九、公钥加密理论 (Pulic-Key Encryptions Principle)

哈尔滨工业大学 张宇 2024春

概览

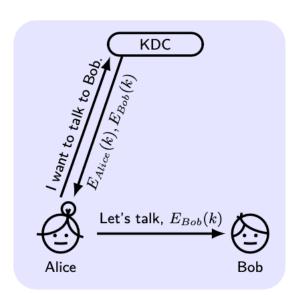
- 1. 公钥加密安全定义
- 2. 基于陷门排列(Trapdoor Permutation)的公钥加密
- 3. CCA安全公钥加密
- 4. 随机预言机模型 (ROM) 中基于TDP的公钥加密

私钥密码学局限性

- □密钥分发需要通信各方在物理上会面;
- □\$U\$个用户的密钥的数量 \$\Theta(U^2)\$;
- □开放系统的安全通信:基于私钥密码学的解决方案无法充分处理开放系统中的安全通信问题,在开放系统中通信各方不能物理上会面,或只能暂时交互;
 - □私钥密码中的一个核心问题就是密钥分发与管理问题

Needham-Schroeder对称密钥协议

- □*Needham Schroeder Symmetric Key Protocol*: 在开放网络中双方通过一个可信的第三方建立一个会话密钥 (session key);
- □密钥分发中心(Key Distribution Center, KDC)作为可信的第三方 (Trusted Third Party, TTP),与通信双方Alice和Bob在事前分别建 立了对称密钥;
- □KDC根据Alice的请求,生成一个新的\$k\$会话密钥(session key), 分别用与Alice和Bob分别共享的密钥来加密并发送给Alice; \$E_{Bob}(k)\$作为一个来访问Bob所需的凭证(ticket);
- □用于MIT's Kerberos 协议 (in Windows);
- □优点:每一方只需要存储一个密钥;不需要 更新通信双方密钥(因为采用新的会话密钥)
- □弱点:单点失效,一旦KDC被破坏,则整个系统都不安全。



默克尔难题(Merkle Puzzles)

o Alice准备 2^{32} 个难题 $Puzzle_i$,并且发送给Bob;难题如下:

Puzzle_i ← Enc_(0⁹⁶||p_i)("Puzzle #" x_i || k_i),,其中 Enc 是 128位加密, p_i ← $\{0,1\}^{32}$ 并且 x_i, k_i ← $\{0,1\}^{128}$ 。

注:每个难题中明文包括一个随机数和一个密钥,用一个密钥加密;

- o Bob随机选择一个难题 Puzzle_j ,并且在 2^{32} 时间内猜测 p_j ,获得 x_j, k_j 并将 x_j 发送给 Alice。
- \circ Alice 按照 x_j 查询谜题,并且使用 k_j 作为密钥。
- \circ 敌手需要 2^{32+32} 时间,是诚实方所需时间复杂性的二次方。
- 在诚实方和敌手之间存在更好的差距吗?如果将加密方法看作是一个黑 盒预言机,那么二次差距是最好的。
- Merkle难题的缺点是谜题数量太大,获得密钥的代价太大;
- 注: Merkle当时是UC的一名本科生,这是他的一门课程设计申请。

公钥革命

- □在1976年, Whitfield Diffie 和 Martin Hellman 发表了"New Directions in Cryptography"(密码学的新方向)。在这篇论文中,提出公钥加密方案、陷门(Trap door)和数字签名等概念。
- □非对称 (Asymmetric) 或公钥 (public-key) 加密方案:
 - □公钥 (Public key) 加密密钥; (注:接收方产生,发送方持有)
 - □私钥(Private key)解密密钥; (注:接收方产生,接收方持有)
- □公钥原语 (Public-key primitives):
 - □公钥加密 (Public-key encryption)
 - □数字签名 (Digital signatures) (不可抵赖性, non-repudiation)
 - □交互式密钥交换 (Interactive key exchange)
- □优点:
 - □在公开信道上密钥分发
 - □减少保存大量密钥的需求
 - □使得在开放系统的安全成为可能
- □缺点: 慢两到三个数量级, 针对公钥分发的主动攻击
 - □注:如何保证Alice得到的公钥真的是Bob的公钥?



□问题: 谁发送了消息?

