Cryptography

Lecture 03

Recap

Consider the following Encryption Schemes:

1. Shift Cipher: Crackable. Keyspace has only 26 elements.
2. Affine Cipher: Crackable. Keyspace has only 312 elements.
3. Vig Cipher: Crackable by repeats and letter freqs.
4. General Sub: Crackable by letter freqs.
5. Matrix Cipher: Crackable if know (Enc,Dec)-pairs.
6. One-Time Pad: Uncrackable!
7. ElGamal: Uncrackable if we make hardness assumptions.
8. RSA: Uncrackable if we make hardness assumptions.

Recap

Consider the following Encryption Schemes:

1. Shift Cipher: Crackable. Keyspace has only 26 elements.
2. Affine Cipher: Crackable. Keyspace has only 312 elements.
3. Vig Cipher: Crackable by repeats and letter freqs.
4. General Sub: Crackable by letter freqs.
5. Matrix Cipher: Crackable if know (Enc,Dec)-pairs.
6. One-Time Pad: Uncrackable!
7. ElGamal: Uncrackable if we make hardness assumptions.
8. RSA: Uncrackable if we make hardness assumptions.

All of the above are true.

Recap

Consider the following Encryption Schemes:

1. Shift Cipher: Crackable. Keyspace has only 26 elements.
2. Affine Cipher: Crackable. Keyspace has only 312 elements.
3. Vig Cipher: Crackable by repeats and letter freqs.
4. General Sub: Crackable by letter freqs.
5. Matrix Cipher: Crackable if know (Enc,Dec)-pairs.
6. One-Time Pad: Uncrackable!
7. ElGamal: Uncrackable if we make hardness assumptions.
8. RSA: Uncrackable if we make hardness assumptions.

All of the above are true.

All of the above are not rigorous!

Recap

Consider the following Encryption Schemes:

1. Shift Cipher: Crackable. Keyspace has only 26 elements.
2. Affine Cipher: Crackable. Keyspace has only 312 elements.
3. Vig Cipher: Crackable by repeats and letter freqs.
4. General Sub: Crackable by letter freqs.
5. Matrix Cipher: Crackable if know (Enc,Dec)-pairs.
6. One-Time Pad: Uncrackable!
7. ElGamal: Uncrackable if we make hardness assumptions.
8. RSA: Uncrackable if we make hardness assumptions.

All of the above are true.
All of the above are not rigorous!
We make them . . . more rigorous

Assumptions

- With few exceptions, cryptography currently requires *computational assumptions*
  - Question: If P$\neq$NP was proven then would we still need to make hardness assumptions? Discuss.

Assumptions

- With few exceptions, cryptography currently requires *computational assumptions*
  - Question: If P≠NP was proven then would we still need to make hardness assumptions? Discuss.

    P≠NP not good enough!
    Factoring and Discrete Log are not NP-complete and are thought to not be NP-complete. Possible that:

    1. $\mathrm{SAT} \notin$P
    2. Factoring is in P.

Assumptions

- With few exceptions, cryptography currently requires *computational assumptions*
  - Question: If P≠NP was proven then would we still need to make hardness assumptions? Discuss.

    P≠NP not good enough!
    Factoring and Discrete Log are not NP-complete and are thought to not be NP-complete. Possible that:

    1. $\mathrm{SAT} \notin$P
    2. Factoring is in P.

- Principle: Need assumptions to be explicit

Importance of clear assumptions

- Allow researchers to (attempt to) *validate* assumptions by studying them
- Allow meaningful *comparison* between schemes based on different assumptions
- Useful to understand minimal assumptions needed
- Practical implications if assumptions are wrong
- Enable proofs of security

Proofs of Security/Limitations

Proofs give an iron-clad guarantee of security
. . . relative to the definition and assumptions!
Provably secure schemes can be broken!

1. If the definition does not correspond to the real-world threat model
2. i.e. if attacker can go “outside the security mode”
3. If the assumption is invalid
4. If the implementation is flawed

All four of these happen in the real world.

Examples

1. Outside the Box: I'm from IT and I'm here to help.
2. Outside the Box: Timing attacks. To quote Wikipedia:

   > *A timing attack in when the attacker compromises a cryptosystem by analyzing the time taken to execute protocols.*

   Example: In RSA the amount of time it takes to decrypt gives a rough idea of the size of the primes involved, cutting down search space.
3. Bad Implementations of Diffie-Helman
   3.1 Pick p or g to small.
   3.2 Pick a or b two small (they are random– how to prevent?)
   3.3 Use same p, g for a year. Eve has a year to build DL tables.
4. Look up the story of the Maginot Line, an immense wall that France build to deter a German Invasion. It didn't work.

