암호 시스템의 증명 가능 안전성에 관한 고찰

백준상, 김광조

한국정보통신대학원대학교

A Note On Provable Security of Cryptosystems

Joonsang Baek, Kwangjo Kim

Information and Communications University

요 약

본 논문에서는 각종 암호 시스템이나 암호 프로토콜의 안전성 검증의 중요한 기법인 증명 가능 안전성(provable security)에 관하여 고찰한다. 증명 가능 암호 시스템을 구성하는 방법으로 랜덤 오라클 패러다임과 랜덤 오라클 모델을 이용한 공개키 암호 시스템의 구성 방법을 살펴보며, 랜덤 오라클과 비교하여 최근 제안된 새로운 암호 개념인 오라클 해쉬의 정의와 특징을 기술한다. 또한 오라클 해쉬의 응용 분야에 대해 논하며 오라클 해쉬와 연관된 증명 가능 안전성에 관한 새로운 응용 문제들을 제시한다.

1. 서론

개방된 시스템의 보안에 대한 요구가 증가함에 따라, 강력한 보안을 제공하기 위한 암호 기술도 급속히 발전하고 있다. 암호 프로토콜의 효율성과 보안에 대한 연구가 계속 진행 중에 있으며, 많은 암호 프로토콜도 제안되었다. 예전에는 안전하다고 여겨지던 암호 프로토콜이나 공개키 암호시스템(이하 간단히 암호 시스템이라 한다.)에 대한 취약점이 끊임없이 발견되고 있으며, 이에 대한 대응책도 계속해서 연구되고 있다[6, 12, 18]. 따라서, 아무도 어떤 암호 시스템이 과연 안전한가에 대한 의구심에 한마디로 대답하기 어려운 것으로 인식되어 왔으나 최근 '증명 가능 안전성(provable security)'이라는 개념이 암호 시스템의 안전성의 분석에 관한 중요한 연구로 부각되고 있다. 어떤 암호 시스템을 제안하고 그것에 대한 취약점이 발견되면 그 취약점을 보완하고, 또 다른 공격이 발견되면 그 취약점을 보완하는 이 전의 암호 연구 방식이었다면, 이 증명 가능 안전성을 제공하는 접근 방식은 암호 시스템을 설계하기 전에 가능한 공격 모델을 정형화하고 보안 요구사항에 대한 명확한 정의를 내린 다음, 여러 암호학적 요소들을(cryptographic primitives) 이용하여 설계한 암호 시스템의 안전성을 수학적으로 엄밀히 증명하는 것이다[1, 4, 16, 17].

90년 대 초에 제안된 랜덤 오라클 패러다임[1]은 효율적이면서도 안전성이 증명 가능한 암호

시스템을 설계하는 방법을 제공하여 준다. 또한 Cramer와 Shoup가 Crypto '98에서 발표한 능동 선택 암호문 공격(adaptive chosen ciphertext attack)에 안전한 암호 시스템[10]은 랜덤 오라클을 사용하지 않고도 효율적인 암호 시스템을 구성할 수 있는 예를 제시하였다.

본 논문의 구성은 다음과 같다. 2장에서는 랜덤 오라클을 고찰하고 3장에서는 랜덤 오라클을 이용한 공개키 암호 시스템의 설계에 대하여 자세히 살펴보며, 4장에서는 랜덤 오라클과 비교하여 최근 제안된 오라클 해쉬의 정의와 성질을 기술한다. 또한 오라클 해쉬에 관련된 새로운 응용 문제를 5장에서 제시한다.

2. 랜덤 오라클 패러다임

2.1 신문 퍼즐 문제(News Paper Puzzle Problem)

랜덤 오라클을 정의하기에 전에 랜덤 오라클이 의미하는 바를 Canetti가 고안한 문제[7]를 예로 들어 설명한다. 신문사의 편집인인 Alice는 신문에 퍼즐을 싣고 싶은데 짧은 길이의 스트링 c를 신문에 붙여 퍼즐을 푼 사람은 자신이 올바른 답을 냈다는 것을 확인할 수 있지만, 퍼즐을 풀지 못한 사람들에게는 정답 x 에 대한 어떤 부분 정보도 누출시키지 않기를 원한다.

이 문제를 암호학적으로 어떻게 해결할 수 있을까? 언뜻 떠오르는 생각은 암호학적 해쉬함수(cryptographic hash function)를 이용하는 것이다. 즉, 정답인 암호학적 해쉬함수 h를 이용하여 x의 해쉬 값 $c(=h(x))$를 계산하여 신문에 싣는 것이다. 암호학적인 해쉬함수는 일방향성(one-wayness)을 가지기 때문에 원상(preimage)인 x를 알 수 없고 또한 충돌회피성(collision freeness)의 성질도 가지고 있어 정답이 아닌 값을 이용하여 같은 해쉬 값을 생성할 수 없다. 따라서 신문 퍼즐 문제의 해답이 될 수 있을 것으로 보인다. 하지만 문제를 좀 더 면밀히 살펴보면 '어떤 부분 정보도 누출시키지 않는다'라는 가정이 있음을 알 수 있다. 어떤 함수가 일방향성을 가지고 있다는 말은 원상의 '전체'를 알기가 계산량적으로 불가능하다(computationally infeasible)는 것이지 부분 정보의 누출이 없다는 것을 의미하지는 않는다. 하지만 만약 해쉬 함수 h가 이상적인 함수라서 입력값에 대해 완전히 랜덤한 출력값을 생성한다면 이 함수는 신문 퍼즐 문제의 해답이 될 수 있다. 사실 실제의 세계에 신문 퍼즐 문제의 해답은 존재하지 않는다. 바로 이 신문 퍼즐 문제의 해답이 랜덤 오라클인 것이다.

