RSA-OAEP IND-CCA2 증명

김동현(wlswudpdlf31@kookmin.ac.kr)

March 25, 2025

Contents

1 논문정보

- 제목: RSA-OAEP is Secure under the RSA Assumption
- 저자: Eiichiro Fujisaki1, Tatsuaki Okamoto, David Pointcheval, and Jacques Stern
- 년도: 2001년
- 초록: 최근 Victor Shoup은 적응적 선택 암호문 공격에 대한 OAEP의 보안성에 관한 널리 받아들여진 결과에 틈이 있음을 지적하였다. 더욱이, 그는 기본 트랩도어 치환의 단방향성만으로는 OAEP의 보안성을 증명할 수 없을 것으로 예상된다는 점을 보였다. 본 논문은 OAEP의 보안성에 대한 또 다른 결과를 제시한다. 즉, 본 논문에서는 무작위 오라클 모델에서, 기본 치환의 부분 영역 단방향성(partial-domain one-wayness) 하에서, OAEP가 적응적 선택 암호문 공격에 대해 의미론적 보안성을 제공함을 증명한다. 따라서, 이는 형식적으로 더 강한 가정을 사용한다. 그럼에도 불구하고, RSA 함수의 부분 영역 단방향성이 (전체 영역) 단방향성과 동치이므로, RSA-OAEP의 보안성은 단순한 RSA 가정만으로도 증명될 수 있음을 알 수 있다. 다만, 그 축소(reduction)는 타이트하지 않다.

2 보안 개념

2.1 OW trapdoor permutation

트랩도어 치환 체계(Trapdoor permutation scheme) $\Psi = (\mathcal{K}, \mathcal{F}, \mathcal{I})$를 다음과 같이 정의한다.

- $\mathcal{K}(1^\lambda)$: 확률론적 키 생성 알고리즘으로, 1^λ를 입력 받아 $(\mathsf{pk}, \mathsf{sk})$를 생성한다.
- $\mathcal{F}_{\mathsf{pk}}(x)$: 결정론적 알고리즘으로, pk와 $x \in \{0,1\}^\lambda$를 입력 받아 $y \in \{0,1\}^\lambda$를 출력한다.
- $\mathcal{I}_{\mathsf{sk}}(y)$: 결정론적 알고리즘으로, sk와 $y \in \{0,1\}^\lambda$를 입력 받아 $x \in \{0,1\}^\lambda$를 출력한다. $\mathcal{K}(1^\lambda)$로 생성한 모든 $(\mathsf{pk}, \mathsf{sk})$와 모든 $x \in \{0,1\}^\lambda$에 대해, $\mathcal{I}_{\mathsf{sk}}(\mathcal{F}_{\mathsf{pk}}(x)) = x$를 만족한다.

동작시간(Running time) τ를 가지는 공격자 $\mathcal{A}$와 트랩도어 치환 체계 Ψ에 대한 일방향성(One-wayness) 실험 $\mathsf{Exp}^{\mathsf{OW}}_{\Psi,\lambda}(\mathcal{A};\tau)$을 다음과 같이 정의한다.

Challenger $\mathcal{C}$	$\xLeftrightarrow{\mathsf{Exp}^{\mathsf{OW}}_{\Psi,\lambda}}$	Adversary $\mathcal{A}$
$(\mathsf{pk}, \mathsf{sk}) \xleftarrow{\$} \mathcal{K}(1^\lambda)$		
$x^* \xleftarrow{\$} \{0,1\}^\lambda$		
$y^* \leftarrow \mathcal{F}_{\mathsf{pk}}(x^*)$	$\xrightarrow{1^\lambda, \mathsf{pk}, y^*}$	
Return $[x' \overset{?}{=} x^*]$	$\xleftarrow{x'}$	$\mathcal{A}$ chooses $x' \in \mathcal{X}$

$\mathcal{A}$의 능력치 $\mathsf{Adv}^{\mathsf{OW}}_{\mathcal{A};\Psi}(\lambda, \tau)$를 다음과 같이 정의한다.

$$\mathsf{Adv}^{\mathsf{OW}}_{\Psi,\lambda}(\mathcal{A};\tau) := \Pr[\mathsf{Exp}^{\mathsf{OW}}_{\Psi,\lambda}(\mathcal{A};\tau) = 1].$$

2.2 Partial-domain OW trapdoor permutation

트랩도어 치환 $\Psi = (\mathcal{K}, \mathcal{F}, \mathcal{I})$에서, $\mathcal{F}_{\mathsf{pk}}(x) : \{0,1\}^\lambda \to \{0,1\}^\lambda$를 다음과 같이 표현한다.

$$\mathcal{F}_{\mathsf{pk}} : \{0,1\}^{n+\lambda_1} \times \{0,1\}^{\lambda_0} \to \{0,1\}^{n+\lambda_1} \times \{0,1\}^{\lambda_0}.$$

이때 $\lambda = n + \lambda_0 + \lambda_1$이다.

