다. 모든 결과들은 그림 1에 제시해 놓았다.

정리 4. $n\geq 1$ 이라 하자. 그러면 임의의 $q\geq 1$에 대하여 $\mathrm{Adv}^{\mathrm{uowhf}}_{\Phi_{21}}(q)\leq \frac{4q}{2^{l-1}}$이 성립한다.

정리 4에 대한 증명은 정리 3의 증명과 유사하므로 생략한다. 나머지 Group-E2 안의 원소들에 대한 증명도 정리 4의 증명과 유사하므로 본 논문에서는 생략한다.

V. Group E1과 E2의 표적 충돌 저항성에 대한 Matching Attack

[2]에서와 마찬가지로, 우리는 42 가지의 모든 경우에 대하여 하한값을 계산할 수 있다. 이 모든 경우에 대한 결과는 그림 1을 참고하기 바란다.

정리 5. ($H_{1,2,3,4,38,40}$에 대한 표적 충돌쌍 찾기). $i\in\{1,2,3,4,38,40\}$, $n\geq 2$, $1\leq l<n$라 하자. 이때 $q\in[1,\cdots,(2^{n-l}+3)/3]$인 임의의 홀수 q에 대하여, $Adv^{uowhf}_{H_i[n]}(q)\geq(q-1)/(2^{n+2}-4)$이 성립한다.

$Perm(n)$은 $\{0,1\}^n$에서의 모든 전단사 함수들의 집합이고, $P_q(\{0,1\}^n)$은 $\{0,1\}^n$에 속하는 q개의 원소로 이루어진 모든 집합을 모은 집합이라고 하자. 정리 5에 대한 증명은 다음의 보조 정리 1로부터 얻어지며, 보조 정리 1에 대한 증명은 [2]에서의 증명과 유사하므로 생략한다.

보조 정리 1. $n\geq 2$이고, $1\leq l<n$이라 하자. 그러면 임의의 $k\in\{0,1\}^n$에 대하여, 다음이 성립한다. $q\in[1,\cdots,(2^{n-l}+3)/3]$인 임의의 홀수 q에 대하여,

$$\Pr[\pi \xleftarrow{R} Perm(n);\ \{x_1,\cdots,x_q\} \xleftarrow{R} P_q(\{0,1\}^{n-l}):$$

$$\exists i\ such\ that\ 2\leq i\leq q \text{ and } \pi(x_1\|k)\oplus(x_1\|k)=\pi(x_i\|k)\oplus(x_i\|k)]\geq(q-1)/(2^{n+2}-4).$$

(정리 5의 증명) H_1의 경우를 고려하자. 또한 $h_0\in\{0,1\}^n$를 고정시킨다. 이때 A를 오라클 E,E^{-1}을 갖는 공격자라고 하자. A는 $m_1 \xleftarrow{R} \{0,1\}^{n-l}$을 선택하여 $M=m_1$을 내놓는다. $\{0,1\}^l$으로부터 무작위로 선택된 키 k를 부여받으면, A는 $y_1=E_{h_0}(m_1\|k)\oplus(m_1\|k)$를 계산한다. 그 이후 A는 $1\leq r<s\leq q$에 대해 $m_r\neq m_s$인 $m_1,m_2\cdots m_q \xleftarrow{R} \{0,1\}^{n-l}$를 선택하고, $2\leq i\leq q$에 대해 $y_i=E_{h_0}(m_i\|k)\oplus(m_i\|k)$를 계산한다. 그리고나서 A가 $y_1=y_r$인 $r\in[2,\cdots,q]$를 찾는다면 형제 메시지 $M'=m_r$를 출력하고, 그렇지 않으면 0(실패)를 출력한다. 그러면 A의 공격 성공확률의 하한값을 계산해 보자. 이를 위해 먼저 $\pi=E_{h_0}$로 놓자. 정의에 의해 π는 $Perm(n)$에서 무작위로 선택된 원소이므로, 우리는 보조 정리 1을 이용하여 A가 $y_1=E_{h_0}(m_r\|k)\oplus(m_r\|k)$을 만족하는 $r\in[2,\cdots,q]$을 찾을 확률이 적어도 $(q-1)/(2^{n+2}-4)$ 임을 알 수 있다. 이와 같은 공격과 분석 방법은 다른 경우에 대해서도 비슷하게 확장될 수 있다.

다음의 정리 6, 7, 8에 대한 증명은 정리 5에 대한 증명 방식과 비슷하게 하므로 생략한다.

