CSE 594 : Modern Cryptography 26 Jan, 2017

Lecture 1: Shannon and Perfect Secrecy

Instructor: Omkant Pandey *Scribe: Him Kalyan Bordoloi, Arian Akhavan Niaki*

1 Symmetric Ciphers

A symmetric cipher consists of the following elements:

1. KG a method for generating random keys k.
2. Enc an encryption algorithm, where Enc encrypts a message m using a secret key k and generate ciphertext c. This is formally shown as:

$$Enc(k, m) \to c$$

3. Dec a decryption algorithm, where Dec should work correctly for every m in the message space M given the ciphertext and the key. This is formally shown as:

$$\forall k, \forall m : Dec(k, Enc(k, m)) = m.$$

Notation: M, K and C are the message space, key space and the ciphertext space and they contain the set of all messages m, all keys k and all ciphertexts c respectively.

1.1 Security of a Cipher

1. Hide the key: hiding the key does not mean hiding the message, for example in Caesar Cipher ATTACK = BUUBDL and DEFEND = EFGFOE. Therefore, the cipher can be broken by checking patterns and without having the key.
2. Hide the message: hiding all possible functions of the message is impossible because some characteristic about the message may be known. For example, a message in English may always start with "Hello".
3. Hide everything that is not known: The ciphertext should not give any **new** information about the message to the adversary.

1.2 Hide everything that is not known

1.2.1 Shannon's Secrecy

The approach of "Hiding Everything that is not known" is represented mathematically as follows

- D is the distribution of messages over the message space M. D consists of the probabilities of all messages m in M.
- $c = Enc(m, k)$ is the cipher text produced by the encryption algorithm where
  - m is the message being encrypted

- k is the key chosen randomly
- Enc induces some additional randomness
- C is the distribution of cipher-text

- For to adversary to not gain any additional knowledge from the encrypted message, his knowledge of D must not increase after observing C

i.e. distribution D and $D|C$ must be identical

Definition 1 *A cipher (M, K,KG, Enc, Dec) is Shannon secure w.r.t a distribution D over M if for all m_1 P M and for all c*

$$Pr[m \leftarrow D : m = m'] = Pr[k \leftarrow KG, m \leftarrow D : m = m' | Enc(m, k) = c]$$

It is Shannon secure if it is Shannon secure w.r.t. all distributions D over M.

1.2.2 Perfect Secrecy

For every pair of messages $m_1 \in M$ and $m_2 \in M$, The distribution of cipher-texts for m_1, $C_1 = \{k \leftarrow KG, output\ Enc(m_1, k)\}$ and for m_2, $C_2 = \{k \leftarrow KG, output\ Enc(m_2, k)\}$ are identical

i.e. The distributions C_1 and C_2 must be identical for every pair of m_1, m_2

Definition 2 *Scheme (M, K,KG, Enc, Dec) is perfectly secure for every pair of messages m_1, m_2 in M and for all c,*

$$Pr[k \leftarrow KG\ :\ Enc(m_1, k) = c] = Pr[k \leftarrow KG\ :\ Enc(m_2, k) = c]$$

Theorem 1 *Equivalence Theorem A private-key encryption scheme is perfectly secure if and only if it is Shannon secure.*

Proof. In order to prove the Equivalence Theorem we need to prove the following

Perfect Secrecy => Shannon Secrecy
And
Shannon Secrecy => Perfect Secrecy

Part 1: Perfect Secrecy => Shannon Secrecy

Given: $\forall (m_1, m_2) \in M \times M$ and every $c \in C$

$$Pr[Enc_k(m_1) = c] = Pr[Enc_k(m_2) = c]$$

Show: for every D over M $m' \in M$, and $c \in C$

$$Pr_k, m[m = m' | Enc_k(m) = c] = Pr_m[m = m']$$

$$
\begin{aligned}
L.H.S &= Pr_{k,m}[m = m'|Enc_k(m) = c] \\
&= \frac{Pr_{k,m}[m = m' \cap Enc_k(m) = c]}{Pr_{k,m}[Enc_k(m) = c]} \\
&= \frac{Pr_{k,m}[m = m' \cap Enc_k(m') = c]}{Pr_{k,m}[Enc_k(m) = c]} \because m = m' \ in \ numerator
\end{aligned}
$$

$$\because Pr[m = m'] \ is \ independent \ of \ k \ and \ Pr[Enc_k(m') = c] \ is \ independent \ of \ m$$