메모 1. 예를 들어, $x = s \parallel t$라고 할 때, $y \leftarrow \mathcal{F}_{\mathsf{pk}}(x)$ 대신 $y \leftarrow \mathcal{F}_{\mathsf{pk}}(s \parallel t)$로 표현할 수 있다.

동작시간 τ를 가지는 공격자 $\mathcal{A}$와 트랩도어 함수 체계 Ψ에 대한 부분 일방향성(Partial-domain one-wayness) 실험 $\mathsf{Exp}^{\mathsf{PD\text{-}OW}}_{\Psi,\lambda}(\mathcal{A};\tau)$을 다음과 같이 정의한다.

$$
\begin{array}{ccc}
\underline{\text{Challenger } \mathcal{C}} & \overset{\mathsf{Exp}^{\mathsf{PD\text{-}OW}}_{\Psi,\lambda}}{\Longleftrightarrow} & \underline{\text{Adversary } \mathcal{A}} \\
(\mathsf{pk},\mathsf{sk}) \overset{\$}{\leftarrow} \mathcal{K}(1^\lambda) & & \\
(s^*, t^*) \overset{\$}{\leftarrow} \{0,1\}^{n+\lambda_1} \times \{0,1\}^{\lambda_0} & & \\
y^* \leftarrow \mathcal{F}_{\mathsf{pk}}(s^*, t^*) & \xrightarrow{1^\lambda, \mathsf{pk}, y^*} & \\
\text{Return } [s' \overset{?}{=} s^*] & \overset{s'}{\longleftarrow} & \mathcal{A} \text{ chooses } s' \in \{0,1\}^{n+\lambda_1}
\end{array}
$$

공격자 $\mathcal{A}$의 능력치 $\mathsf{Adv}^{\mathsf{PD\text{-}OW}}_{\Psi,\lambda}(\mathcal{A};\tau)$를 다음과 같이 정의한다.

$$\mathsf{Adv}^{\mathsf{PD\text{-}OW}}_{\Psi,\lambda}(\mathcal{A};\tau) := \Pr[\mathsf{Exp}^{\mathsf{PD\text{-}OW}}_{\Psi,\lambda}(\mathcal{A};\tau) = 1].$$

2.3 Set partial-domain OW trapdoor permutation

동작시간 τ를 가지고 l개의 원소를 출력하는 공격자 $\mathcal{A}$와 트랩도어 함수 체계 Ψ에 대한 집합 부분 일방향성(Set partial-domain one-wayness) 실험 $\mathsf{Exp}^{\mathsf{S\text{-}PD\text{-}OW}}_{\Psi,\lambda}(\mathcal{A};\tau,l)$을 다음과 같이 정의한다.

$$
\begin{array}{ccc}
\underline{\text{Challenger } \mathcal{C}} & \overset{\mathsf{Exp}^{\mathsf{S\text{-}PD\text{-}OW}}_{\Psi,\lambda}}{\Longleftrightarrow} & \underline{\text{Adversary } \mathcal{A}} \\
(\mathsf{pk},\mathsf{sk}) \overset{\$}{\leftarrow} \mathcal{K}(1^\lambda) & & \\
(s^*, t^*) \overset{\$}{\leftarrow} \{0,1\}^{n+\lambda_1} \times \{0,1\}^{\lambda_0} & & \\
y^* \leftarrow \mathcal{F}_{\mathsf{pk}}(s^*, t^*) & \xrightarrow{1^\lambda, \mathsf{pk}, y^*} & \\
\text{Return } [s^* \overset{?}{\in} S'] & \overset{S'}{\longleftarrow} & \mathcal{A} \text{ chooses } S' = \{s_i \in \{0,1\}^{n+\lambda_1} : \\
 & & 0 \le i \le l-1\}
\end{array}
$$

공격자 $\mathcal{A}$의 능력치 $\mathsf{Adv}^{\mathsf{S\text{-}PD\text{-}OW}}_{\Psi,\lambda}(\mathcal{A};\tau,l)$를 다음과 같이 정의한다.

$$\mathsf{Adv}^{\mathsf{S\text{-}PD\text{-}OW}}_{\Psi,\lambda}(\mathcal{A};\tau,l) := \Pr[\mathsf{Exp}^{\mathsf{S\text{-}PD\text{-}OW}}_{\Psi,\lambda}(\mathcal{A};\tau,l) = 1].$$

2.4 IND security against CCA2

공개키 암호 체계(Public-key encryption scheme) $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$를 다음과 같이 정의한다.

- $\mathcal{K}(1^\lambda)$: 확률론적 키 생성 알고리즘으로, 1^λ를 입력 받아 $(\mathsf{pk}, \mathsf{sk})$를 생성한다.
- $\mathcal{E}_{\mathsf{pk}}(m)$: 암호화 알고리즘으로, pk와 $m \in \mathcal{M}$를 입력 받아 $c \in \mathcal{C}$를 출력한다. 확률론적 알고리즘으로, $r \overset{\$}{\leftarrow} \Omega$를 추가로 입력 받아 $\mathcal{E}_{\mathsf{pk}}(m;r)$으로 표현할 수도 있다.
- $\mathcal{D}_{\mathsf{sk}}(c)$: 결정론적 복호화 알고리즘으로, sk와 $c \in \mathcal{C}$를 입력 받아 $m \in \mathcal{M}$를 출력한다.