2.2 랜덤 오라클의 의미

2.1의 신문 퍼즐 문제에서 언급된 바와 같이 랜덤 오라클은 주어진 입력에 대해 완전히 랜덤한 출력값을 생성하는 성질이 있다. 랜덤 오라클을 이용하여 암호 시스템의 안전성을 증명할 때에는 공격자 또는 정직한 참가자가 랜덤 오라클에 접근하여 질문(query)을 하고(질문의 회수는 다항식 유계(polynomial bound)를 갖는다.) 랜덤 오라클은 이 질문에 대한 답(answer)을 주는 형태를 많이 이용한다. 즉, 랜덤 오라클 $H(\cdot)$가 x를 첫 번째 질문으로 받았다면 랜덤 오라클은 x의 길이 k번의 동전 던지기 결과값을 길이 k인 스트링으로 출력한다. 그리고 두 번째 질문 x'에 대하여 $x'=x$

이면 이전의 질문 x에 해당하는 결과값과 같은 값을, $x' \neq x$ 이면 다시 k번의 동전 던지기 결과값을 길이가 k인 스트링으로 출력한다.

공격자가 랜덤 오라클에 q_H번의 질문을 했다고 가정하면 랜덤 오라클이 가지는 완전 랜덤성 때문에 랜덤 오라클에 질문을 던지지 않고 출력을 추측할 수 있는 확률은 정확히 $q_H/2^k$(k는 입력의 길이)가 된다. 랜덤 오라클은 이와 같이 입력에 대한 출력이 결정적이며 완전한 랜덤성이라는 서로 양립하기는 힘든 두 성질을 동시에 가지고 있다. 앞서 열거된 특징으로 인해 랜덤 오라클은 '이상적인 랜덤 함수(ideal random function)'라고 불리는 것이다.

이 이상적인 랜덤 오라클을 이용하여 암호 프로토콜의 안전성을 증명하는 것을 랜덤 오라클 모델[1]이라하며 다음은 랜덤 오라클 모델의 형식적인 정의이다.

정의 1 [랜덤 오라클 모델] Ω를 유한한 길이의 스트링의 집합 $\{0,1\}^*$에서 무한한 길의의 스트링의 집합 $\{0,1\}^\infty$으로의 사상들의 집합이라 하자. Ω에서 유한한 길이(예를 들어, $\{0,1\}^a$)의 스트링의 집합을 유한한 길이의 스트링의 집합 (예를 들어, $\{0,1\}^b$)으로의 사상 H를 균일하고, 랜덤하게 선택했다고 가정하자. 어떤 프로토콜 Π 내의 함수들이 이 H에 접근이 가능하다면, Π는 랜덤 오라클 모델 하에서 정의된다고 한다.

3. 랜덤 오라클의 응용 - 공개키 암호

랜덤 오라클은 여러 암호 시스템에 사용[5, 15, 16] 되고 있지만, 공개키 암호의 안전성 증명에 매우 폭넓게 응용되고 있다[1, 2, 3, 4, 13]. 랜덤 오라클 모델하에서 설계된 증명 가능 공개키 암호시스템은 효율성이 매우 높기 때문에 실용적인 가치가 있다. 실제로 랜덤 오라클 모델 하에서 안전성이 증명 가능한 대표적인 공개키 암호 시스템 구성 방법인 OAEP(Optimal Asymmetric Encryption Padding)는 PKCS(Public-Key Cryptography Standards) #1에 표준화 되어있고 최근의 Fujisaki-Okamoto 기법(이하, F-O 기법)[13]도 표준화 추진 중에 있다.

원래 공개키 암호 시스템의 안전성에 대한 연구는 80년 대 초반부터 계속되고 있다. 초기의 공개키 암호 시스템의 안전성은 단순히 평문을 공개키로 암호화하면 비밀키를 알기 전에는 암호문을 복호화할 수 없다는 것을 의미했다. 그러나 Goldwasser가 확률적 암호 시스템을 제안[14]하면서 제기한 계산량적 구별불가능성(computational indistinguishability)은 공개키 암호의 안전성 연구가 새 국면을 맞이하는 전환점이 되었다. 간단히 말해 계산 불가능성이란 공격자가 두 평문 m_1과 m_2을 임의로 선택하여 암호화한 암호문 $E_{pk}(m_b)$($b \in_R \{0,1\}$)을 보고 그것이 평문 m_1을 암호화한 것인지 m_2를 암호화한 것인지 맞출 확률이 거의 $1/2$로 같다는 것이다. 이것은 암호 시스템의 안전성을 논하기 위해서는 단순히 주어진 함수값의 원상을 알아내는 것이 계산량적으로 어렵다는 일방향성의 성질만으로는 안전성을 논하기 불충분하다는 것을 시사한다. 이 계산량적 구별 불가능성은 공격자는 공개키 pk에 대응되는 비밀키 sk를 모르더라도 평문에 대한 부분 정보를 얻을 가능성이 충분히 있다는 사실을 잘 반영한 개념이라고 할 수 있다.