I'M SURE YOU'VE HEARD ALL ABOUT THIS SORDID AFFAIR IN THOSE GOSSIPY CRYPTOGRAPHIC PROTOCOL SPECS WITH THOSE BUSYBODIES SCHNEIER AND RIVEST, ALWAYS TAKING ALICE'S SIDE, ALWAYS LABELING ME THE ATTACKER.



YES, IT'S TRUE. I BROKE BOB'S PRIVATE KEY AND EXTRACTED THE TEXT OF HER MESSAGES. BUT DOES ANYONE REALIZE HOW MUCH IT HURT?



HE SAID IT WAS NOTHING, BUT EVERYTHING FROM THE PUBLIC-KEY AUTHENTICATED SIGNATURES ON THE FILES TO THE LIPSTICK HEART SMEARED ON THE DISK SCREAMED "ALICE."

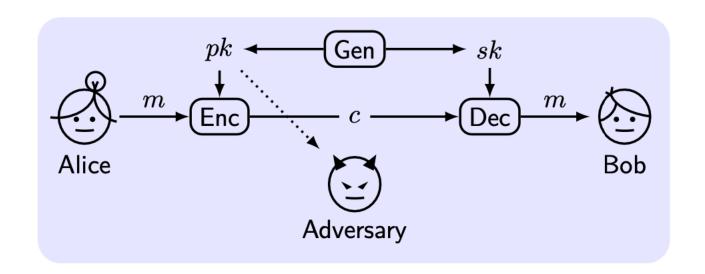


DIDN'T WANT TO BELIEVE. OF COURSE ON SOME LEVEL 1 REALIZED IT WAS A KNOWN-PLAINTEXT ATTACK. BUT 1 COULDN'T ADMIT IT UNTIL / I SAW FOR MYSELF.



SO BEFORE YOU SO QUICKLY LABEL ME A THIRD PARTY TO THE COMM-UNICATION, JUST REMEMBER: I LOVED HIM FIRST. WE HAD SOMETHING AND SHE / TORE IT AWAY. SHE'S THE ATTACKER, NOT ME. NOT EVE.

公钥加密词法



- **Key-generation** algorithm: $(pk, sk) \leftarrow \text{Gen}$, key length $\geq n$.
- **Plaintext space** \mathcal{M} is associated with pk.
- **Encryption** algorithm: $c \leftarrow \mathsf{Enc}_{pk}(m)$.
- **Decryption** algorithm: $m := \mathsf{Dec}_{sk}(c)$, or outputs \bot .
- **Requirement**: $\Pr[\mathsf{Dec}_{sk}(\mathsf{Enc}_{pk}(m)) = m] \ge 1 \mathsf{negl}(n)$.

公钥加密窃听者=CPA

□由于公钥是公开的,敌手不仅能窃听,且能够加密任意明文。

The eavesdropping indistinguishability experiment PubK $_{\mathcal{A},\Pi}^{\mathsf{eav}}(n)$:

- 2 \mathcal{A} is given input \mathbf{pk} and so oracle access to $\mathsf{Enc}_{\mathbf{pk}}(\cdot)$, outputs m_0, m_1 of the same length.
- $b \leftarrow \{0,1\}.$ $c \leftarrow \mathsf{Enc}_{pk}(m_b)$ (challenge) is given to \mathcal{A} .
- 4 Continues to have access to $Enc_{\mathbf{pk}}(\cdot)$ and outputs b'.
- If b' = b, \mathcal{A} succeeded PubK^{eav}_{\mathcal{A},Π} = 1, otherwise 0.

Definition 1

 Π is **CPA-secure** if \forall PPT \mathcal{A} , \exists negl such that

$$\Pr\left[\mathsf{PubK}^{\mathsf{cpa}}_{\mathcal{A},\Pi}(n) = 1
ight] \leq rac{1}{2} + \mathsf{negl}(n).$$

课堂练习(公钥加密的安全性)

- □对称加密可以加密32比特消息,产生32比特密文,例如,使用一次一密。在公钥系统中能够做到同样的吗?
- □一个确定性的公钥加密方案在窃听者出现时是安全的?
- □如果 \$\Pi\$ 在窃听者出现时是安全的,那么 \$\Pi\$ 也是 CPA安全的?是否是多重加密安全的?
- □完美保密的公钥加密是可能的吗?