동작시간 τ를 가지고 복호화 오라클에 q회 질의하는 공격자 $\mathcal{A}$와 공개키 암호 체계 Π에 대해, 선택 암호문 공격(Adaptive chosen ciphertext attack, 이하 CCA2)에 대한 구별불가능성(Indistinguishability) 실험 $\mathsf{Exp}^{\mathsf{IND\text{-}CCA2}}_{\Pi,\lambda}(\mathcal{A};\tau,q)$을 다음과 같이 정의한다.

Challenger $\mathcal{C}$	$\xLeftrightarrow{\mathsf{Exp}^{\mathsf{IND\text{-}CCA2}}_{\Pi,\lambda}}$	Adversary $\mathcal{A}$
$(\mathsf{pk},\mathsf{sk}) \leftarrow \mathcal{K}(1^\lambda)$	$\xrightarrow{1^\lambda,\mathsf{pk}}$	
	$\mathcal{O}^{\mathcal{D}_{\mathsf{sk}}}$ Query Phase 1	
	$\xleftarrow{m_0^*,m_1^*}$	$\mathcal{A}$ chooses $m_0^*, m_1^* \in \mathcal{M}$ such that $\lvert m_0^*\rvert = \lvert m_1^*\rvert$ and $m_0^* \neq m_1^*$
$b \xleftarrow{\$} \{0,1\}, r^* \xleftarrow{\$} \Omega$ $c^* \leftarrow \mathcal{E}_{\mathsf{pk}}(m_b^*; r^*)$	$\xrightarrow{c^*}$	
	$\mathcal{O}^{\mathcal{D}_{\mathsf{sk}}}$ Query Phase 2	
Return $[b' \stackrel{?}{=} b]$	$\xleftarrow{b'}$	$\mathcal{A}$ chooses $b' \in \{0,1\}$

공격자 $\mathcal{A}$의 능력치 $\mathsf{Adv}^{\mathsf{IND\text{-}CCA2}}_{\Pi,\lambda}(\mathcal{A};\tau,q)$를 다음과 같이 정의한다.

$$\mathsf{Adv}^{\mathsf{IND\text{-}CCA2}}_{\Pi,\lambda}(\mathcal{A};\tau,q) = 2\cdot \Pr[\mathsf{Exp}^{\mathsf{IND\text{-}CCA2}}_{\Pi,\lambda}(\mathcal{A};\tau,q) = 1] - 1.$$

2.5 IND security against CCA2 in ROM

동작시간 τ를 가지고 복호화 오라클에 $q_{\mathcal{D}}$회, 랜덤 오라클에 q_H회 질의하는 공격자 $\mathcal{A}$와 공개키 암호 체계 Π에 대해, 랜덤 오라클 모델(Random oracle model)에서의 CCA2에 대한 구별불가능성 실험 $\mathsf{Exp}^{\mathsf{IND\text{-}CCA2\text{-}ROM}}_{\Pi,\lambda}(\mathcal{A};\tau,q_{\mathcal{D}},q_H)$을 다음과 같이 정의한다.

Challenger $\mathcal{C}$	$\xLeftrightarrow{\mathsf{Exp}^{\mathsf{IND\text{-}CCA2\text{-}ROM}}_{\Pi,\lambda}}$	Adversary $\mathcal{A}$
$\mathcal{O}_H \xleftarrow{\$} \Phi$ $(\mathsf{pk},\mathsf{sk}) \leftarrow \mathcal{K}(1^\lambda)$	$\xrightarrow{1^\lambda,\mathsf{pk}}$	
	$\mathcal{O}^H, \mathcal{O}^{\mathcal{D}_{\mathsf{sk}}}$ Query Phase 1	
	$\xleftarrow{m_0^*,m_1^*}$	$\mathcal{A}$ chooses $m_0^*, m_1^* \in \mathcal{M}$ such that $\lvert m_0^*\rvert = \lvert m_1^*\rvert$ and $m_0^* \neq m_1^*$
$b \xleftarrow{\$} \{0,1\}, r^* \xleftarrow{\$} \Omega$ $c^* \leftarrow \mathcal{E}_{\mathsf{pk}}(m_b^*; r^*)$	$\xrightarrow{c^*}$	
	$\mathcal{O}^H, \mathcal{O}^{\mathcal{D}_{\mathsf{sk}}}$ Query Phase 2	
Return $[b' \stackrel{?}{=} b]$	$\xleftarrow{b'}$	$\mathcal{A}$ chooses $b' \in \{0,1\}$

공격자 $\mathcal{A}$의 능력치 $\mathsf{Adv}_{\Pi,\lambda}^{\mathsf{IND\text{-}CCA2\text{-}ROM}}(\mathcal{A};\tau,q_{\mathcal{D}},q_H)$를 다음과 같이 정의한다.