Nevertheless. . .

- This does not detract from the importance of having formal definitions in place
- This does not detract from the importance of proofs of security

Defining secure encryption

Crypto definitions (generally)

- Security guarantee/goal
  - What we want to achieve and/or what we want to prevent the attacker from achieving
- Threat model
  - What (real-world) capabilities the attacker is assumed to have

Threat models for encryption

- Ciphertext-only attack (CTA). As name indicates, Eve only has access to the ciphertext. Eve can crack Shift, Affine, Vig, Gen. Matrix might be an open problem.

- Known-plaintext attack (KPA). Eve has access to previous plaintexts and what the ciphertext was. Matrix can be cracked this way if text is long enough.

- Chosen-plaintext attack (CPA). Eve can fool Alice into encoding a particular plaintext.

- Chosen ciphertext attack (CCA). Eve can fool Bob into telling her what a particular ciphertext decodes to.

Goal of secure encryption?

- How would you define what it means for encryption scheme (Gen, Enc, Dec) over message space $\mathcal{M}$ to be secure?
    - Against a (single) ciphertext-only attack

Secure encryption?

- “Impossible for the attacker to learn the key”
  - The key is a *means to an end*, not the end itself
  - Necessary (to some extent) but not sufficient
  - Easy to design an encryption scheme that hides the key completely, but is insecure
  - Can design schemes where most of the key is leaked, but the scheme is still secure

Secure encryption?

- "Impossible for the attacker to learn the plaintext from the ciphertext"
  - What if the attacker learns 90% of the plaintext?

Secure encryption?

- "Impossible for the attacker to learn any character of the plaintext from the ciphertext"
    - What if the attacker is able to learn (other) partial information about the plaintext?
        - e.g. salary is greater than $75K
    - What if the attacker guesses a character correctly?

Perfect Secrecy

Perfect secrecy

- “Regardless of any *prior* information the attacker has about the plaintext, the ciphertext should leak no *additional* information about the plaintext”
  - The right notion!
  - How to formalize?

Probability review

- *Random variable (r.v.):* variable that takes on (discrete) values with certain probabilities
- Probability distribution for a r.v. specifies the probabilities with which the variable takes on each possible value
  - Each probability must be between 0 and 1
  - The probabilities must sum to 1

Probability review

- *Event:* a particular occurrence in some experiment
  - $\Pr[E]$: probability of event E
- Conditional probability: probability that one event occurs, *given that some other event occurred*
  - $\Pr[A|B] = \frac{\Pr[A \wedge B]}{Pr[B]}$
- To RV's X, Y are *independent* if for all x, y: $\Pr[X = x|Y = y] = \Pr[X = x]$

  Important: X, Y independent if Knowing that $Y = y$ does not help you figure out if $X = x$.
  Discuss: Why will this notion be important for defininig perfect security?

Probability distributions

Normally a RV returns a number. We allow it to be a message. Below are all the messages I could send, with the prob that I send them (unrelated to whether they are true).

1. Today 456 went well. $\Pr = \frac{1}{2}$.
2. Today 456 went badly. $\Pr = \frac{1}{100}$.
3. All 456 students submitted HW Monday. $\Pr = \frac{1}{100}$.
4. I proved a new result in crypto. $\Pr = \frac{1}{50}$.
5. I proved a new result about The Muffin Problem. $\Pr = \frac{9}{25}$.
6. I saw a student in office hours. $\Pr = \frac{1}{10}$

Note:

Should we assume that Eve knows this distribution? Discuss.

Probability distributions

Normally a RV returns a number. We allow it to be a message. Below are all the messages I could send, with the prob that I send them (unrelated to whether they are true).

1. Today 456 went well. $\Pr = \frac{1}{2}$.
2. Today 456 went badly. $\Pr = \frac{1}{100}$.
3. All 456 students submitted HW Monday. $\Pr = \frac{1}{100}$.
4. I proved a new result in crypto. $\Pr = \frac{1}{50}$.
5. I proved a new result about The Muffin Problem. $\Pr = \frac{9}{25}$.
6. I saw a student in office hours. $\Pr = \frac{1}{10}$

Note:

Should we assume that Eve knows this distribution? Discuss. YES. Recall that we assume Eve knows distribution of letters in English.