$$
\begin{aligned}
&= \frac{Pr_m[m = m'].Pr_k[Enc_k(m') = c]}{Pr_{k,m}[Enc_k(m) = c]} \\
&= \frac{Pr_m[m = m']}{Pr_{k,m}[Enc_k(m) = c]} \times \frac{Pr_k[Enc_k(m') = c]}{Pr_{k,m}[Enc_k(m) = c]} \\
&= Pr_{m,k}[m = m'|Enc_k(m) = c] \times \frac{Pr_k[Enc_k(m') = c]}{Pr_{k,m}[Enc_k(m) = c]} \\
&= R.H.S \times \frac{Pr_k[Enc_k(m') = c]}{Pr_{k,m}[Enc_k(m) = c]}
\end{aligned}
$$

Now we need to prove that

$$\frac{Pr_k[Enc_k(m') = c]}{Pr_{k,m}[Enc_k(m) = c]} = 1$$

The probability that we get a cipher-text c from any message m is the sum of the probabilities of each test in the message set M leading to c on encryption using Enc

$$\therefore Pr_{k,m}[Enc_k(m) = c] = \sum_{m'' \in M} Pr_m[m = m'']Pr_k[Enc_k(m'') = c]$$

$$\because \ probability \ of \ getting \ cipher-text \ c \ is \ equal \ for \ every \ message \ in \ M$$

$$
\begin{aligned}
&= \sum_{m'' \in M} Pr_m[m = m'']Pr_k[Enc_k(m'') = c] \\
&= Pr_k[Enc_k(m') = c] \sum_{m'' \in M} Pr_m[m = m''] \\
&= Pr_k[Enc_k(m') = c] \times 1
\end{aligned}
$$

$$\therefore \frac{Pr_k[Enc_k(m') = c]}{Pr_{k,m}[Enc_k(m) = c]} = 1$$

Part 2: Shannon Secrecy => Perfect Secrecy

Given: $\forall(m_1, m_2) \in M \times M$ and $\forall c$

Show: $Pr_k[Enc_k(m_1) = c] = Pr_k[Enc_k(m_2) = c]$

We will only look at uniform distribution for this proof
Let D be the uniform distribution over m_1, m_2 so that:

$$Pr_m[m = m_1] = Pr_m[m = m_2] = \frac{1}{2}$$

Since we are assuming this to be Shannon secure w.r.t D

$$Pr_{k,m}[m = m_1 | Enc_k(m) = c] = Pr_m[m = m_1] \ and$$
$$Pr_{k,m}[m = m_2 | Enc_k(m) = c] = Pr_m[m = m_2]$$

$$\therefore Pr_{k,m}[m = m_1 | Enc_k(m) = c] = Pr_{k,m}[m = m_2 | Enc_k(m) = c]$$

$$\begin{aligned} L.H.S &= Pr_{k,m}[m = m_1 | Enc_k(m) = c] \\ &= \frac{Pr_{k,m}[m = m_1 \cap Enc_k(m) = c]}{Pr_{k,m}[Enc_k(m) = c]} \\ &= \frac{Pr_{k,m}[m = m_1 \cap Enc_k(m_1) = c]}{Pr_{k,m}[Enc_k(m) = c]} \because m = m_1 \ in \ numerator \end{aligned}$$

$$\because Pr[m = m_1] \ is \ independent \ of \ k \ and \ Pr[Enc_k(m_1) = c] \ is \ independent \ of \ m$$

$$\begin{aligned} &= \frac{Pr_m[m = m_1].Pr_k[Enc_k(m_1) = c]}{Pr_{k,m}[Enc_k(m) = c]} \\ &= \frac{\frac{1}{2}.Pr_k[Enc_k(m_1) = c]}{Pr_{k,m}[Enc_k(m) = c]} \end{aligned}$$

$$\begin{aligned} &Similarly \\ R.H.S &= Pr_{k,m}[m = m_2 | Enc_k(m) = c] \\ &= \frac{\frac{1}{2}.Pr_k[Enc_k(m_2) = c]}{Pr_{k,m}[Enc_k(m) = c]} \end{aligned}$$

$$\because L.H.S = R.H.S$$
$$\frac{\frac{1}{2}.Pr_k[Enc_k(m_1) = c]}{Pr_{k,m}[Enc_k(m) = c]} = \frac{\frac{1}{2}.Pr_k[Enc_k(m_2) = c]}{Pr_{k,m}[Enc_k(m) = c]}$$