$$\mathsf{Adv}_{\Pi,\lambda}^{\mathsf{IND\text{-}CCA2\text{-}ROM}}(\mathcal{A};\tau,q_{\mathcal{D}},q_H) = 2\cdot\Pr[\mathsf{Exp}_{\Pi,\lambda}^{\mathsf{IND\text{-}CCA2\text{-}ROM}}(\mathcal{A};\tau,q_{\mathcal{D}},q_H)=1]-1.$$

3 RSA-OAEP

다음과 같은 트랩도어 치환 $\mathcal{F}$를 고려한다.

$$\mathcal{F}_{\mathsf{pk}} : \{0,1\}^{n+\lambda_1} \times \{0,1\}^{\lambda_0} \to \{0,1\}^{n+\lambda_1} \times \{0,1\}^{\lambda_0}.$$

그리고 두 해시 함수 H, G를 다음과 같이 준비한다.

$$H : \{0,1\}^{\lambda_0} \to \{0,1\}^{\lambda-\lambda_0} \quad G : \{0,1\}^{\lambda-\lambda_0} \to \{0,1\}^{\lambda_0}.$$

트랩토어 치환 체계 $\Psi = (\mathcal{K}, \mathcal{F}, \mathcal{I})$를 포함하는 OAEP 변환 $(\mathcal{K}, \mathcal{E}, \mathcal{D})$는 다음과 같이 동작한다.

- $\mathcal{K}(1^\lambda)$: $(\mathsf{pk}, \mathsf{sk})$를 생성한다. pk는 이후 트랩도어 치환 $\mathcal{F}$에서 사용하며, sk는 $\mathcal{I}$에서 사용한다.
- $\mathcal{E}_{\mathsf{pk}}(m; r)$: $m \in \{0,1\}^n$과 $r \xleftarrow{\$} \{0,1\}^{\lambda_0}$가 주어졌을 때, s, t를 다음과 같이 계산한다.

  $$s = (m \parallel 0^{\lambda_1}) \oplus G(r), \quad t = r \oplus H(s).$$

  s, t를 계산하는 과정을 도식화하면 그림 1와 같다. 이후 암호문 $c = \mathcal{F}_{\mathsf{pk}}(s, t)$를 출력한다.
- $\mathcal{D}_{\mathsf{sk}}(c)$: $(s, t) = \mathcal{I}_{\mathsf{sk}}(c)$을 계산한 후, r, M을 다음과 같이 계산한다.

  $$r = t \oplus H(s) \quad M = s \oplus G(r).$$

  만약 $[M]_{\lambda_1} = 0^{\lambda_1}$이면 $[M]^n$을 출력하고, 아니라면 “Reject”를 출력한다.

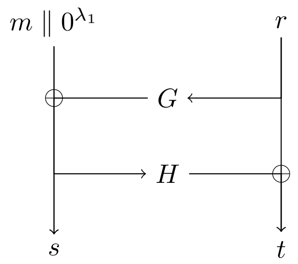

Figure 1: $\mathcal{E}_{\mathsf{pk}}(m; r)$에서 s, t를 계산하는 과정

4 증명

보조정리 1. *공격자 $\mathcal{A}$를 OAEP 변환 $(\mathcal{K}, \mathcal{E}, \mathcal{D})$에 대해 동작시간 τ를 가지고, 복호화 오라클 $\mathcal{O}^{\mathcal{D}}$와 랜덤 오라클 $\mathcal{O}^H, \mathcal{O}^G$에 각각 $q_{\mathcal{D}}, q_H, q_G$회 질의하는 IND-CCA2 공격자라 하자. 이때, 다음을 만족하는 S-PD-OW 공격자 $\mathcal{B}$가 존재한다.*

$$\mathsf{Adv}^{\mathsf{S\text{-}PD\text{-}OW}}_{\Psi,\lambda}(\mathcal{B}; \tau', q_H) \geq \frac{\mathsf{Adv}^{\mathsf{IND\text{-}CCA2}}_{\Pi,\lambda}(\mathcal{A}; \tau, q_{\mathcal{D}}, q_H, q_G)}{2} - \frac{2q_{\mathcal{D}}q_G + q_{\mathcal{D}} + q_G}{2^{\lambda_0}} - \frac{2q_{\mathcal{D}}}{2^{\lambda_1}}.$$

여기서, $\tau' \leq \tau \cdot q_H \cdot q_G \cdot (T_{\mathcal{F}} + O(1))$이고, $T_{\mathcal{F}}$는 트랩도어 치환 $\mathcal{F}$의 시간 복잡도를 의미한다.