다른 한 가지의 공개키 암호의 안전성을 논의하는 데 중요한 개념은 비유연성(non-malleability)이다. 비유연성이란 Dolev[11]등에 의해 제시되었는데, 공격자가 공격하고자하는 암호문 c의 평문 m과 연관된 평문 m'의 암호문 c'을 만드는 것이 계산량적으로는 불가능하다는 것을 나타낸다. 예를 들어 엘가말(ElGamal)공개키 암호 시스템은 매우 유연하다. Alice는 평문 m을 $(\alpha^k, h^k m)$(여기서, $h(=\alpha^x)$는 Bob의 공개키, α는 주어진 유한군의 생성원이다.)로 암호화하여 Bob에게 전달한다고 하자. 만일 공격자 Charlie가 도청을 하여 이 암호문을 알아내어 $(\alpha^k, l \cdot h^k m)$을 생성한 후(Charlie는 단지 암호문의 두 번 째 원소에 l만을 곱함에 유의하자.) Bob에게 보내면 Bob은 $l \cdot h^k m/(\alpha^k)^x = l \cdot h^k m/h^k = l \cdot m$을 계산하여 원래의 평문 m과 연관된 다른 평문을 얻는다.

Bellare 등은 Crypto'98[4] 에서 이 계산량적 구별불가능성과 비유연성을 공개키 암호시스템의 안전성이 추구해야할 목표(goal)로, 또한 공격자가 어느 정도의 힘을 가질 수 있는

가를 나타내는 공격모델(attack model)을 결합하여 공개키 암호 시스템의 안전성에 관한 개념을 정리했다. 즉, 목표를 {IND(INDistinguishability), NM(Non-Malleability)}로, 공격 모델을 {CPA(Chosen Plaintext Attack), CCA1(non-adaptive Chosen Ciphertext Attack), CCA2(adaptive Chosen Ciphertext Attack)}로 나누어 이들의 6가지 조합을 정의했다. CPA는 공격자가 자신이 임의로 고른 평문을 암호화하여 이루어지는 공격을 말하고(사실, 공개키 암호 시스템에서는 암호화키가 공개되기 때문에 이 공격은 항상 가능하다.), CCA1와 CCA2에서는 공격자가 임의의 암호문을 선택한 후 이 암호문들에 대한 복호화를 요청하여 암호문들에 대한 평문을 얻을 수 있다고 가정한다. 그런데 CCA1에서는 공격자가 공격하고자하는 암호문과 독립적인 암호문만을 선택하여 복호화한다(즉, 공격하고자하는 암호문을 얻지 못한 경우이다.)고 가정하고, CCA2는 공격하고자 하는 암호문을 얻은 후, 선택암호문 공격을 하는 경우이다. CCA2는 매우 강력한 공격이긴 하지만, 최근의 공개키 암호 시스템은 이 CCA2를 공격 모델로 가정하여 분석되고 설계된다.

앞서 언급한 6가지 조합은 IND-CPA, IND-CCA1, IND-CCA2, NM-CPA, NM-CCA1, NM-CCA2 이며 본 논문에서는 'IND-공격 모델'에 초점을 맞추어 설명한다. 우선 'IND-공격모델'의 형식적인 정의는 다음과 같다.

정의 2 [IND-CPA, IND-CCA1, IND-CCA2] $\Pi=(K,E,D)$를 공개키 암호 시스템이라 하고, $A=(A_1,A_2)$를 공격자라고 하자. 그리고, 공격 모델$\in\{\mathrm{cpa},\mathrm{cca1},\mathrm{cca2}\}$과, $k\in N$에 대하여 $\mathrm{Adv}^{\mathrm{ind-공격\ 모델}}_{A,\Pi}(\mathrm{k})$을 다음과 같이 정의하자.

$$2 \cdot \Pr[(pk, sk) \leftarrow K(1^k)\ ;\ (x_0, x_1, s) \leftarrow A_1^{O_1}(pk)\ ;\ b \leftarrow_R \{0,1\}\ ;\ y \leftarrow E_{pk}(x_b)\ :\ A_2^{O_2}(x_0, x_1, s, y) = b] - 1$$

단, 공격 모델=cpa이면, $O_1(\cdot)=\lambda$, $O_2(\cdot)=\lambda$ 이고 (λ는 'null'을 나타낸다.)

