One-Time Pad

CS 5158/6058 Data Security and Privacy

Spring 2018

Instructor: Boyang Wang

Space & Random Variable

- Key space $\mathcal{K}$
- *K* be a Random Variable for keys
- Pr[*K*=k]: the probability of a key is k

- Message space $\mathcal{M}$, ciphertext space $\mathcal{C}$
- *M* is a RV for messages, *C* is a RV for ciphertexts
- Pr[*M*=m]: the probability of a message is m
- Pr[*C*=c]: the probability of a ciphertext is c

Random Variable

- Example: message space $\mathcal{M}$ = {**`a`**, **`b`**, **`c`**},
  - **`a`** (0.5), **`b`** (0.4), **`c`** (0.1)
  - *M* is RV for the message space

$$\Pr[M = a] = 0.5$$
$$\Pr[M = b] = 0.4$$
$$\Pr[M = c] = 0.1$$

$$\sum_{m \in \mathcal{M}} \Pr[M = m] = 1$$

- RV *K* and RV *M* are <u>independent</u>

$$\Pr[(K = k) \cap (M = m)] = \Pr[K = k] \cdot \Pr[M = m]$$

Random Variable

- Practice: Shift Cipher
  - $\mathcal{M}$={**`a`**, **`b`**, **`c`**}, **`a`** (0.5), **`b`** (0.3), **`c`** (0.2)
  - $\mathcal{K}$={0, 1, …, 25}, Pr[K=k]=1/26, each
  - What is the probability of ciphertext is **`F`**?
- Three cases: 1) M=**`a`** and K=5; 2) M=**`b`** and K=4; 3) M=**`c`** and K=3

$$\Pr[C = F] = 0.5 \cdot \frac{1}{26} + 0.3 \cdot \frac{1}{26} + 0.2 \cdot \frac{1}{26} = \frac{1}{26}$$

Random Variable

- <u>Practice</u>: Shift Cipher
  - $\mathcal{M}$={**`a`**, **`b`**, **`c`**}, **`a`** (0.5), **`b`** (0.3), **`c`** (0.2)
  - $\mathcal{K}$={0, 1, …, 25}, Pr[*K*=k]=1/26, each
  - What is the probability of ciphertext is **`F`**?

- There are 3 messages in message space, so 3 different messages that could lead to **`F`**

$$\begin{aligned}\Pr[C=F] &= \Pr[C=F|M=a]+\Pr[C=F|M=b]\\ &+ \Pr[C=F|M=c]\end{aligned}$$

Random Variable

- Case 1: given M = **a**, then ciphertext C = **F**

—> Enc_K(**a**) = **F** (in Shift Cipher)

—> (M = **a**) AND (K = 5)

$$\begin{aligned}\Pr[C = F | M = a] &= \Pr[\mathsf{Enc}_K(a) = F] \\ &= \Pr[(M = a) \cap (K = 5)]\end{aligned}$$

- M and K are independent

$$\Pr[(M = a) \cap (K = 5)] = \Pr[M = a] \times \Pr[K = 5] = 0.5 \times \frac{1}{26}$$

- Try Case 2(M = **b**); Case 3(M = **c**) yourself

Random Variable

- Case 2: given M `= b`, then ciphertext C `= F`

$$
\begin{aligned}
\Pr[C = F | M = b] &= \Pr[\mathsf{Enc}_K(b) = F] \\
&= \Pr[(M = b) \cap (K = 4)] \\
&= \Pr[M = b] \cdot \Pr[K = 4] \\
&= 0.3 \cdot \frac{1}{26}
\end{aligned}
$$

Random Variable

- Case 3: given M `= c`, then ciphertext C `= F`

$$
\begin{aligned}
\Pr[C = F | M = c] &= \Pr[\mathsf{Enc}_K(c) = F] \\
&= \Pr[(M = c) \cap (K = 3)] \\
&= \Pr[M = c] \cdot \Pr[K = 3] \\
&= 0.2 \cdot \frac{1}{26}
\end{aligned}
$$