Now cancel $\frac{\frac{1}{2}}{Pr_{k,m}[Enc_k(m)=c]}$ from both sides to get:

$$Pr_k[Enc_k(m_1) = c] = Pr_k[Enc_k(m_2) = c]$$

■

Remark 1 *As noted in the class, it is not necessary to assume that m_1 and m_2 occur with equal probability $\frac{1}{2}$. We can work with any D over the message space M such that support of D is equal to M. To see this, observe that "LHS" is also equal* $\Pr[m = m_1]$ *so we can divide by* $\Pr[m = m_1]$ *(which is not 0) to get that* $\Pr_k[Enc_k(m_1) = c] = \Pr_{k,m \leftarrow D}[Enc_k(m) = c]$*. Do the same to the term in "RHS" to get the same equation for m_2 and observe that they come out to be equal.*

2 One Time Pad

- n is an integer which is equal to the length of the plaintext message.
- $M := \{0,1\}^n$ is the Message space which is an n bit binary string.
- $K := \{0,1\}^n$ is the Key space. Therefore the key is as long as the message.

Definition 3 *OTP Algorithm:*

- *KG sample a key k uniformly at random. $k \leftarrow \{0,1\}^n$*
- *$Enc(m, k) = c$ is a bit-by-bit XOR*
  if $m = m_1m_2...m_n$ and $k = k_1k_2...k_n$ the output ciphertext $c = c_1c_2...c_n$ is generated by $c_i = m_i \oplus k_i$.
- *$Dec(c, k) = m$ is a bit-by-bit XOR as well*
  where $m_i = c_i \oplus k_i$ for ever i.
- *the key must have the following conditions:*
  - *The key can be only used once.*
  - *It must be sampled uniformly every time.*
  - *The key must be the same length as the message. This will be a problem when encrypting large amounts of data. (Ex: 80 GB hard drive)*

Theorem 2 *Perfect Security of OTP One Time Pad is a perfectly secure symmetric cipher encryption scheme.*

Proof. Perfect secrecy: for a fix $m \in \{0,1\}^n$ and $c \in \{0,1\}^n$.
We know that $Enc(m, k) = m \oplus k$ therefore:

$$Pr_k[Enc_k(m) = c] = Pr[m \oplus k = c]$$

By applying $\oplus m$ to both sides of $m \oplus k = c$:

$$Pr[m \oplus k = c] = Pr[k = m \oplus c] = 2^{-n}$$

For all c that are not an n bit binary string ($\forall c \notin \{0,1\}^n$):

$$Pr_k[Enc_k(m) = c] = 0$$

$\Rightarrow \forall (m_1, m_2) \in \{0,1\}^{n \times n}$ and $\forall c$:

$$Pr_k[Enc_k(m_1) = c] = Pr_k[Enc_k(m_2) = c]$$

■

Theorem 3 *Shannon's Theorem For every perfectly secure cipher (Enc, Dec) with message space M and key space K, it holds that $|K| \geq |M|$.*

Remark 2 *Note that message length n and, key length l are $n = lg|M|$ and, $l = lg|K|$ respectively. Taking log on both sides, we get $l \geq n$, i.e., keys must be as long as the messages for perfect secrecy.*

Proof. If we assume the contrary $|K| \leq |M|$ and fix any message m_0 and any key k_0.
Let: $c_0 = Enc(m_0, k_0)$

$$\Rightarrow Pr_k[Enc(m_0, k) = c_0] > 0.$$

If we decrypt c_0 with each key one by one we get a set of messages defined as below:

$$S = \{Dec(c_0, k) : k \in |K|\}$$

We know that $|S| \leq |K|$ and from our assumption $|K| < |M|$, therefore we have:

$$|S| < |M|$$

This means that there exists a message $m_1 \in |M|$ such that $m_1 \notin |S|$. If we encrypt m_1 with key $k \in |K|$:

$$\forall k \in |K| : Enc(m_1, k) \neq c_0.$$

$$\Rightarrow Pr_k[Enc(m_1, k) = c_0] = 0.$$

Therefore, there exists m_0, m_1, and c_0 such that:

$$Pr_k[Enc(m_0, k) = c_0] \neq Pr_k[Enc(m_1, k) = c_0].$$

The statement above contradicts perfect secrecy. ■