공격 모델=cca1이면, $O_1(\cdot)=D_{sk}(\cdot)$, $O_2(\cdot)=\lambda$ 이며,

공격 모델=cca2이면, $O_1(\cdot)=D_{sk}(\cdot)$, $O_2(\cdot)=D_{sk}(\cdot)$ 이다.

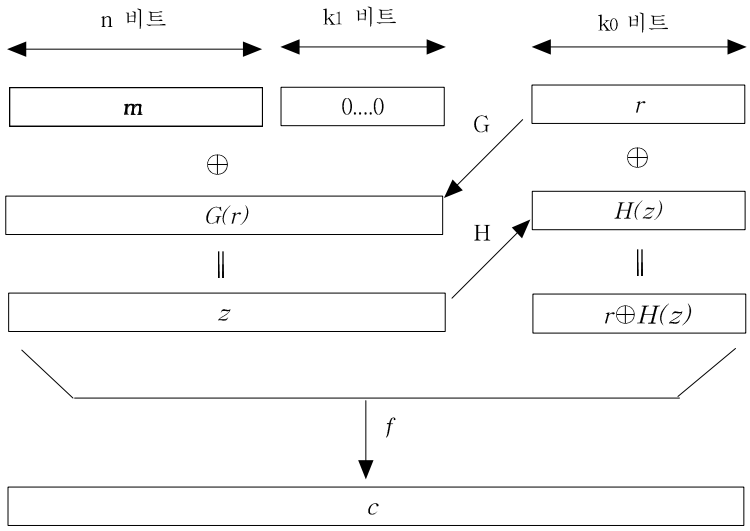
$\mathrm{Adv}_{A,\Pi}^{\mathrm{ind-공격\ 모델}}(\mathrm{k})$가 무시할 수 있다(negligible)면 공개키 암호 시스템 Π가 'IND-공격 모델'에 대하여 안전하다고 한다.

(참고) K는 키공간, E는 암호화 알고리즘, D는 복호화 알고리즘이며, pk는 공개키, sk는 비밀키이다. k는 키 파라미터이며, s는 단계(state) 정보, E_{pk}는 공개키 pk를 사용하는 암호화 함수, D_{sk}는 비밀키 sk를 사용하는 복호화 함수를 나타낸다.

위의 정의는 다음과 같이 풀이된다. 먼저, 키 파라미터 k에 의하여 공개키 비밀키 쌍이 생성($(pk, sk) \leftarrow K(1^k)$)되고, 공격자(A_1)가 추측 단계(guess-stage)에서 평문 x_0, x_1을 선택한다. 그리고 $0,1$중 하나가 랜덤하게 선택된다.(이것은 공격자가 선택하는 것이 아님에 유의하자.) 그 다음, 선택된 평문 x_b의 암호문이 계산된다. (이것 역시 공격자가 계산하는 것이 아니다.) 선정 단계(find-stage)에서 공격자(A_2)는 평문, x_0, x_1과 암호문 y을 입력으로 받아 b를 선정한다. 즉, 암호문이 x_0이 암호화된 것인지, x_1이 암호화된 것인지 맞추는 것이다. **정의 2**가 의미하는 바는 이 사건(event, 공격자가 $b=0$ 또는 $b=1$을 올바르게 선정하는 사건)이 일어날 확률이 무시할 수 있을 정도로 작을 때 주어진 공개키 암호시스템이 'IND-공격 모델'에 대하여 안전하다는 것이다.

한편 $O_i(\cdot)$는 공격자 A_i가 추측 단계와 선정 단계에서 가지는 부가적인 알고리즘을 말한다. 예를 들어, cpa에서는 추측 단계와 선정 단계에서 공격자에 제공되는 부가적인 알고리즘이 없고, cca2에서는 추측 단계와 선정 단계에서 공격자에게 복호화 알고리즘 D_{sk}가 제공된다. 한편, 랜덤 오라클 모델에서 IND-공격을 정의하는 것은 매우 쉽다. $O_i(\cdot)$ ($i \in \{1,2\}$)에 랜덤 오라클 $H_j(\cdot)$ ($j \in N$)를 첨가하는 것이다. 예를 들어 두 개의 랜덤 오라클을 가정한 CCA2에 안전한 공개키 암호 시스템인 경우 $O_1(\cdot) = \{D_{sk}(\cdot), H_1(\cdot), H_2(\cdot)\}$, $O_2(\cdot) = \{D_{sk}(\cdot), H_1(\cdot), H_2(\cdot)\}$이 된다.

실제로, 두 개의 랜덤 오라클을 가정한 CCA2에 안전한 증명 가능 공개키 암호 시스템으로 Bellare와 Rogaway의 OAEP[2]가 있다. 다음의 그림은 OAEP 기법을 나타낸 그림이다.