Random Variable

- The overall probability that ciphertext C `= F`

$$
\begin{aligned}
\Pr[C = F] &= \Pr[C = F|M = a] + \Pr[C = F|M = b] \\
&+ \Pr[C = F|M = c] \\
&= 0.5 \cdot \frac{1}{26} + 0.3 \cdot \frac{1}{26} + 0.2 \cdot \frac{1}{26} \\
&= \frac{1}{26}
\end{aligned}
$$

Random Variable

- Practice: Shift Cipher
  - $\mathcal{M}$={**ab**, **aa**, **cc**}, **ab** (0.5), **aa** (0.3), **cc** (0.2)
  - $\mathcal{K}$={0, 1, …, 25}, Pr[K=k]=1/26, each
  - What is the probability of ciphertext is **EE**?

- Hint: There are 3 messages in message space, so 3 different messages that could lead to **EE**

$$\begin{aligned}\Pr[C=EE] &= \Pr[C=EE|M=ab]+\Pr[C=EE|M=aa] \\ &+ \Pr[C=EE|M=cc]\end{aligned}$$

Random Variable

- Case 1: given M = **`ab`**, then ciphertext C = **`EE`**

—> Enc_K(**`ab`**) = **`EE`** (in Shift Cipher)

—> there is no k can do this!!!!

$$\Pr[C = EE | M = ab] = \Pr[\mathsf{Enc}_K(ab) = EE] = 0$$

- Case 2: given M = **`aa`**, then ciphertext C = **`EE`**

$$\begin{aligned}
\Pr[C = EE | M = aa] &= \Pr[\mathsf{Enc}_K(aa) = EE] \\
&= \Pr[(M = aa) \cap (K = 4)] \\
&= \Pr[M = aa] \cdot \Pr[K = 4] = 0.3 \cdot \frac{1}{26}
\end{aligned}$$

Random Variable

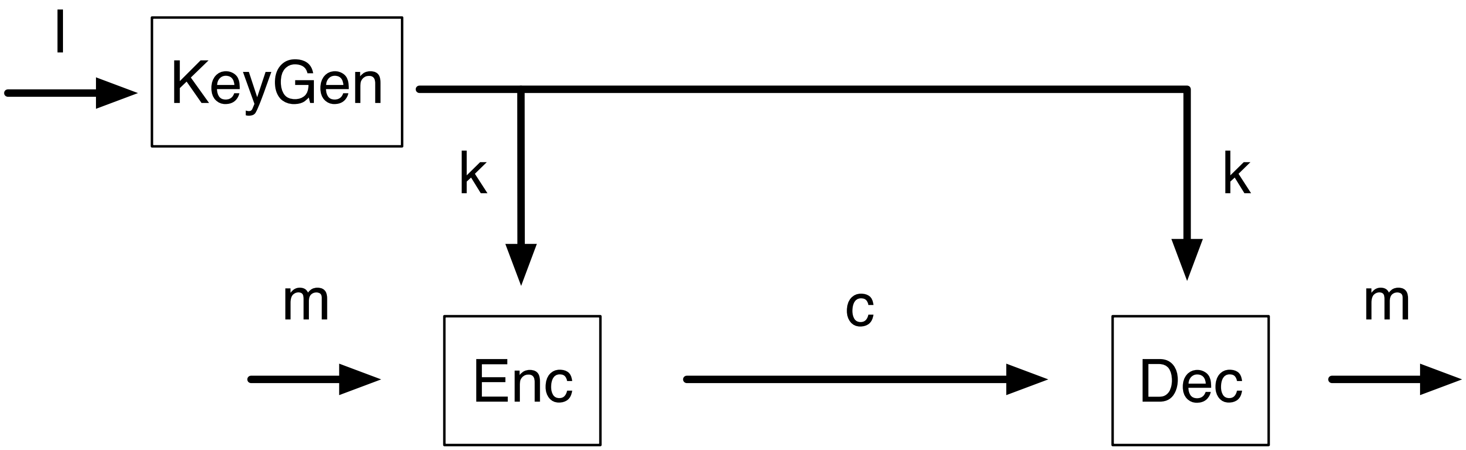
- Case 3: given M = **cc**, then ciphertext C = **EE**

$$\begin{aligned} \Pr[C = EE | M = cc] &= \Pr[\mathsf{Enc}_K(cc) = EE] \\ &= \Pr[(M = cc) \cap (K = 2)] \\ &= \Pr[M = cc] \cdot \Pr[K = 2] = 0.2 \cdot \frac{1}{26} \end{aligned}$$

