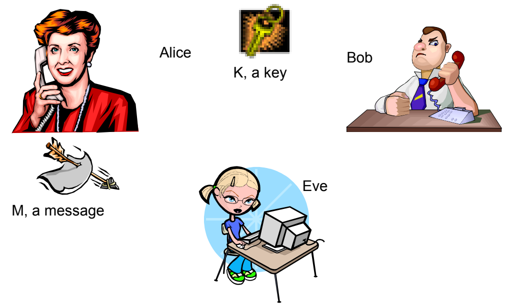
Symmetric Cryptography

15-859I
Spring 2003

Cast of Characters:



Introduction

- Alice wants to send M to Bob
- Eve wants to find out what M is
- Alice and Bob don't want her to.
- Previously, Alice and Bob chose K (together) randomly, so that no one else would know it.
- Can they use one secret (K) to keep another secret (M)?

Encryption Schemes

- Alice and Bob want an *Encryption Scheme*:
- An encryption scheme is a triple $\mathcal{SE} = (G,E,D)$ of Algorithms:
  - $G(1^k)$: generates a key of length k
  - E_K: $P \rightarrow C$ maps an input message space (*plaintexts*) to an output message space (*ciphertexts*)
  - D_K:$C \rightarrow P$ maps an ciphertexts to plaintexts
- For all K, for all $M \in P$, we require that $D_K(E_K(M)) = M$.

Security of Encryption schemes

- What does it mean for $\mathcal{SE}$ to be secure?
- Of course, given $E_K(M)$, Eve should not be able to guess M.
- We will call an attack where Eve recovers M from only $E_K(M)$ a *plaintext recovery* (pr) attack.
- What if M comes from very small subset of P?
- Ideally, we would like Eve to "get no information about M from $E_K(M)$."

This problem is solved unconditionally

- Let $P = \{0,1\}^k$, define $\mathcal{OTP} = (G,E,D)$ as follows:
  - $G(1^k)$ = return $K \leftarrow U_k$.
  - $E_K(M) = K \oplus M$
  - $D_K(C) = K \oplus C$
- It is not hard to see that for M chosen from any distribution on P,
- $H(M|E_K(M)) = H(M)$
- i.e., $E_K(M)$ gives no information about M.

Problem

- We can only use K once, to encrypt |K| bits.
- This means we have to know, beforehand, how many bits we plan to exchange (or an upper bound)
- Then we have to generate that many bits and keep them all secret.
- If we are never in a secure location again, we can never extend the number of bits we can transmit

Solution

- Instead of considering arbitrarily powerful Eve, we constrain Eve to run in polynomial time.
- This suggests that *pseudorandomness* may be useful
- What should it mean for a polytime Eve to learn no information from $E_K(M)$?

Security against Plaintext Recovery

- Suppose Eve plays the following game:
- $Exp^{pr}(Eve)$ =
  - Choose $K \leftarrow U_k$
  - Choose $M \leftarrow U_m$
  - If $Eve^{E_K(\cdot)}(E_K(M)) = M$ output 1 else output 0
- Define $Adv^{pr}(Eve) = Pr[Exp^{pr}(Eve) = 1]$
- Define $Insec^{pr}(\mathcal{SE},t,q,l) = max_{Eve}\{Adv^{pr}(Eve)\}$
- Where we take the max over all Eve running in t operations, making q queries of L bits to $E_k(.)$

Security against Plaintext Recovery

We say $\mathcal{SE}$ is *(t,q,l,ε)-secure against plaintext recovery* if

$$Insec^{pr}(\mathcal{SE}, t,q,l) \leq \varepsilon$$

Asymptotically, $\mathcal{SE}$ is *secure against plaintext recovery* (PR-CPA) if for every polynomial time Eve, $Adv^{pr}(Eve)$ is negligible as a function of k.

Problem with plaintext recovery

- If Eve can reliably recover m/2 bits of the plaintext, she might be satisfied, and $\mathcal{SE}$ would still be secure against plaintext recovery.
- Need a stronger definition, which is equivalent to the information-theoretic notion of not being able to learn a single bit about the plaintext.