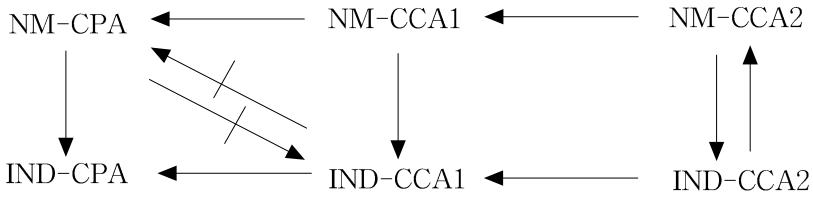


(그림 1)OAEP

위의 그림에 의하면, OAEP를 이용하여 평문 m을 암호화 하면 암호문(c)은 $f((m\|0^{k_1})\oplus G(r)\ \|\ r\oplus H(z))$ (단, $z=(m\|0^{k_1})\oplus G(r)$, f는 RSA 함수) 이 된다. 여기서, G와 H는 랜덤 오라클이다. OAEP가 IND-CPA를 만족하는 직관적인 이유는 다음과 같다. 만약 공격자가 $n+k_0$비트인 H의 원상 z를 모두 알아내지 못한다면, 공격자는 해쉬 값인 $H(z)$를 전혀 알 수 없다. 그런데 $H(z)$값을 알아내지 못한다면 $r\oplus H(z)$의 값을 알아낼 수 없고 따라서, $G(r)$도 알아낼 수 없으므로, $m\|0^{k_1}(=G(r)\oplus z)$를 알 수 없다. 그렇다면, 공격자가 z를 모두 알아냈다고 가정하자. 그러나 공격자가 r을 알아내지 못한다면, $G(r)$을 알아낼 수 없고, 따라서, $m\|0^{k_1}(=G(r)\oplus z)$을 알 수 없다.

한편, Bellare와 Rogaway는 IND-CCA2에 대한 안전성을 평문 인식성(Plaintext Awareness, PA)이라는 개념을 도입하여 증명하였다. 간단히 말해, 평문 인식성이란 평문을 알지 못하고는 그에 상응하는 암호문을 만들기가 극히 어렵다는 것이다. 이 평문 인식성은 랜덤 오라클 모델에서만 정의된다는 특징이 있다.

본 논문에서는 자세히 언급하지 않겠지만, 비유연성도 구별불가능성과 같이 'NM-공격 모델' 형태로 정의될 수 있다[4]. 'NM-공격 모델'과 'IND-공격 모델'은 밀접한 관련이 있다는 것이 [4]에 증명되어 있다. 특히, CCA2에 대해서는 'NM-공격 모델'과 'IND-공격 모델' 은 완전히 같다. 다음의 그림은 이들 간의 연관성을 보여주는 그림이다. 그림에서 $A \rightarrow B$는 'A가 성립하면 B가 성립한다'는 것을 나타내고, $A \nrightarrow B$는 'A가 성립한다는 것이 B가 성립한다는 것을 의미하지 않는다'는 것을 나타낸다.



(그림 2)공개키 암호 안전성의 각 개념들간의 관계

최근의 공개키 암호 시스템의 안전성에 대한 연구는 대부분이 암호문 선택공격(CCA2)에 초점이 맞추어지고 있다. 랜덤 오라클을 사용하지 않는 실용적인 공개키 암호 시스템은 Cramer와 Shoup에 의해 Crypto'98에서 제안 되었는데 매우 높은 이론적 가치를 가지고 있다[10]. 하지만 여기서 '실용적'이라는 말은 다분히 상대적인 개념이다. 실제 응용에 곧 적용하기가 수월한 기법은 앞서 언급한 OAEP와 PKC'99에서 Fujisaki와 Okamoto가 제안한 F-O 기법(F-O 기법은 임의의 CPA에 안전한 공개키 암호 시스템을 CCA2에 안전한 시스템으로 바꾸는 일반적인 방법으로 그 응용 범위가 매우 넓다.)[13]등이 있다.

4. 오라클 해쉬

4.1 동기

랜덤 오라클 패러다임은 안전성이 증명 가능한 효율적인 암호 프로토콜을 설계하는 방법을 제공하기는 하지만, 랜덤 오라클의 가정이 너무 강력하다는 지적을 받고 있다. 실제로는 랜덤 오라클과 같은 이상적인 해쉬함수가 존재하지 않기 때문에 MD5나 SHA-1과 같은 암호학적 해쉬함수로 대체되어야 한다. 따라서, 증명에서 포함되어 있던 수학적 의미는 많이 사라지게 되는 것이다. 이 문제에 대해서는 Goldreich 등이 연구한 바 있다.[9] Crypto'97에서 Canetii는 랜덤 오라클에 비교될 수 있는 해쉬 기법인 오라클 해쉬(Oracle hash, 다른 말로 Perfectly probabilistic one-way 함수[8]라고 불리기도 한다.)를 제안하였다[7]. 이전의 해쉬함수는 모두 결정적인(deterministic)함수이기 때문에 랜덤 오라클과는 달리 해쉬값은 입력값에 대한 정보를 상당량 유출할 수 밖에 없었다. 하지만 이 오라클 해쉬 기법은 이전의 해쉬 함수가 가지고 있는 충돌회피성, 일방향성 등의 성질을 유지하면서 입력값의 부분 정보를 숨길 수 있다는 특징을 가지고 있다. 이를 위해서 Canetti는 해쉬 함수의 출력값을 랜덤화(randomization)하는 방법을 생각하였다. 즉, 해쉬 함수에 랜덤화 파라미터를 삽입하였는데, 결과적으로 오라클 해쉬값은 하나의 입력에 대하여 여러 가지가 될 수 있다. 그리고 간단히 말해, 공격자가 오라클 해쉬값들을 보고, 이 값들이 어떤 입력값의 해쉬 값들인지 구별할 수 없을 때 오라클 해쉬가 안전하다고 정의한다. 이에 대해서는 다음 절에서 자세히 다루도록 한다.