- The overall probability that ciphertext C = **EE**

$$\begin{aligned} \Pr[C = EE] &= \Pr[C = EE | M = ab] + \Pr[C = EE | M = aa] \\ &+ \Pr[C = EE | M = cc] \\ &= 0 + 0.3 \cdot \frac{1}{26} + 0.2 \cdot \frac{1}{26} = \frac{1}{52} \end{aligned}$$

One-Time Pad



- KeyGen(1^l): $k \leftarrow \{0,1\}^l$, return k
- Enc$_k(m)$: $c \leftarrow k \oplus m$, return c
- Dec$_k(c)$: $m \leftarrow k \oplus c$, return m

1) k is as long as m
2) Use each k once

XOR Truth Table

0 XOR 0 = 0

0 XOR 1 = 1

1 XOR 0 = 1

1 XOR 1 = 0

Correctness

- For any k and any m

$$
\begin{aligned}
\mathsf{Dec}_k(\mathsf{Enc}_k(m)) &= k \oplus (k \oplus m) \\
&= (k \oplus k) \oplus m \\
&= \{0\}^{\lambda} \oplus m \\
&= m
\end{aligned}
$$

- Example:

```
m: 01110
k: 11010
c: 10100
```

1) k is as long as m
2) Use each k once

XOR Truth Table
0 XOR 0 = 0
0 XOR 1 = 1
1 XOR 0 = 1
1 XOR 1 = 0

Security

- One-Time Pad is perfectly secret.
    - Informally, an adversary absolutely learns nothing about the plaintext that was encrypted

- Assumption about an adversary:
    - Knows distribution over message space $\mathcal{M}$
    - Knows Enc and Dec algorithm
    - Can eavesdrop, unlimited computation power
    - Does not know key k

Perfect Secrecy

- Observing ciphertext c has <u>no effect</u> on an adversary's knowledge regarding message m

Theorem An encryption scheme $(\mathsf{KeyGen}, \mathsf{Enc}, \mathsf{Dec})$ with message space $\mathcal{M}$ is **perfectly secret** if for every probability distribution over $\mathcal{M}$, every message $m \in \mathcal{M}$, and every ciphertext $c \in \mathcal{C}$:

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

Perfect Secrecy

- The distribution of the ciphertext does not depend on distribution of the plaintext

Lemma An encryption scheme (KeyGen, Enc, Dec) with message space $\mathcal{M}$ is **perfectly secret** if for every $m, m' \in \mathcal{M}$ and every $c \in \mathcal{C}$

$$\Pr[\mathsf{Enc}_K(m) = c] = \Pr[\mathsf{Enc}_K(m') = c]$$

where the probabilities are over choice of K and any randomness of Enc.

Security Game

- Ciphertexts of m_0, m_1 are <u>indistinguishable</u>.

Given $\Pi = (\mathsf{KeyGen}, \mathsf{Enc}, \mathsf{Dec})$, **security game** $\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi}$:

1. Adversary $\mathcal{A}$ outputs m_0, $m_1 \in \mathcal{M}$ with $|m_0| = |m_1|$
2. Challenger flips a (fair) coin $b \in \{0, 1\}$, compute $c_b \leftarrow \mathsf{Enc}_k(m_b)$, where $k \leftarrow \mathsf{KeyGen}(1^l)$, and return c_b to $\mathcal{A}$
3. $\mathcal{A}$ guesses a bit b'
4. Output 1 if $b' = b$, otherwise 0; $\mathcal{A}$ wins if it is 1