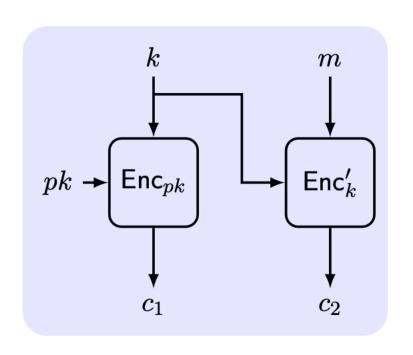
密钥长度

□NIST (美国国家标准技术研究所)推荐可比较的密钥长度 (按比特)。NIST 认为一个112比特的有效密钥长度直到 2030年是可接受的,但是推荐 128 比特或更长的密钥。

AES	$RSA\ (N)/DH\ (p)$	ECC (order q)
56	512	112
80	1024	160
112	2048	224
128	3072	256
192	7680	384
256	15360	512

混合加密(Hybrid Encryption)构造

□为了加速加密,采用私钥加密方案 \$\Pi'\$(数据封装机制,data-encapsulation mechanism, DEM) 与公钥加密方案 \$\Pi\$(密钥封装机制, key-encapsulation mechanism, KEM) 一起。



Construction 5

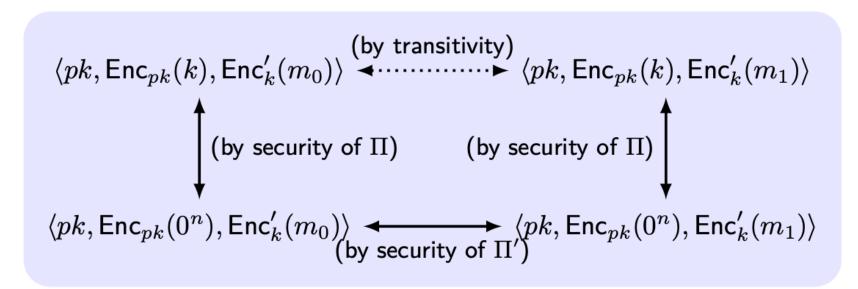
 $\Pi^{\mathsf{h}\mathsf{y}} = (\mathsf{Gen}^{\mathsf{h}\mathsf{y}}, \mathsf{Enc}^{\mathsf{h}\mathsf{y}}, \mathsf{Dec}^{\mathsf{h}\mathsf{y}})$:

- $\operatorname{\mathsf{Gen}}^{\mathsf{hy}}$: $(pk, sk) \leftarrow \operatorname{\mathsf{Gen}}(1^n)$.
- \blacksquare Enc^{hy}: pk and m.
 - 1 $k \leftarrow \{0,1\}^n$.
 - $c_1 \leftarrow \mathsf{Enc}_{pk}(k), \\ c_2 \leftarrow \mathsf{Enc}'_k(m).$
- Dec^{hy}: sk and $\langle c_1, c_2 \rangle$.

 - $m := \mathsf{Dec}'_k(c_2).$
- □混合加密是公钥加密还是私钥加密?

混合加密安全性

- □定理:如果\$\Pi\$是一个CPA安全的公钥加密方案,并且\$\Pi'\$是窃听者不可区分的私钥加密方案,那么\$\Pi^{\mathsf{hy}}\$是CPA安全的公钥加密方案。
- □这里对于私钥加密方案的安全性要求只是窃听者不可区分的,不要求是CPA安全的,因为私钥加密密钥是每次加密时随机产生的新密钥,私钥加密预言机无法被利用。



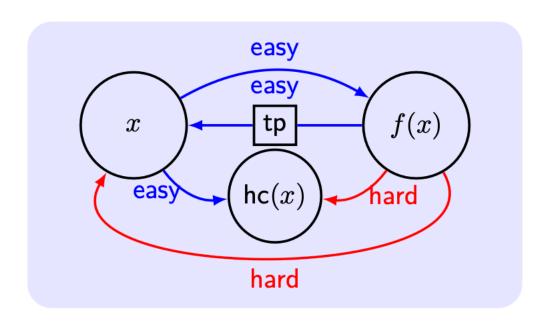
课堂练习 (混合加密应用)

□共享文件访问: 在一个现代文件系统中,一个人 (Alice) 加密的文件只有指定的某些人 (Bob、Chalie) 才能查看。如何用混合加密方案实现?

□密钥托管:基于以上文件系统方案,一个公司运行一个密钥托管服务器,并产生一个公钥对。一天,Alice出去旅行并且联系不上了;她的主管Bob要访问一个Alice的文件。如何用混合加密方案实现?