정리 6 ($H_{5,6,\cdots,19,20,27,29,32,33,36,39,41,42}$에 대한 표적 충돌쌍 찾기). $i\in\{5,6,\cdots,19,20,27,29,32,33,36,39,41,42\}$, $n\geq 2$, $1\leq l<n$라 하자. 이때 $q\in[1,\cdots,2^{n-l}]$인 임의의 q에 대하여, $Adv^{uowhf}_{H_i[n]}(q)\geq 0.632(q-1)/2^n$이 성립한다.

정리 7 ($H_{21,22,26,28,30,37}$에 대한 표적 충돌쌍 찾기). $i\in\{21,22,26,28,30,37\}$, $n\geq 2$, $1\leq l<n$라 하자. 이때 $q=1$에 대하여, $Adv^{uowhf}_{H_i[n]}(q)\geq 1/2^n$이 성립한다.

정리 8 ($H_{23,24,25,31,34,35}$에 대한 표적 충돌쌍 찾기). $i\in\{23,24,25,31,34,35\}$, $n\geq 2$, $1\leq l<n$라 하자. 이때 $q=1$ 대하여, $Adv^{uowhf}_{H_i[n]}(q)\geq 1/2^l$이 성립한다.

참 고 문 헌

[1] M. Bellare, and P. Rogaway, "Collision-resistent hashing: towards making UOWHFs practical," *Proceedings of CRYP-*

TO'97, pp 470-484, 1997.

[2] J. Black, P. Rogaway and T. Shrimpton, "Black-Box Analysis of the Block-Cipher-Based Hash-Function Constructions from PGV," *CRYPTO'02*, LNCS Vol.2442, pp. 320-335, 2002.

[3] F. Chabaud and A. Joux. "Differential Collisions in SHA-0," *CRYPTO'98*, LNCS 1462, Springer-Verlag, pp. 56-71, 1998

[4] Donghoon Chang, Jaechul Sung, Soohak Sung, Sangjin Lee, Jongin Lim. "Full-Round Differential Attack on the Original Version of the Hash Function Proposed at PKC'98," *SAC2002*, pp. 168-182, 2002.

[5] I. B. Damgard. "Collision free hash functions and public key signature schemes," *EUROCRYPT'87*, LNCS,304, pp. 203-216, 1988.

[6] I. B. Damgard. "A design principle for hash functions," *CRYPTO'89*, LNCS 435, pp.416-427, 1990.

[7] H. Dobbertin. "Cryptanalysis of MD4," *Fast Software Encryption*, LNCS 1039, Springer- Verlag, pp. 53-69, 1996.

[8] S. Even and Y. Mansour, "A Construction of a cipher from a single pseudo-random permutation," *ASIACRYPT'91*, LNCS Vol. 739, pp. 210-224, 1992.

[9] J. Kilian and P. Rogaway, "How to protect DES against exhaustive key search," *Journal of Cryptology*, 14(1): pp. 17-35, 2001. Earlier version in CRYPTO' 96.

[10] R. C. Merkle. "One way hash functions and DES," *CRYPTO'89*, 1989.

[11] I. Mironov. "Hash functions: from Merkle-Damgard to Shoup," *EUROCRYPT 2001*, pp. 166-181, 2001.

[12] M. Naor and M. Yung. "Universal one-way hash functions and their cryptographic applications," *ACM Symposium on Theory of Computing*, pp. 33-43, 1989.

[13] B. Preneel, R. Govaerts and J. Vandewalle, "Hash functions based on block ciphers:A synthetic approach," *CRYPTO'93*, LNCS, pp. 368-378, 1994.

[14] W. Lee, D. Chang, S. Lee, S. Sung, and M. Nandi, "New Prallel Domain Extender for UOWHF," *ASIACRYPT '03*, LNCS 2894, pp. 208-227, Dec 2003.

[15] C. Shannon, "Communication Theory of Secrecy Syetems," *Bell Systems Technical Journal*, 28(4): pp. 656-715, 1949.

[16] V.Shoup. "A composition theorem for universal one-way hash functions," *EUROCRYPT 2000*, pp. 445-452, 2000.

[17] R. Winternitz, "A secure one-way hash function built from DES," *In Proceedings of the IEEE Symposium on Information Security and Privacy*, IEEE Press, pp. 88-90, 1984.

〈著者紹介〉

이 원 일 (Wonil Lee) 정회원

1998년 2월: 고려대학교 수학과 학사
2000년 2월: 고려대학교 수학과 석사
2003년 8월: 고려대학교 수학과 박사
2000년 8월~현재: 고려대학교 정보보호기술연구센터 연구원
〈관심분야〉 해쉬 함수, 블록 암호, 스트림 암호, 암호 프로토콜