4.1 증명: Reduction and simulation

Challenger $\mathcal{C}$	$\xleftrightarrow{\mathsf{Exp}^{\mathsf{S\text{-}PD\text{-}OW}}_{\Psi,\lambda}}$	Adversary $\mathcal{B}$	$\xleftrightarrow{\mathsf{Exp}^{\mathsf{IND\text{-}CCA2\text{-}ROM}}_{\Pi,\lambda}}$	Adversary $\mathcal{A}$
$(\mathsf{pk},\mathsf{sk}) \leftarrow \mathcal{K}(1^\lambda)$				
$(s^*,t^*) \xleftarrow{\$} \{0,1\}^{\lambda-\lambda_0} \times \{0,1\}^{\lambda_0}$ $c^* \leftarrow \mathcal{F}_{\mathsf{pk}}(s^*,t^*)$	$\xrightarrow{1^\lambda,\mathsf{pk},c^*}$			
		$\mathcal{L}_H, \mathcal{L}_G \leftarrow \{\},\{\}$	$\xrightarrow{1^\lambda,\mathsf{pk}}$	
		$\mathcal{O}^{\mathcal{D}_{\mathsf{sk}}}, \mathcal{O}^H, \mathcal{O}^G$ Query Phase 1		
			$\xleftarrow{m_0^*,m_1^*}$	$\mathcal{A}$ chooses $m_0^*, m_1^* \in \mathcal{M}$ such that $\lvert m_0^*\rvert = \lvert m_1^*\rvert$ and $m_0^* \neq m_1^*$
		$b \xleftarrow{\$} \{0,1\}$	$\xrightarrow{c^*}$	
		$\mathcal{O}^{\mathcal{D}_{\mathsf{sk}}}, \mathcal{O}^H, \mathcal{O}^G$ Query Phase 2		
Return $s^* \overset{?}{\in} S'$	$\xleftarrow{S'}$	$\mathcal{B}$ chooses $S' = \{s \in \mathcal{L}_H\}$	$\xleftarrow{b'}$	$\mathcal{A}$ chooses $b' \in \{0,1\}$

먼저, 공격자 $\mathcal{B}$가 $\mathcal{O}^H$를 동작시키는 시뮬레이션을 정의한다. 공격자 $\mathcal{A}$가 랜덤 오라클 $\mathcal{O}^H$에 δ를 질의했다고 하자. 공격자 $\mathcal{B}$는 다음과 같이 H_δ를 응답한다.

1. 만약 δ가 $\mathcal{L}_H$에 있다면, δ에 대응하는 H_δ를 응답한다. (즉, $(\delta, H_\delta) \in \mathcal{L}_H$)

2. 만약 δ가 $\mathcal{L}_H$에 없다면, $H_\delta \xleftarrow{\$} \{0,1\}^{\lambda_0}$을 수행한 후 H_δ를 응답한다. 이후 $\mathcal{L}_H \leftarrow \mathcal{L}_H \cap (\delta, H_\delta)$를 수행한다.

다음으로, 공격자 $\mathcal{B}$가 $\mathcal{O}^G$를 동작시키는 시뮬레이션을 정의한다. 공격자 $\mathcal{A}$가 랜덤 오라클 $\mathcal{O}^G$에 γ를 질의했다고 하자. 공격자 $\mathcal{B}$는 다음과 같이 G_γ를 응답한다.

1. 만약 γ가 $\mathcal{L}_G$에 있다면, γ에 대응하는 G_γ를 응답한다.

2. 만약 γ가 $\mathcal{L}_G$에 없다면, 다음 과정을 진행한다.

   (a) 어떤 $(\delta, H_\delta) \in \mathcal{L}_H$에 대해, 만약 $c^* = \mathcal{F}_{\mathsf{pk}}(\delta, \gamma \oplus H_\delta)$라면, 우리는 여전히 G를 올바르게 시뮬레이션할 수 있다. 이 때 응답은 $G_\gamma \leftarrow \delta \oplus (m_b \parallel 0^{\lambda_1})$이다. $\delta = s^\star$이고 $s^\star$가 균등하게 분포하므로 G_γ는 균등 분포된 값이 된다.

   (b) 모든 $(\delta, H_\delta) \in \mathcal{L}_H$에 대해, 만약 $c^* \neq \mathcal{F}_{\mathsf{pk}}(\delta, \gamma \oplus H_\delta)$라면, $G_\gamma \xleftarrow{\$} \{0,1\}^{n+\lambda_1}$를 수행한다.

   (c) G_γ를 응답한 후, $\mathcal{L}_G \leftarrow \mathcal{L}_G \cap (\gamma, G_\gamma)$를 수행한다.

마지막으로, 공격자 $\mathcal{B}$가 $\mathcal{O}^{\mathcal{D}}$를 동작시키는 시뮬레이션을 정의한다. 공격자 $\mathcal{A}$가 랜덤 오라클 $\mathcal{O}^{\mathcal{D}}$에 $c = \mathcal{F}_{\mathsf{pk}}(s,t)$를 질의했다고 하자. 공격자 $\mathcal{B}$는 다음과 같이 응답한다.