Security Game

- Random guess is 1/2, but cannot do better

Def. Encryption scheme $\Pi = (\mathsf{KeyGen}, \mathsf{Enc}, \mathsf{Dec})$, with message space $\mathcal{M}$ is **perfectly indistinguishable** if for every $\mathcal{A}$ it holds that

$$\Pr[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1] = \frac{1}{2}$$

- Adversary does not have <u>any advantage</u>

$$\mathrm{Adv}^{\mathsf{eav}}_{\mathcal{A},\Pi} = \left| \Pr[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1] - \frac{1}{2} \right| = 0$$

Vigenere Cipher

- Vigenere Cipher is <u>not</u> perfectly indistinguishable
- Example: message $\mathcal{M}$ = {`aa`, `ab`}, key is a string of 1 or 2 (key length is uniformly chosen)

1. Adversary $\mathcal{A}$ chooses $m_0 = aa$ and $m_1 = ab$
2. Challenger flips a coin, obtains b and $c_b \leftarrow \mathsf{Enc}_k(m_b)$
3. Given $c_b = c_{b1}c_{b2}$, Adversary $\mathcal{A}$ guesses $b' = 0$ if $c_{b1} = c_{b2}$; otherwise $b' = 1$
4. $\mathcal{A}$ wins iff $b' = b$

Analysis on Vigenere Cipher

- Adversary A wins if b'=0|b=0 or b'=1|b=1
- Random guess 1/2, prove A can win <u>greater than1/2</u>

$$\begin{aligned}
& \Pr[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\mathsf{VC}} = 1] \\
= \quad & \Pr[b=0] \cdot \Pr[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\mathsf{VC}} = 1 | b=0] \\
& + \quad \Pr[b=1] \cdot \Pr[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\mathsf{VC}} = 1 | b=1] \\
= \quad & \frac{1}{2} \cdot \Pr[b'=0|b=0] + \frac{1}{2} \cdot \Pr[b'=1|b=1]
\end{aligned}$$

Analysis on Vigenere Cipher

- $b'=0|b=0$ ($c_b=c_{b1}c_{b2}, c_{b1}=c_{b2}|m_0=$**`aa`**) has two cases:
  - Key length is 1 (k_1), any k_1 in {0,1, …, 25}
    - E.g., **`aa`** + k_1k_1 = **`xx`**
  - Key length is 2 (k_1k_2), and <u>k_1,k_2 are same</u>
    - E.g., **`aa`** + $k_1k_2 = k_1k_1$ = **`xx`**

$$\Pr[b' = 0|b = 0] = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{26} \approx 0.52$$

Analysis on Vigenere Cipher

- $b'=1|b=1(c_b=c_{b1}c_{b2}, c_{b1}!=c_{b2}|m_1=$**`ab`**$)$ has two cases:
  - Key length is 1 (k_1), any k_1 in {0,1, …, 25}
    - E.g., **`ab`** + k_1k_1 = **`XY`**
  - Key length is 2 (k_1k_2), and <u>k_2 is not k_1-1</u>
    - E.g, **`ab`** + k_1k_2 = $k_1(k_1\text{-}1)$ = **`XX`**

- <u>Practice:</u> Pr[b'=1|b=1] = ?

$$\Pr[b'=1|b=1]=\frac{1}{2}+\frac{1}{2}\cdot(1-\frac{1}{26})\approx 0.98$$

Analysis on Vigenere Cipher

- Put everything together

$$
\begin{aligned}
& \Pr[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\mathsf{VC}} = 1] \\
= \quad & \frac{1}{2} \cdot \Pr[b' = 0 | b = 0] + \frac{1}{2} \cdot \Pr[b' = 1 | b = 1] \\
\approx \quad & \frac{1}{2} \cdot 0.52 + \frac{1}{2} \cdot 0.98 \\
= \quad & 0.75 \quad > \quad \frac{1}{2}
\end{aligned}
$$

Additional Reading

Chapter 2, *Introduction to Modern Cryptography, Drs. J. Katz and Y. Lindell, 2nd edition*