4.2 오라클 해쉬의 정의 및 성질

오라클 해쉬는 완전성(Completeness), 정확성(Correctness)과 오라클 은닉성(Oracle Secrecy)의 세 가지 성질로 정의된다.

4.1에서 언급되었듯이 주어진 입력값 x에 대한 오라클 해쉬값(c)이 여러개가 될 수 있기 때문에 일반적인 해쉬 함수와는 달리 해쉬값을 확인(verify)할 수 있는 성질을 잃게 된다. 따라서, 오라클 해쉬 기법에서는 별도의 해쉬값 다항식 시간 확인 알고리즘 V를 정의한다. 완전성은 이 "확인 알고리즘 V가 (무시할 수 없을 정도의 큰 확률로) $(x, c=(\hat{H}(x)))$ (단, $\hat{H}$는 오라클 해쉬)의 쌍을 받아들인다"는 것을 의미한다. 형식적인 정의는 다음과 같다.

정의 3 [오라클 해쉬의 완전성] 충분히 큰 k와 모든 입력 x와 랜덤수 $r\in R_k$에 대하여 $\Pr[V(x, \hat{H}(x, r))\neq 1]$이 무시할 수 있을 정도(negligible)로 작을 때 오라클 해쉬 $\hat{H}$가 완전성을 만족한다고 한다.

정확성은 일반적인 암호학적 해쉬함수의 충돌회피성과 유사한 개념인데, "x에 오라클 해쉬 $\hat{H}$를 적용하지 않고 나온 값 c를 공격자 A가 V로 하여금 받아들이게 할 확률이 무시할 수 있을 정도로 작다"는 것을 의미한다. 형식적인 정의는 다음과 같다.

정의 4 [오라클 해쉬의 정확성] 임의의 PPT(Probabilistic Polynomial Time) 공격자 A가 입력 k에 대하여 $V(x, c)=V(y, c)=1$인 c, x, y ($x\neq y$)를 출력할 확률이 무시할 수 있을 정도로 작을 때 오라클 해쉬 $\hat{H}$가 정확성을 만족한다고 한다.

오라클 은닉성이란 "$c=\hat{H}(x)$를 가지고 입력 x에 대한 정보(부분 정보라도)를 알아내는 것은 x의 정의역(domain)에서 $V(z, c)$가 받아들이는 z를 전수 조사하는 방법 외에는 없다"라는 것을 의미한다. 이는 다음과 같이 좀 더 형식화 될 수 있다. I_x를 $z=x$인 질문(query)에 대하여 1을 출력하고, $z\neq x$인 질문에 대해서는 0을 출력하는 오라클이라 하자. 따라서 위의 조건은 "주어진 $\hat{H}(x)$에 대하여 x에 관한 정보를 찾는 것은 오라클 I_x에 접근하여 정보를 알아내는 방법 밖에 없다"로 재해석 될 수 있다. 다음은 오라클 은닉성의 형식적인 정의이다.

정의 5 [오라클 은닉성] 모든 다항식 시간 공격자 C과 다항식 $p(\cdot)$, 모든 분포 앙상블(distribution ensemble) $\{X_k\}$에 대해서

$$\Pr[C(\hat{H}(x, r))=P(x)]-\Pr[C^{I_x}()=P(x)]<\frac{1}{p(k)}$$

를 만족하는 다항식 시간 공격자 C가 존재하면 오라클 은닉성을 만족한다고 한다. (단, r은 랜덤수이며, $r\in R_k$이며, $P(\cdot)$은 다항식 시간 predicate이며, $C^{I_x}(\cdot)$은 공격자 C가 오라클 I_x를 이용하여 '질문-답'의 행위를 수행하며 연산을 할 수 있다는 것을 말한다.)

참고로, 오라클 은닉성은 **정의 5**와 다른 여러 가지 형태로 정의될 수 있다.[7, 8] 이들 서로 다른 정의들의 등가성(equivalence)과 의미에 대해서는 [7,8]에 자세히 언급되어 있다.

4.3 오라클 해쉬의 구성 및 응용

[7]에서 오라클 해쉬를 구성할 수 있는 몇 가지 예가 제시되었다. 첫 번째의 예는 결정 Diffie-Hellman 문제(Decisional Diffie-Hellman, DDH)를 이용한 것인데 $\hat{H}(x, r) = (r, r^x)$로 구성된다. 여기서 r은 랜덤수이고 x는 입력값이다. 다른 하나는 암호학적 해쉬함수를 이용한 것인데, $(r, h(r, h(x)))$이다. 한편, [8]에서는 범용 일방향 해쉬함수(Universal one-way hash function)를 사용하여 오라클 해쉬를 구성할 수 있음을 보였다.