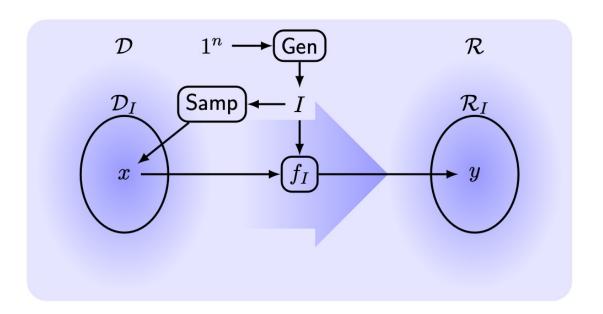
陷门排列(Trapdoor Function)

- □陷门函数 (Trapdoor function): 易于计算,在缺乏特定信息(陷门)时难以求逆,即带有陷门的单向函数。
- □1982年,姚期智在论文《Theory and Applications of Trapdoor Functions》中提出,从任意陷门函数中可构造一个公钥加密方案。



函数族

□ 这里强调采样算法是因为后面要学习的数论难题的输入是要满足某些 条件的。



Definition 7

 $\Pi = (\mathsf{Gen}, \mathsf{Samp}, f)$ is a **family of functions** if:

- **1 Parameter-generation** algorithm: $I \leftarrow \text{Gen}(1^n)$.
- **2** sampling algorithm: $x \leftarrow \mathsf{Samp}(I)$.
- **3** The deterministic **evaluation** algorithm: $y := f_I(x)$.

陷门排列族

- \circ 一组多项式时间算法 $\Pi=(\mathsf{Gen},\mathsf{Samp},f,\mathsf{Inv})$ 是一个陷门排列族(family of trapdoor permutations,TDP),如果:
 - 参数生成(parameter generation)算法 **Gen**, 输入 1^n ,输出 (I, td) 有 $|I| \geq n$ 。其中, (I, td) 定义了集合 $\mathcal{D}_I = \mathcal{D}_{td}$ 。注:陷门排列族是一个函数集合,参数生成算法产生一个具体陷门排列所需的参数。
 - Gen_I 只输出 I 。 $(\mathsf{Gen}_I, \mathsf{Samp}, f)$ 是 OWP。其中的 Samp 是采样函数,用于获得函数的输入 $x \leftarrow \mathcal{D}_I$ 。
 - 一个确定性求逆算法 Inv,对于 $\forall (I, td)$ 并且 $\forall x \in \mathcal{D}_I$, $Inv_{td}(f_I(x)) = x$ 。 $ag{注: 可求逆}$
- 。 核心断言:确定性多项式时间算法 $\operatorname{hc} \, \exists \, \operatorname{negl} \,$ 的一个核心断言(hard-core predicate),如果 $\forall \, \operatorname{ppt} \, \mathcal{A}, \, \exists \, \operatorname{negl} \,$ 使得 $\operatorname{Pr}[\mathcal{A}(I,f_I(x))=\operatorname{hc}_I(x)]\leq \frac{1}{2}+\operatorname{negl}(n)$ 。
- 。 定理: 给定一个陷门排列族 $\Pi = (\mathsf{Gen}, \mathsf{Samp}, f, \mathsf{Inv})$,则存在一个带有核心断言的陷门排列族 $\widehat{\Pi} = (\widehat{\mathsf{Gen}}, \mathsf{Samp}, f, \mathsf{Inv})$ 。注:证明与单向函数部分关于核心断言的定理类似。

课堂练习

Let f with < I, td > be a TDP. Which of the following f' is also a TDP?

- $f'(x) = f(x) \| \mathsf{td}$
- f'(x) = f(x) || I
- $f'(x||x') = f(x)||\operatorname{Inv}_{\mathsf{td}}(f(x'))|$
- f'(x||x') = f(x)||f(x')||

Is the following public-key encryption scheme from any TDP is secure?

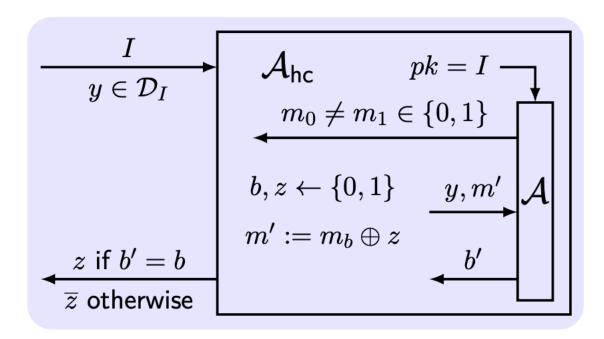
$$\operatorname{Enc}_I(m) = f_I(m)$$
, $\operatorname{Dec}_{\mathsf{td}}(c) = f_I^{-1}(c)$.