1. $\mathcal{L}_G$의 질의 응답 쌍 $(\gamma, G_\gamma) \in \mathcal{L}_G$ 및 $\mathcal{L}_H$의 $(\delta, H_\delta) \in \mathcal{L}_H$를 조회하고, 각 리스트에서 선택된 쌍에 대해 다음과 같이 정의한다.

$$\sigma = \delta, \quad \tau = \gamma \oplus H_\delta, \quad \mu = G_\gamma \oplus \delta.$$

   만약 $c = \mathcal{F}_{\mathsf{pk}}(\sigma,\tau)$이면서 $[\mu]_{\lambda_1} = 0^{\lambda_1}$라면, $[\mu]^n$을 응답한다.

2. 그 외에는 Reject를 응답한다.

4.2 증명: 사건 정의

Table 1: 오라클 관련 사건 정의

AskG	r^*가 $\mathcal{O}^G$에 질의되었을(has been asked) 사건.
AskH	s^*가 $\mathcal{O}^H$에 질의되었을 사건.
GBad	$\mathcal{O}^G$에 r^*를 질의했지만, $\mathcal{O}^G$의 응답이 $s^* \oplus (m_b \parallel 0^{\mathsf{sk}})$가 아닌 사건. GBad가 발생하면, AskG도 발생한다.
DBad	~~CPA 시나리오에서 복호화가 실패하는 사건.~~
Bad	~~GBad ∨ DBad.~~

공격자 $\mathcal{A}$는 복호화 오라클 $\mathcal{O}^{\mathcal{D}}$에 암호문 $c = \mathcal{F}_{\mathsf{pk}}(s,t)$를 질의할 수 있다. 질의한 암호문 c와 관련된 사건을 다음 표와 같이 정의한다.

Table 2: 복호화 시뮬레이션 관련 사건 정의

SBad	$s = s^*$인 사건.
RBad	$r = r^*$인 사건. 즉, $H(s) \oplus t = H(s^*) \oplus t^*$인 사건.
CBad	SBad ∨ RBad.
AskR	r이 $\mathcal{O}^G$에 질의되었을 사건. 즉, $H(s) \oplus t$이 질의되었을 사건
AskS	s가 $\mathcal{O}^H$에 질의되었을 사건.
AskRS	AskR ∧ AskS
Fail	복호화 오라클이 질의 c에 대해 잘못 응답하는 사건. i번째 질의 c_i에 대해서는 Fail_i로 나타낸다. 여기서 $i = 1, \cdots, q_{\mathcal{D}}$이다. 어떤 i에 대해서도 Fail_i의 확률을 균등하게 평가(evaluate)할 수 있으므로, 여기서는 사용하지 않는다. Fail 사건은 평문 추출기(plaintext extractor)가 실제 복호화 오라클에서는 허용될 암호문을 거부하는 경우로 제한된다. 실제로, 추출기가 암호문을 허용하는 순간, 해당 암호문이 유효하며 출력 평문과 일치함을 알 수 있다.

4.3 증명: Analysis of the Decryption Oracle Simulation

보조정리 2. *$s^\star$가 $\mathcal{O}^H$에 질의되지 않았을 때, $\mathcal{O}^\mathcal{D}$는 질의된 암호문 c $(c \neq c^\star)$에 대해 출력을 정확히 생성할 수 있으며, 이 확률은 다음보다 크거나 같다.*

$$1 - \left(\frac{2}{2^{k_1}} + \frac{2q_G + 1}{2^{k_0}}\right).$$

또한, 시간 제한 $t' \leq q_G \cdot q_H \cdot (T_\mathcal{F} + \mathcal{O}(1))$ 내에서 이를 수행할 수 있다.

Proof. 본 증명에서는 다음이 참임을 보인다.

$$\Pr[\mathsf{Fail} \mid \neg\mathsf{AskH}] \leq \frac{2}{2^{\lambda_1}} + \frac{2q_G + 1}{2^{\lambda_0}}.$$

$\Pr[\mathsf{Fail} \mid \neg\mathsf{AskH}]$는 다음과 같이 표현 가능하다.

$$\Pr[\mathsf{Fail} \wedge \mathsf{CBad} \mid \neg\mathsf{AskH}] + \Pr[\mathsf{Fail} \wedge \neg\mathsf{CBad} \mid \neg\mathsf{AskH}].$$

본 증명에서는 먼저 $\Pr[\mathsf{Fail} \wedge \mathsf{CBad} \mid \neg\mathsf{AskH}]$를 구한다. $\mathsf{CBad} = \mathsf{SBad} \vee (\mathsf{RBad} \wedge \neg\mathsf{SBad})$를 이용하여, 이 확률을 다음과 같이 표현한다.