오라클 해쉬는 그 자체가 암호 요소로서 의미가 있지만, 응용 범위를 넓히는 것도 중요하다. Canetti는 Bellare와 Rogaway가 [1]에서 제안한 랜덤 오라클 모델에서 CPA에 안전한 공개키 암호 시스템의 랜덤 오라클을 오라클 해쉬로 대체하여 안전성을 증명하였다[7].

한편, 오라클 해쉬는 디지털 서명에서 서명되는 메시지의 부분 정보를 효과적으로 숨길 수 있는 암호 요소로서 이용될 수 있을 것이다. 또한, 안전한 키 분배(key distribution)에 있어서도 이용이 가능한데, 예를 들어, 키 분배 후 분배자는 각 실체들이 분배한 키와 같은 키를 소유하고 있는지를 키의 오라클 해쉬 값을 공개하여 각 실체들이 그것을 확인함으로서 자신의 키가 올바른 키인지 확인할 수 있다.

하지만 오라클 해쉬는 랜덤 오라클에 비하여 제약성이 많다. 첫째로, 랜덤 오라클과 같이 결정적이면서 입력값의 정보를 모두 숨기는 성질을 가지고 있는 것이 아니라, 출력값이 여러 가지로 나올 수 있기 때문에 분석이 복잡해 질 수 있다. 또한 랜덤 오라클은 공격자가 독립적인 질문을 하여 그에 응답을 얻는 모형으로 분석을 할 수 있지만, 오라클 해쉬에서는 공격자가 질문을 하지 않았다고 해서 나오는 출력값을 예상하지 못한다는 보장이 없으므로 공격자의 행동을 '질문-응답'의 형태로 나타내는 것이 매우 복잡해 진다.

5. 오라클 해쉬의 응용에 관련된 문제

오라클 해쉬의 응용과 관련하여 많은 문제들이 제기될 수 있다. 참고로 지금까지는 오라클 해쉬를 사용하는 공개키 암호시스템이 거의 제안된 바가 없다.

첫째 문제는 오라클 해쉬의 가정 하(랜덤 오라클이 아닌)에서 CCA2에 증명 가능 안전성이 보장되는 공개키 암호 시스템을 구성할 수 있는가하는 문제이다. Canetti는 두 개의 랜덤 오라클을 이용하고 있는 Bellare와 Rogaway가 제안한 CCA2에 안전한 공개키 암호 시스템의 랜덤 오라클들을 오라클 해쉬로 전환하려 했지만 성공하지 못했다[7]. 이는 두 개의 오라클 해쉬를 다루어야하는 데 오는 분석의 복잡함 때문이라 여겨진다. 그러나 최근 제안된 F-O 기법은 하나의 랜덤오라클을 사용하며 CCA2에 대한 안전성을 제공하기 때문에 이 시도의 가장 좋은 대상이 될 수 있을 것이다. (F-O 기법은 $E_{pk}(m) = e_{pk}(m||r, h(m||r))$으로 간단히 표현된다. 즉 평문 m을 암호화 하기 위

해 랜덤수 r을 패딩한 후 이 값을 랜덤 오라클에 적용한 후 CPA에 안전한 암호화 함수 e_{pk}로 암호화한다. 이 랜덤오라클 h를 오라클 해쉬로 대치하는 것이 제안하는 문제이다.) 하지만 F-O 기법은 랜덤오라클 모델에서만 정의되는 평문 인식성을 증명에 강하게 이용하고 있으므로 문제가 발생할 수 있다. 그러나 CCA2의 정의만을 사용하여 증명할 수 있는 방법을 찾을 가능성은 존재한다.

둘째는 오라클 해쉬의 가정하에서 디지털 서명의 안전성 분석에 핵심이되는 Forking Lemma(즉, 주어진 디지털 서명에 대한 PPT 공격자가 효과적으로 서명 $(m, \sigma_1, h, \sigma_2)$ (h는 m과 σ_1의 랜덤 오라클 값, σ_2는 σ_1과 h에 종속적인 값)을 생성해 냈다면, $h(\cdot)$와 다른 랜덤 오라클 $h'(\cdot)$을 되풀이(replay)하여 $(m, \sigma_1, h', \sigma_2)$을 효과적으로 생성해 낼 수 있다는 정리[15, 16].)를 증명하는 것이다. 이 문제를 해결하기 위해서는 오라클 되풀이 공격이라는 Forking Lemma의 기본 개념을 오라클 해쉬의 모델 하에서 효과적으로 정의하는 것이 선행되어야 할 것이다.

셋째는 오라클 해쉬를 이용한 키 분배 메커니즘에 관한 것이다. 4.3에서도 언급되었듯이 각 실체들이 자신의 키를 상대방의 키의 오라클 해쉬값과 비교하여 자신의 키를 확인할 수 있는 효과적인 방법을 찾는 것이다.