从TDP到公钥加密

- 。 从一个带有核心断言hc的陷门排列族 $\widehat{\Pi}=(\widehat{\mathsf{Gen}},\mathsf{Samp},f,\mathsf{Inv})$ 来构造一个公钥加密方案:
 - Gen: $(I, td) \leftarrow \widehat{Gen}$ 输出公钥 I 和私钥 td。
 - Enc: 输入 I 和 $m \in \{0,1\}$,选择一个 $x \leftarrow \mathcal{D}_I$ 并且输出 $\langle f_I(x), \mathsf{hc}_I(x) \oplus m \rangle$ 。
 - Dec: 输入 td 和 $\langle y, m' \rangle$,计算 $x := f_I^{-1}(y)$ 并且输出 $\mathsf{hc}_I(x) \oplus m'$ 。
- 。 定理: 如果 $\widehat{\Pi}=(\widehat{Gen},f)$ 是 TDP,并且 \mathbf{hc} 是 $\widehat{\Pi}$ 的 HCP , 那么构造 Π 是 CPA安全的。
- 。 问题:这个方案是安全的吗? $\mathsf{Enc}_I(m) = f_I(m), \;\; \mathsf{Dec}_{\mathsf{td}}(c) = f_I^{-1}(c).$

证明

Idea: $hc_I(x)$ is pseudorandom. Reduce \mathcal{A}_{hc} for hc to \mathcal{A} for Π .



$$\Pr[\mathcal{A}_{\mathsf{hc}}(I,f_I(x)) = \mathsf{hc}_I(x)] = rac{1}{2} \cdot (\Pr[b' = b|z = \mathsf{hc}_I(x)] + \Pr[b'
eq b|z \neq \mathsf{hc}_I(x)]).$$

证明 (续)

$$\Pr[b'=b|z=\mathsf{hc}_I(x)]=\Pr[\mathsf{PubK}^{\mathsf{eav}}_{\mathcal{A},\Pi}(n)=1]=rac{1}{2}+arepsilon(n).$$

If $z \neq \mathrm{hc}_I(x)$, $m' = m_b \oplus \overline{\mathrm{hc}}_I(x) = m_{\overline{b}} \oplus \mathrm{hc}_I(x)$, which means $m_{\overline{b}}$ is encrypted.

$$\Pr[b'=b|z
eq \mathsf{hc}_I(x)] = \Pr[\mathsf{PubK}^{\mathsf{eav}}_{\mathcal{A},\Pi}(n)=0] = rac{1}{2} - arepsilon(n).$$

$$\Pr[b'
eq b|z
eq \mathsf{hc}_I(x)] = rac{1}{2} + arepsilon(n).$$

$$\Pr[\mathcal{A}_{\mathsf{hc}}(I,f_I(x)) = \mathsf{hc}_I(x)] = rac{1}{2} + arepsilon(n).$$

□至此,我们学习了基于陷门排列的公钥加密方案,但只能加密一个比特,如何加密一个更长的明文?后面学习随机预言机模型设定下的公钥加密方案。

对公钥加密方案的CCA攻击

- □在公钥设定中CCA情景
 - □敌手A观察由S发送给R的密文c。
 - □A以S或自己的名义发送c'给R。
 - □A根据从c'中解密出的m'来推断m。
- □情景
 - □用口令来登陆在线银行: 试错, 从银行反馈中获得信息。
 - □邮件回复中包含解密出的文本的引用。
 - □密文的可锻造性,例如,在拍卖中将其他人的出价翻倍。

CCA安全定义

The CCA/CCA2 indistinguishability experiment PubK^{cca}_{A,Π}(n):

- 2 \mathcal{A} is given input pk and oracle access to $Dec_{sk}(\cdot)$, outputs m_0, m_1 of the same length.
- 3 $b \leftarrow \{0,1\}.$ $c \leftarrow \mathsf{Enc}_{pk}(m_b)$ is given to \mathcal{A} .
- 4 A have access to $Dec_{sk}(\cdot)$ except for c in CCA2² and outputs b'.
- If b' = b, \mathcal{A} succeeded PubK^{cca}_{\mathcal{A},Π} = 1, otherwise 0.