$$\begin{aligned}
&\Pr[\mathsf{Fail} \wedge \mathsf{CBad} \mid \neg\mathsf{AskH}] \\
&= \Pr[\mathsf{Fail} \wedge (\mathsf{SBad} \vee (\mathsf{RBad} \wedge \neg\mathsf{SBad})) \mid \neg\mathsf{AskH}] \\
&= \Pr[(\mathsf{Fail} \wedge \mathsf{SBad}) \vee (\mathsf{Fail} \wedge \mathsf{RBad} \wedge \neg\mathsf{SBad}) \mid \neg\mathsf{AskH}] \\
&\leq \Pr[\mathsf{Fail} \wedge \mathsf{SBad} \mid \neg\mathsf{AskH}] + \Pr[\mathsf{Fail} \wedge \mathsf{RBad} \wedge \neg\mathsf{SBad} \mid \neg\mathsf{AskH}] \\
&\leq \Pr[\mathsf{Fail} \wedge \mathsf{SBad} \mid \neg\mathsf{AskH}] + \Pr[\mathsf{RBad} \wedge \neg\mathsf{SBad} \mid \neg\mathsf{AskH}] \\
&\leq \Pr[\mathsf{Fail} \mid \mathsf{SBad} \wedge \neg\mathsf{AskH}] + \Pr[\mathsf{RBad} \mid \neg\mathsf{SBad} \wedge \neg\mathsf{AskH}]. \\
&= \Pr[\mathsf{Fail} \wedge \mathsf{AskR} \mid \mathsf{SBad} \wedge \neg\mathsf{AskH}] + \Pr[\mathsf{Fail} \wedge \neg\mathsf{AskR} \mid \mathsf{SBad} \wedge \neg\mathsf{AskH}] + \Pr[\mathsf{RBad} \mid \neg\mathsf{SBad} \wedge \neg\mathsf{AskH}]. \\
&\leq \Pr[\mathsf{AskR} \mid \mathsf{SBad} \wedge \neg\mathsf{AskH}] + \Pr[\mathsf{Fail} \mid \neg\mathsf{AskR} \wedge \mathsf{SBad} \wedge \neg\mathsf{AskH}] + \Pr[\mathsf{RBad} \mid \neg\mathsf{SBad} \wedge \neg\mathsf{AskH}].
\end{aligned}$$

세 번째 사건은 $s \neq s^\star$이고 공격자 $\mathcal{A}$가 $s^\star$에 대해 $\mathcal{O}^H$에 질의하지 않았을 때 RBad가 발생함을 의미한다. $s^\star$가 $\mathcal{O}^H$에 질의되지 않았고 $s \neq s^\star$일 때, $H(s^\star)$는 예측 불가능(unpredictable)하며 $H(s)$뿐 아니라 t, $t^\star$와도 독립적이다. 이때 RBad 사건, $H(s^\star) = H(s) \oplus t \oplus t^\star$ 는 최대 $2^{-\lambda_0}$의 확률로 발생한다. 즉, 다음과 같다.

$$\Pr[\mathsf{RBad} \mid \neg\mathsf{SBad} \wedge \neg\mathsf{AskH}] \leq 2^{-\lambda_0}.$$

첫 번째 사건은 $s = s^\star$이며 $H(s^\star)$는 예측 불가능할 때, r이 $\mathcal{O}^G$에 대해 질의되었을 사건을 의미한다. 이때, $H(s)$ 또한 예측 불가능하다. 즉, $r = H(s) \oplus t$가 예측 불가능하므로, r이 $\mathcal{O}^G$에 질의되었을 확률은 최대 $q_G \cdot 2^{-\lambda_0}$이다. 즉, 다음과 같다.

$$\Pr[\mathsf{AskR} \mid \mathsf{SBad} \wedge \neg\mathsf{AskH}] \leq q_G \cdot 2^{-\lambda_0}.$$

두 번째 사건은 복호화 시뮬레이션에서 $H(s)$는 예측 불가능하고 r은 $\mathcal{O}^G$에 질의되지 않았을 때, 유효한 암호문 c를 거부하는 경우이다. 페이스텔 네트워크(Feistel network)의 일대일 성질에 따라 $s = s^\star$이면 $r \neq r^\star$이고, 따라서 $G(r)$는 예측 불가능하다. 그러므로 이 경우 중복 조건은 $2^{-\lambda_1}$보다 큰 확률로 성립할 수 없다. 즉, 다음과 같다.

$$\Pr[\mathsf{Fail} \mid \neg\mathsf{AskR} \wedge \mathsf{SBad} \wedge \neg\mathsf{AskH}] \leq 2^{-\lambda_1}.$$

세 식을 결합하면, 다음과 같다.

$$\Pr[\mathsf{Fail} \wedge \mathsf{CBad} \mid \neg\mathsf{AskH}] \le 2^{-k_1} + (q_G + 1) \cdot 2^{-k_0}.$$

다음으로, $\Pr[\mathsf{Fail} \wedge \neg\mathsf{CBad} \mid \neg\mathsf{AskH}]$를 계산하고 본 증명을 마친다. 만약 $\neg\mathsf{CBad} \wedge \mathsf{AskRS}$가 성립한다면, 복호화 시뮬레이션은 실패하지 않는다. 따라서 이 식은 아래와 같이 표현 가능하다.