Indistinguishability under chosen plaintext attack

Define the oracle $LR_K(b,.,.)$ as follows:

$LR(b,m_0,m_1)$ =

If $|m_0| \neq |m_1|$, return “”

Else return $E_K(m_b)$

Suppose Eve is allowed to choose m_0,m_1. Then given $LR_K(b,.,.)$ for randomly chosen b, she has one bit of uncertainty about $D_K(LR_K(b,m_0,m_1))$.

Indistinguishability under chosen plaintext attack

In a *chosen plaintext attack*, Eve plays this game:

$Exp^{cpa}(b,Eve)$ =

Choose $K \leftarrow U_k$

Return $Eve^{LR_K(b,.,.)}(1^k)$.

Define the advantage of Eve, $Adv^{cpa}(Eve)$, by

$Pr[Exp^{cpa}(1,Eve) = 1] – Pr[Exp^{cpa}(0,Eve) = 1]$

And $Insec^{cpa}(\mathcal{SE}, t,q,l) = max_{Eve}\{Adv^{cpa}(Eve)\}$

Indistinguishability under chosen plaintext attack

$\mathcal{SE}$ is called *(t,q,l,ε)-indistuingishable under chosen plaintext attack* if $Insec^{cpa}(\mathcal{SE},t,q,l) \leq \varepsilon$

It is called indistinguishable under chosen plaintext attack (IND-CPA) if for every polynomial-time Eve, $Adv^{cpa}(Eve)$ is negligible in k.

IND-CPA is stronger than PR-CPA

- Suppose we are given an Eve such that $Adv^{pr}(Eve)$ is non-negligible. Then we will construct an IND-CPA adversary A which has

  $Adv^{cpa}(A) \geq Adv^{pr}(Eve) – 1/2^m$

- This means that if we prove that $\mathcal{SE}$ is IND-CPA then it is also PR-CPA.

IND-CPA is stronger than PR-CPA

- A works as follows:
  - Randomly choose $M_0, M_1 \leftarrow U_m$.
  - Compute $C = LR_K(M_0,M_1)$
  - Run Eve(C), responding to oracle queries X with $LR_K(X,X)$.
  - Let M = output of Eve(C).
  - If $(M = M_1)$, output 1, else output 0.
- Then: if b = 1, $Pr[A^{LR}(1^k) = 1] = Adv^{pr}(Eve)$
- If b = 0, $Pr[A^{LR}(1^k) = 1] \leq 1/2^m$ (M_1 is independent of Eve's view)

IND-CPA is stronger than PR-CPA

- So $Adv^{cpa}(A) \geq Adv^{pr}(Eve) – 1/2^m$
- Giving $Insec^{pr}(\mathcal{SE},t,q,l) \leq Insec(\mathcal{SE},t,q+1,l+m) + 2^{-m}$
- But in general it is much smaller…

Example where PR-CPA is much weaker than IND-CPA

- Suppose P_k is a strong pseudorandom permutation family on $\{0,1\}^k$. Let the message space be $\{0,1\}^k$.
- Define the scheme $\mathcal{ECB}$ = (G,E,D) as follows:
  - $G(1^k)$ = choose $K \leftarrow U_k$
  - $E_K(M) = P_K(M)$
  - $D_K(C) = P_K^{-1}(C)$.
- Claim: $Insec^{pr}(\mathcal{ECB},t,q,l) \leq Insec^{prp}(P,t,q) + q2^{-k}$
- Yet $Insec^{cpa}(\mathcal{ECB},O(k),2,2k) = 1$

IND-CPA encryption: CTR

- Let $F_K : \{0,1\}^L \to \{0,1\}^l$ be a collection of pseudorandom functions.
- Define the *stateful* encryption scheme $\mathcal{CTR}$ as follows:
  - $G(1^k)$ = Choose $K \leftarrow U_k$
  - $E_K(m_0,m_1,\ldots,m_l)$ =
    - Let $c_i = F_K(j+i) \oplus m_i$
    - update $j = j + l$
    - return $c_0,c_1,\ldots,c_l$.
  - $D_K(c_0,c_1,\ldots,c_l) = E_K(c_0,c_1,\ldots,c_l)$

IND-CPA security of $\mathcal{CTR}$

- Claim: Given any Eve which makes at most $q < 2^L$ queries of at most $\mu < l2^L$ bits, we can design a PRF Adversary A with

  $\mathrm{Adv}^{prf}(A) = ½\, \mathrm{Adv}^{cpa}(\mathrm{Eve})$.
- This gives us

  $\mathrm{Insec}^{cpa}(\mathcal{CTR},t,q,\mu) \le 2\mathrm{Insec}^{prf}(F,t,\mu/l)$

  So if F is a secure PRF than $\mathcal{CTR}$ is IND-CPA

Proof of claim

- Given Eve, we define the PRF adversary A as follows:

$A^g(1^k)$ =

Choose $b \leftarrow U_1$.