Recall

- A *private-key encryption scheme* is defined by a message space $\mathcal{M}$ and algorithms (Gen, Enc, Dec):
  - *Gen* (key generation algorithm) generates k
  - *Enc* (encryption algorithm): takes key k and message $m \in \mathcal{M}$ as input; outputs ciphertext c

    $$c \leftarrow Enc_k(m)$$

  - *Dec* (decryption algorithm): takes key k and ciphertext c as input; outputs m

    $$m \leftarrow Dec_k(c)$$

Notation

- $\mathcal{K}$ (key space) — set of all possible keys
- $\mathcal{M}$ (message space) — set of all possible messages
- $\mathcal{C}$ (ciphertext space) — set of all possible ciphertexts

Distribution on Keys

- Let K be the random variable denoting the key
  - K ranges over $\mathcal{K}$
- Fix some encryption scheme (Gen, Enc, Dec)
  - Gen defines a probability distribution for K:

$$\Pr[K = k] = \Pr[\text{Gen outputs key k}]$$

Usually Uniform.

Message and Key Independent

- Random variables M and K are *independent*
  - i.e., the message that a party sends does not depend on the key used to encrypt that message

Distribution on Ciphertext

- Fix some encryption scheme (Gen, Enc, Dec) and some distribution for M
- Consider the following (randomized) experiment:
    1. Choose a message m, according to the given distribution
    2. Generate a key k using Gen
    3. Compute $c \leftarrow Enc_k(m)$
- This defines a distribution on the ciphertext!
- Let C be a random variable denoting the value of the ciphertext in this experiment

Example 1

- Consider the shift cipher
  - So for all k $\in \{0, \ldots, 25\}$, $\Pr[K = k] = \frac{1}{26}$
- Say $\Pr[M = a] = 0.7$, $\Pr[M = z] = 0.3$. So the message is ONLY 1 character.
- What is $\Pr[C = b]$? Discuss?

Example 1

- Consider the shift cipher
    - So for all k $\in \{0, \ldots, 25\}$, $\Pr[K = k] = \frac{1}{26}$
- Say $\Pr[M = a] = 0.7$, $\Pr[M = z] = 0.3$. So the message is ONLY 1 character.
- What is $\Pr[C = b]$? Discuss?
    - Either $M = a$ and $K = 1$, or $M = z$ and $K = 2$

$$\begin{aligned}\Pr[C = b] &= \Pr[M = a] \cdot \Pr[K = 1] + \Pr[M = z] \cdot \Pr[K = 2] \\ &= 0.7 \cdot \frac{1}{26} + 0.3 \cdot \frac{1}{26} = \frac{1}{26}\end{aligned}$$

$\frac{1}{26}$? Hmmm? What if we had diff prob of messages.
If replace 0.7 with p and 0.3 with $1 - p$ do we still get $\frac{1}{26}$?
Discuss.

Example 2

Consider the following distribution. Shift Cipher being used.
$\Pr[M = ab] = \frac{1}{3}$, $\Pr[M = dg] = \frac{2}{3}$

$\Pr[C = ef] =$ Discuss

Example 2

Consider the following distribution. Shift Cipher being used.
$\Pr[M = ab] = \frac{1}{3}$, $\Pr[M = dg] = \frac{2}{3}$

$\Pr[C = ef] =$ Discuss

$\Pr[C = ef | M = ab] \cdot \Pr[M = ab] + \Pr[C = ef | M = dg] \cdot \Pr[M = dg]$

$$= \frac{1}{26} \cdot \frac{1}{3} + 0 \cdot \frac{2}{3} = \frac{1}{78}$$

Example 2

Consider the following distribution. Shift Cipher being used.
$\Pr[M = ab] = \frac{1}{3}$, $\Pr[M = dg] = \frac{2}{3}$

$\Pr[C = ef] =$ Discuss

$\Pr[C = ef \,|\, M = ab] \cdot \Pr[M = ab] + \Pr[C = ef \,|\, M = dg] \cdot \Pr[M = dg]$

$$= \frac{1}{26} \cdot \frac{1}{3} + 0 \cdot \frac{2}{3} = \frac{1}{78}$$

What is $\Pr[C = ei]$? Discuss

Example 2

Consider the following distribution. Shift Cipher being used.
$\Pr[M = ab] = \frac{1}{3}$, $\Pr[M = dg] = \frac{2}{3}$