Definition 10

 Π has **CCA/CCA2-secure** if \forall PPT \mathcal{A} , \exists negl such that

$$\Pr\left[\mathsf{PubK}^{\mathsf{cca}}_{\mathcal{A},\Pi}(n) = 1\right] \leq \frac{1}{2} + \mathsf{negl}(n).$$

²CCA is also called Lunchtime attacks; CCA2 is also called Adaptive CCA.

课堂练习

Let (Gen, E, D) be CCA-secure on message space $\{0, 1\}^{128}$. Which of the following is also CCA-secure?

- $E'(pk,m) = (E(pk,m),0^{128})$ $D'(sk,(c_1,c_2)) = \left\{ egin{array}{ll} D(sk,c_1) & ext{if } c_2 = 0^{128} \ ot & ext{otherwise} \end{array}
 ight.$
- $E'(pk, m) = (E(pk, m), E(pk, 0^{128}))$ $D'(sk, (c_1, c_2)) = D(sk, c_1)$

CCA安全技术进展

- □零知识证明 (Zero-Knowledge Proof): 复杂并不可实践。 (例如, Dolev-Dwork-Naor)
- □随机预言机模型 (Random Oracle model): 有效,但并不踏实(将 CRHF 当作 RO)。(例如, RSA-OAEP 和 Fujisaki-Okamoto)
- □DDH (决策性Diffie-Hellman) 假设和UOWHF (全域单向哈希函数): 大小扩展2倍,但可以在没有RO和ZKP场景下证明安全 (例如, Cramer-Shoup system)。
- □CCA2安全意味着明文感知 (Plaintext-aware): 敌手在不知道明文的情况下,不能产生有效的密文。
- □开放问题:如何构造一个与"书本上RSA"一样有效的,基于RSA问题的CCA2安全的方案。

随机预言机模型 (ROM)

- □为了在实践中实现CPA安全和CCA安全的公钥加密方案,引入了一个更强大的随机对象,称为随机预言机 (Random Oracle Model)。
- □随机预言机 (RO): 一个真随机函数 (H) 对每个可能的查询回答一个随机应答。
 - □一致性:如果H曾经在运行中为一个输入 x 输出 y, 那么它一直对相同的输入输出相同的答案。
 - □无人"知道"整个函数 H。
- □随机预言机模型 (ROM): 存在一个公开的RO。与此相对的,不存在RO的情况,称作标准模型。
- □方法论: 在ROM中构造可证明的安全。很容易实现可证明安全,同时通过正确的实例化来保持高效。
 - □ 1. 在ROM中, 一个方案被设计并被证明是安全的。
 - □ 2. 将 \$H\$ 用一个哈希函数 \$\hat{H}\$, 例如 SHA256。
- □无人严格地声明随机预言机存在。存在某些方案,在ROM中被证明是 安全的,但无论如何将随机预言机实例化都不是安全的。

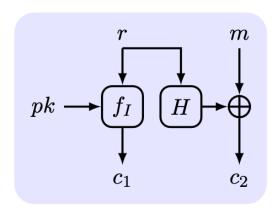
ROM例子

- 由于RO "强大的随机性",其可以充当或构造之前学习过得密码学原语,包括为单向 函数、抗碰撞哈希函数、伪随机函数等。
- 一个 RO 将 n_1 比特输入映射为 n_2 比特输出。
- RO 作为 OWF, 进行如下实验:
 - 1. 选择一个ROH;
 - 2. 选择一个随机的 $x \in \{0,1\}^{n_1}$,并且赋值 y := H(x) ;
 - 3. 敌手 A 被给予 y,如果输出 x': H(x') = y,则成功;解释:如果敌手成功求逆,则意味着敌手"事先"询问过RO;
- RO 作为 CRHF, 进行如下实验:
 - 1. 选择一个ROH;
 - 2. 敌手 \mathcal{A} 成功,如果其输出 x, x' 满足 H(x) = H(x') ,但是 $x \neq x'$;解释:如果敌手找到碰撞,则意味着H不是随机的,因为两个随机输出不可能相同。
- 。 从一个RO构造PRF: $n_1=2n$, $n_2=n$.
 - $lacksquare F_k(x) \stackrel{
 m def}{=} H(k\|x), |k| = |x| = n.$