Run Eve, responding to query $m_0,m_1,\ldots,m_l$ with $g(j)\oplus m_0, g(j+1)\oplus m_1,\ldots,g(j+l)\oplus m_l$, and updating j appropriately.

If Eve outputs b, output 1, else output 0.

Proof of CTR security

- What is $\mathrm{Adv}^{prf}(A)$?
- First, notice that $\Pr[A^{\mathcal{F}(L,l)} = 1] = ½$
  - If g is a random function, then there is no correlation between the bit b and the responses to Eve's queries
- Claim: $\Pr[A^{F_K}=1] = ½ + ½\, \mathrm{Adv}^{cpa}(\mathrm{Eve})$
  - $\Pr[A^F=1|b=0] = \Pr[\mathrm{Eve}^{LR(0,\ldots)} = 0]$
  - $\Pr[A^F=1|b=1] = \Pr[\mathrm{Eve}^{LR(1,\ldots)} = 1]$
  - So $\Pr[A^F=1] = ½(\Pr[\mathrm{Eve}^{LR(0,\ldots)} = 0] + \Pr[\mathrm{Eve}^{LR(1,\ldots)} = 1])$
  - $= ½((1-\Pr[\mathrm{Eve}^{LR(0,\ldots)}=1]) + \Pr[\mathrm{Eve}^{LR(1,\ldots)}=1])$
  - $= ½ + ½\, \mathrm{Adv}^{cpa}(\mathrm{Eve})$

Randomized (stateless) CTR

- Define the scheme $\mathcal{R}\text{-}\mathcal{CTR}$ as follows:
  - $G(1^k)$ = Choose $K \leftarrow U_k$.
  - $E_K(m_0,m_1,\ldots,m_l)$ =

    Choose $r \leftarrow U_L$

    Set $c_i = F_K(r+i)\oplus m_i$

    Return $r,c_0,c_1,\ldots,c_l$
  - $D_K(r,c_0,c_1,\ldots,c_l)$ =

    Set $m_i = F_K(r+i)\oplus c_i$

    Return $m_0,m_1,\ldots,m_l$.

$\mathcal{R}\text{-}\mathcal{CTR}$ is IND-CPA

- Theorem:

  $\mathrm{Insec}^{cpa}(\mathcal{R}\text{-}\mathcal{CTR},t,q,\mu) \le$

  $2\mathrm{Insec}^{prf}(F,t,\mu/l) + \mu q/l2^L$.

- Proof: Given an adversary Eve, define the prf adversary A as before. It still holds that when A is given a pseudorandom oracle, it outputs 1 with probability $½ + ½\, \mathrm{Adv}^{cpa}(\mathrm{Eve})$.

$\mathcal{R}$-$\mathcal{CTR}$ is IND-CPA

- It remains to bound the probability that A outputs 1 given a random function
  - If no input to the random function is repeated, then Pr[A outputs 1] = ½, as in previous argument.
  - If some input is repeated, A outputs 1 with probability at most 1. Call this event (a repeated input to the random function) COL.
  - So $\Pr[A^f=1] \leq ½ + \Pr[COL]$

Claim: $\Pr[COL] < q(\mu/l)2^{-L}$.

- Notice that there are at most (μ/l) inputs to the random function.
- Let n_i = the number of inputs to f as a result of query i.
- Suppose up to query i-1 there have been no repeated inputs to f.
- What is the probability of a collision on query i?
- We get a collision with the j^{th} query if $r_j - n_i < r_i < r_j+n_j+1$, ie, with probability $n_i+n_j/2^L$

Pr[COL]

- Thus the probability of collision on the i^{th} query is at most

  $((i-1)n_i + n_1 + n_2 + \ldots + n_{i-1})/2^L$
- So the probability of a collision on any query is at most

  $q(\mu/l)2^{-L}$, as claimed.