$\Pr[C = ef] =$ Discuss

$\Pr[C = ef|M = ab] \cdot \Pr[M = ab] + \Pr[C = ef|M = dg] \cdot \Pr[M = dg]$

$$= \frac{1}{26} \cdot \frac{1}{3} + 0 \cdot \frac{2}{3} = \frac{1}{78}$$

What is $\Pr[C = ei]$? Discuss

$\Pr[C = ei|M = ab] \cdot \Pr[M = ab] + \Pr[C = ei|M = dg] \cdot \Pr[M = dg]$

$$= 0 \cdot \frac{1}{3} + 0 \cdot \frac{2}{3} = 0$$

Example 2- Discussion

We had

$$\Pr[M = ab] = \frac{1}{3}, \Pr[M = dg] = \frac{2}{3}$$

Example 2- Discussion

We had

$$\Pr[M = ab] = \frac{1}{3}, \Pr[M = dg] = \frac{2}{3}$$

If we had

$$\Pr[M = ab] = \frac{1}{2}, \Pr[M = dg] = \frac{1}{2}$$

Would the cipher be more secure? Discuss.

Example 2- Discussion

We had

$$\Pr[M = ab] = \frac{1}{3}, \Pr[M = dg] = \frac{2}{3}$$

If we had

$$\Pr[M = ab] = \frac{1}{2}, \Pr[M = dg] = \frac{1}{2}$$

Would the cipher be more secure? Discuss.

No.

Example 2- Discussion

We had

$$\Pr[M = ab] = \frac{1}{3}, \Pr[M = dg] = \frac{2}{3}$$

If we had

$$\Pr[M = ab] = \frac{1}{2}, \Pr[M = dg] = \frac{1}{2}$$

Would the cipher be more secure? Discuss.

No.

Before seeing the ciphertext Eve knew that

$$\Pr[M = ab] = \frac{1}{3}, \Pr[M = dg] = \frac{2}{3}$$

We want that after seeing the ciphertext she knows that

Example 2- Discussion

We had

$$\Pr[M = ab] = \frac{1}{3}, \Pr[M = dg] = \frac{2}{3}$$

If we had

$$\Pr[M = ab] = \frac{1}{2}, \Pr[M = dg] = \frac{1}{2}$$

Would the cipher be more secure? Discuss.

No.

Before seeing the ciphertext Eve knew that

$$\Pr[M = ab] = \frac{1}{3}, \Pr[M = dg] = \frac{2}{3}$$

We want that after seeing the ciphertext she knows that

$$\Pr[M = ab] = \frac{1}{3}, \Pr[M = dg] = \frac{2}{3}$$

Example 2- Discussion

We had

$$\Pr[M = ab] = \frac{1}{3},\ \Pr[M = dg] = \frac{2}{3}$$

If we had

$$\Pr[M = ab] = \frac{1}{2},\ \Pr[M = dg] = \frac{1}{2}$$

Would the cipher be more secure? Discuss.

No.

Before seeing the ciphertext Eve knew that

$$\Pr[M = ab] = \frac{1}{3},\ \Pr[M = dg] = \frac{2}{3}$$

We want that after seeing the ciphertext she knows that

$$\Pr[M = ab] = \frac{1}{3},\ \Pr[M = dg] = \frac{2}{3}$$

If so then Eve has not learned NOTHING from ciphertext!

Perfect secrecy

Informal: Let m be a message.
Before Eve sees the ciphertext she knows $\Pr(M = m)$.
After Eve sees the ciphertext we want her to

not gain any knowledge whatsoever

Formal: Encryption scheme (Gen, Enc, Dec) with message space $\mathcal{M}$ and ciphertext space $\mathcal{C}$ is *perfectly secret* if for every distribution over $\mathcal{M}$, every $m \in \mathcal{M}$, and every $c \in \mathcal{C}$ with $\Pr[C = c] > 0$, it holds that

$$\Pr[M = m | C = c] = \Pr[M = m]$$

The distribution of M does not change conditioned on observing he ciphertext.

Be Impressed!

In Mathematics often getting the right definition is the hard part!

That we have a way of formally defining Perfect Secrecy is very impressive!

This definition is one of the things that marks the line between Classical and Modern Cryptography.

Shift Cipher

Does the Shift Cipher have perfect secrecy? Vote

1. YES
2. NO
3. OTHER

Shift Cipher

Does the Shift Cipher have perfect secrecy? Vote

1. YES
2. NO
3. OTHER

OTHER- the question is ill defined. Need to know the distribution of messages.

1-letter Shift Cipher IS Perfectly Secure

Assume $(\forall x)[\Pr(M = x) = \frac{1}{26}]$.

Need to calculate $\Pr[M = a | C = d]$.