$$\begin{aligned}
&\Pr[\mathsf{Fail} \wedge \neg\mathsf{CBad} \mid \neg\mathsf{AskH}] \\
&= \underbrace{\Pr[\mathsf{Fail} \wedge \neg\mathsf{CBad} \wedge \mathsf{AskRS} \mid \neg\mathsf{AskH}]}_{=0} + \Pr[\mathsf{Fail} \wedge \neg\mathsf{CBad} \wedge \neg\mathsf{AskRS} \mid \neg\mathsf{AskH}] \\
&= \Pr[\mathsf{Fail} \wedge \neg\mathsf{CBad} \wedge \neg\mathsf{AskRS} \mid \neg\mathsf{AskH}].
\end{aligned}$$

이제 $\neg\mathsf{AskH}$를 잠시 고려하지 않고, 위 확률을 다음과 같이 계산한다.

$$\begin{aligned}
&\Pr[\mathsf{Fail} \wedge \neg\mathsf{CBad} \wedge \neg\mathsf{AskRS}] \\
&= \Pr[\mathsf{Fail} \wedge \neg\mathsf{RBad} \wedge \neg\mathsf{SBad} \wedge (\neg\mathsf{AskR} \vee \neg\mathsf{AskS})] \\
&= \Pr[\mathsf{Fail} \wedge \neg\mathsf{RBad} \wedge \neg\mathsf{SBad} \wedge (\neg\mathsf{AskR} \vee (\neg\mathsf{AskS} \wedge \mathsf{AskR}))] \\
&= \Pr[(\mathsf{Fail} \wedge \neg\mathsf{RBad} \wedge \neg\mathsf{SBad} \wedge \neg\mathsf{AskR}) \vee (\mathsf{Fail} \wedge \neg\mathsf{RBad} \wedge \neg\mathsf{SBad} \wedge (\neg\mathsf{AskS} \wedge \mathsf{AskR}))] \\
&\le \Pr[\mathsf{Fail} \wedge \neg\mathsf{RBad} \wedge \neg\mathsf{SBad} \wedge \neg\mathsf{AskR}] + \Pr[\mathsf{Fail} \wedge \neg\mathsf{RBad} \wedge \neg\mathsf{SBad} \wedge \neg\mathsf{AskS} \wedge \mathsf{AskR}] \\
&\le \Pr[\mathsf{Fail} \wedge \neg\mathsf{RBad} \wedge \neg\mathsf{AskR}] + \Pr[\mathsf{Fail} \wedge \mathsf{AskR} \wedge \neg\mathsf{AskS} \wedge \neg\mathsf{SBad}] \\
&\le \Pr[\mathsf{Fail} \wedge \neg\mathsf{RBad} \mid \neg\mathsf{AskR}] + \Pr[\mathsf{Fail} \wedge \mathsf{AskR} \mid \neg\mathsf{AskS} \wedge \neg\mathsf{SBad}] \\
&\le \Pr[\mathsf{Fail} \mid \neg\mathsf{RBad} \wedge \neg\mathsf{AskR}] + \Pr[\mathsf{AskR} \mid \neg\mathsf{AskS} \wedge \neg\mathsf{SBad}].
\end{aligned}$$

첫 번째 사건에서, r이 $\mathcal{O}^G$에 대해 질의되지 않았고, 추가로 $r \neq r^*$인 사건을 고려하면, $G(r)$는 예측할 수 없으며, 따라서 $[s \oplus G(r)]_{\lambda_1} = 0^{\lambda_1}$이 될 확률은 $2^{-\lambda_1}$보다 작다. 그리고 두 번째 사건에서, $H(s)$에 대한 정보 없이 r이 $\mathcal{O}^G$에 대해 질의될 확률은 $q_G \cdot 2^{-\lambda_0}$보다 작다. 또한, 이 사건은 AskH와 독립적이므로 다음이 성립한다.

$$\Pr[\mathsf{Fail} \wedge \neg\mathsf{CBad} \wedge \neg\mathsf{AskRS} \mid \neg\mathsf{AskH}] \le 2^{-\lambda_1} + q_G \cdot 2^{-\lambda_0}.$$

그러므로, 다음과 같다.

$$\begin{aligned}
\Pr[\mathsf{Fail} \mid \neg\mathsf{AskH}] &= \Pr[\mathsf{Fail} \wedge \mathsf{CBad} \mid \neg\mathsf{AskH}] + \Pr[\mathsf{Fail} \wedge \neg\mathsf{CBad} \mid \neg\mathsf{AskH}] \\
&\le (2^{-\lambda_1} + (q_G + 1) \cdot 2^{-\lambda_0}) + (2^{-\lambda_1} + q_G \cdot 2^{-\lambda_0}). \\
&= \frac{2}{2^{\lambda_1}} + \frac{2q_G + 1}{2^{\lambda_0}}.
\end{aligned}$$

이 시뮬레이터의 실행 시간은 가능한 모든 쌍에 대해 $\mathcal{F}_{\mathsf{pk}}(\sigma, \tau)$를 계산하는 시간만 포함되며, 따라서 그 시간은 다음으로 상한된다.

$$q_G \cdot q_H \cdot (T_{\mathcal{F}} + \mathcal{O}(1)).$$

$\square$