解释:如果F不是伪随机的,则H也可以与真随机相区分。

基于ROM的CPA安全

- □思路: PubK CPA = PrivK + (Secret Key = TDP + RO)
- □实现CPA安全的公钥加密方案,可以基于一个安全的私钥加密方案, 其中私钥加密的密钥由RO得到,通过TDP传递生成密钥所用的随机量;
- □定理:如果\$f\$ 是TDP, 并且\$H\$ 是RO,则构造是CPA安全的。
 - □解释: 私钥加密方案只需要是窃听下安全,因为每次加密都是概率性的,每次私钥加密密钥都是重新生成的。该方案不是CCA安全的,因为篡改密文可以直接影响明文。
- □用RO的必要性:由于\$r\$的部分信息可能通过TPD泄漏,如果以一个PRG来替换掉RO,则由于种子的部分信息已知,PRG的输出也不在是伪随机的,加密方案也不再安全。

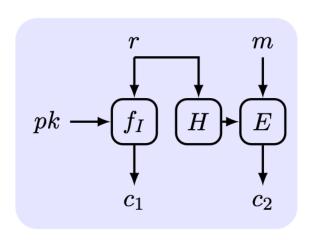


Construction 11

- lacksquare Gen: pk = I, $sk = \mathsf{td}$
- Enc: $r \leftarrow \{0,1\}^*$, output $\langle c_1 = f_I(r), c_2 = H(r) \oplus m \rangle$
- Dec: $r:=f_{\sf td}^{-1}(c_1)$, output $H(r)\oplus c_2$

ROM中基于私钥加密安全的CCA安全

- □ 思路: PubK CCA = PrivK CCA + (Secret Key = TPD + RO)
- □实现CCA安全的公钥加密方案,可以基于一个CCA安全的私钥加密方案,其中私钥加密密钥由RO得到,通过TDP传递生成密钥所用的随机量;-定理:如果\$f\$是TDP,\$\Pi'\$是CCA安全的,并且\$H\$是RO,那么构造是CCA安全的。
- □解释:公钥加密方案的CCA安全性来自私钥加密方案的CCA安全性。

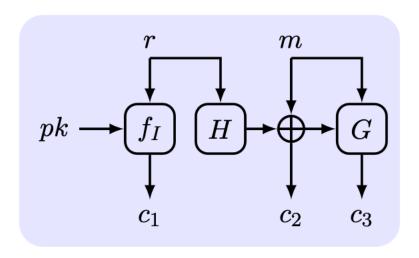


Construction 13

- $\blacksquare \Pi'$ is PrivK
- Gen: pk = I, sk = td.
- Enc: $k := H(r), r \leftarrow D_I$, output $\langle c_1 = f_I(r), c_2 = \operatorname{Enc}'_k(m) \rangle$.
- Dec: $r:=f_{\mathsf{td}}^{-1}(c_1)$, k:=H(r), output $\mathsf{Dec}_k'(c_2)$.

ROM中基于TDP的CCA安全

- □ 思路: PubK CCA = TDP + 2 RO (一个用于加密, 一个用于MAC)
- □实现CCA安全的公钥加密方案,可以通过RO来构造一个CPA安全的公钥加密方案,以明文和密文一起作为输入来生成MAC标签。
- □定理:如果\$f\$ 是TDP, \$G,H\$ 是RO, 那么构造是CCA 安全的。
- □解释: 其CCA安全性在于对密文的任何篡改, 都无法通过MAC验证。



Construction 15

- lacksquare Gen: pk=I, $sk=\mathsf{td}$
- Enc: $r \leftarrow D_I$, output $\langle c_1 = f_I(r), c_2 = H(r) \oplus m, c_3 = G(c_2 \| m)
 angle$
- $lackbox{lack} ext{ Dec} : r := f_{\mathsf{td}}^{-1}(c_1), \ m := H(r) \oplus c_2. ext{ If } \ G(c_2 \| m) = c_3 ext{ output } m, \ otherwise ot$

本节小结

	Private Key	Public Key
Secret Key	both parties	receiver
Weakest Attack	Eav	CPA
Probabilistic	CPA/CCA	always
Assumption against CPA	OWF	TDP
Assumption against CCA	OWF	TDP+RO
Efficiency	fast	slow