Need $\Pr[M = a \wedge C = d] = \Pr[M = a] \times \Pr[k = 3] = \frac{1}{26} \times \frac{1}{26}$

$\Pr[C = d] = \Pr[M = a]\Pr[k = 3] + \cdots + \Pr[M = z]\Pr[k = 4]$
$= 26 \times \frac{1}{26} \times \frac{1}{26} = \frac{1}{26}$.

So
$\Pr[M = a | C = d] = \frac{\Pr[M=a \wedge C=d]}{\Pr[C=d]} = (\frac{1}{26})^2 / \frac{1}{26} = \frac{1}{26} = \Pr[M = a]$

What I did for $M = a$ and $C = d$ works for any pair.

1-letter Shift Cipher IS Perfectly Secure

Assume $(\forall x)[\Pr(M = x) = \frac{1}{26}]$.

Need to calculate $\Pr[M = a|C = d]$.

Need $\Pr[M = a \wedge C = d] = \Pr[M = a] \times \Pr[k = 3] = \frac{1}{26} \times \frac{1}{26}$

$\Pr[C = d] = \Pr[M = a]\Pr[k = 3] + \cdots + \Pr[M = z]\Pr[k = 4]$
$= 26 \times \frac{1}{26} \times \frac{1}{26} = \frac{1}{26}$.

So
$\Pr[M = a|C = d] = \frac{\Pr[M=a \wedge C=d]}{\Pr[C=d]} = (\frac{1}{26})^2/\frac{1}{26} = \frac{1}{26} = \Pr[M = a]$

What I did for $M = a$ and $C = d$ works for any pair.
Is this enough to prove that 1-letter Shift is Perfectly Secure?
Discuss.

1-letter Shift Cipher IS Perfectly Secure

Assume $(\forall x)[\Pr(M = x) = \frac{1}{26}]$.

Need to calculate $\Pr[M = a | C = d]$.

Need $\Pr[M = a \wedge C = d] = \Pr[M = a] \times \Pr[k = 3] = \frac{1}{26} \times \frac{1}{26}$

$\Pr[C = d] = \Pr[M = a]\Pr[k = 3] + \cdots + \Pr[M = z]\Pr[k = 4]$
$= 26 \times \frac{1}{26} \times \frac{1}{26} = \frac{1}{26}$.

So
$\Pr[M = a | C = d] = \frac{\Pr[M=a \wedge C=d]}{\Pr[C=d]} = (\frac{1}{26})^2 / \frac{1}{26} = \frac{1}{26} = \Pr[M = a]$

What I did for $M = a$ and $C = d$ works for any pair.
Is this enough to prove that 1-letter Shift is Perfectly Secure?
Discuss.
NO- we need this to work for ANY distribution. May be HW.

2-letter Shift Cipher NOT Perfectly Secure

Consider the shift cipher with 2-letter messages and distribution

1. $\Pr[M = ab] = \frac{1}{2}$,
2. $\Pr[M = ac] = \frac{1}{2}$

Take $m = ab$ and $c = pr$

$\Pr[M = ab | C = pr] = 0 \neq \Pr[M = ab]$

Conclusion

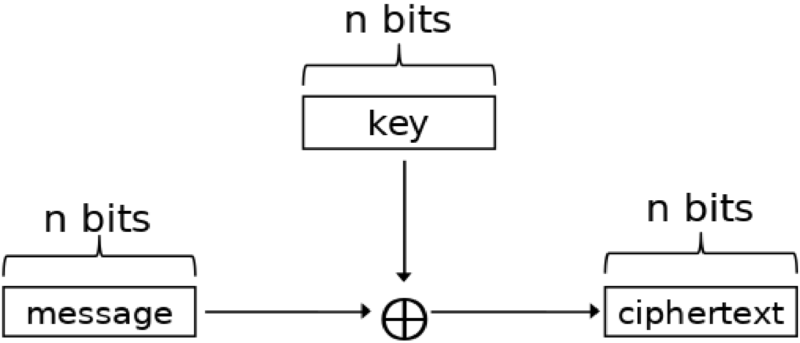
- The shift cipher is not perfectly secret!
  - At least not for 2-character messages
- How to construct a perfectly secret scheme?

Recall the One-time pad

- Let $\mathcal{M} = \{0,1\}^n$
- *Gen*: choose a uniform key $k \in \{0,1\}^n$
- $Enc_k(m) = k \oplus m$
- $Dec_k(c) = k \oplus c$
- Correctness:

$$\begin{aligned} Dec_k(Enc_k(m)) &= k \oplus (k \oplus m) \\ &= (k \oplus k) \oplus m \\ &= m \end{aligned}$$

One-time pad


n bits
key
n bits
message
n bits
ciphertext