6. 결론

본 논문에서는 최근 암호학 연구에서 새로운 이슈로 부각되고 있는 증명 가능 안전성에 대하여 논하였다. 그리고 안전성 증명의 매우 효율적인 도구가 될 수 있는 랜덤 오라클과 랜덤 오라클과 비교하여 오라클 해쉬에 대하여 살펴보았다. 또한 랜덤 오라클과 오라클 해쉬의 응용 분야에 대해 고찰하고 그에 관계되는 문제를 제기하였다.

증명 가능 안전성에 관한 연구는 계산 복잡도 이론 등을 취급하는 전산학, 정수론, 확률론 등의 수학적 이론을 바탕으로 한 여러 가지 암호 요소에 대한 이해를 전제하고 있기 때문에 연구에 어려움이 있지만, 국내에서도 지속적인 연구를 진행해야할 분야이다.

참고문헌

[1] M. Bellare , P. Rogaway, "Random Oracles are Practical: A Paradigm for Designing Efficient Protocols", Proc. First Annual Conference on Computer and Communications Security, ACM, pp.62-73, 1993.

[2] M. Bellare , P. Rogaway, "Optimal Asymmetric Encryption - How to Encrypt with RSA", Advances in Cryptology - Eurocrypt '94, LNCS, Vol. 950, Springer-Verlag, pp. 92-111, 1995.

[3] M. Bellare, "Practice-Oriented Provable-Security", Proc. of First International Workshop on Information Security(ISW 97), LNCS, Vol. 1396, Springer-Verlag, pp221-231, 1998.

[4] M. Bellare, A. Desai, D. Pointcheval and P.Rogaway, "Relations Among Notions of

Security for Public-Key Encryption Schemes", Advances in Cryptology - Crypto '98, LNCS, Vol. 1462, Springer-Verlag, pp.26-45, 1998.

[5] S. Blake-Wilson, D. Johnson and A. Menezes, "Key agreement protocols and their security analysis", Proc. of the 6th IMA International Conference on Cryptography and Coding, LNCS, Vol. 1355 , pp.30-45, Springer-Verlag, 1997.

[6] D. Bleichenbacher "Chosen Ciphertext Attacks against Protocols Based on RSA Encryption Standard PKCS #1" , Advances in Cryptology -Crypto '98, LNCS, Vol. 1462, Springer-Verlag, pp.1-12, 1998.

[7] R. Canetti, "Towards Realizing Random Oracles: Hash Functions That Hide All Partial Information", Advances in Cryptology - Crypto '97, LNCS, Vol. 1294, Springer-Verlag, pp. 455-469, 1997.

[8] R. Canetti, D. Micciancio, O. Reingold, "Perfectly One-Way Probabilistic Hash Functions", Proc. of the 30th ACM, STOC '98, 1998.

[9] R Canetti, O Goldreich and S. Halevi, "The Random Oracle Methodology, Revisited," Proc. of the 30th ACM STOC '98 ,1998.

[10] R. Cramer, V. Shoup, "A Practical Public Key Cryptosystem Provably Secure against Adaptive Chosen Ciphertext Attack" ,Advances in Cryptology-Crypto '98, LNCS, Vol. 1462, Springer-Verlag, pp.13-25, 1998.

[11] D. Dolev, C. Dwork and M. Naor, "Non-Malleable Cryptography", Manuscript, 1998, (preliminary version appeared in STOC '91).

[12] Y. Frankel, M. Yung, "Cryptanalysis of the Immunized LL Public Key Systems",Advances in Cryptology - Crypto '95, LNCS, Vol. 963, Springer-Verlag, pp. 288-296, 1995.

[13] E. Fujisaki, T. Okamoto, "How to Enhance the Security of Public-Key Encryption at Minimum Cost", Proc. of PKC '99, LNCS, Vol. 1560, Springer-Verlag, pp.53-68, 1999.

[14] M. Blum, S. Goldwasser, "An Efficient Probabilistic Public-Key Encryption Scheme Which Hides All Partial Information,", Advances in Cryptology - Crypto '84, LNCS, Vol. 196, Springer-Verlag, pp. 289-299, 1985.

[15] D. Pointcheval J. Stern, "Security Proofs for Signature Schemes", Advances in Cryptology - Eurocrypt '96, LNCS, Vol. 1070, Springer-Verlag, pp.387-398, 1996.

[16] D. Pointcheval J. Stern, "Security Arguments for Digital Signatures and Blind Signatures", to appear in Journal of Cryptology, Springer-Verlag, 1999.

[17] P. Rogaway, Lecture Notes on Modern Cryptography, UC Davis Computer Science Course ECS 227, 1999, Available at http://www.cs.ucdavis.edu/~rogaway/teaching.html

[18] V. Shoup, "Why chosen ciphertext security matters", IBM Research Report RZ 3076, November, 1998.

[19] Y. Zheng, J. Seberry, "Immunizing Public Key Cryptosystems Against Chosen Ciphertext Attack", IEEE Journal of Selected Areas in Communications, pp. 715-724, 1993.