Semantic Security

Rohit Musti

CUNY - Hunter College

February 9, 2022

Overview

• Alice and Bob share a secret key k.

Overview

- Alice and Bob share a secret key k.
- Alice wants to send a message m to Bob through an untrusted medium.

Overview

- Alice and Bob share a secret key k.
- Alice wants to send a message m to Bob through an untrusted medium.
- Mechanism: Shannon Ciphers.

• Father of Information Theory

- Father of Information Theory
- Friends with Alan Turing (they actually compared notes on Turing's famous paper the Universal Turing machine)

- Father of Information Theory
- Friends with Alan Turing (they actually compared notes on Turing's famous paper the Universal Turing machine)
- Estimated the complexity of chess

- Father of Information Theory
- Friends with Alan Turing (they actually compared notes on Turing's famous paper the Universal Turing machine)
- Estimated the complexity of chess
- First to formalize the notion of security

• **Definition:** a Shannon Cipher is a pair of functions $\mathcal{E} = (E, D)$

- **Definition:** a Shannon Cipher is a pair of functions $\mathcal{E} = (E, D)$
- The *E* function takes in two arguments: a key *k* and a message *m*, and produces a ciphertext *c*

$$c=E(k,m)$$

- **Definition:** a Shannon Cipher is a pair of functions $\mathcal{E} = (E, D)$
- The E function takes in two arguments: a key k and a message m, and produces a ciphertext c

$$c=E(k,m)$$

 The D function takes in two arguments: a key k and a ciphertext c, and produces a plaintext m

- **Definition:** a Shannon Cipher is a pair of functions $\mathcal{E} = (E, D)$
- The *E* function takes in two arguments: a key *k* and a message *m*, and produces a ciphertext *c*

$$c = E(k, m)$$

- The D function takes in two arguments: a key k and a ciphertext c, and produces a plaintext m
- While E maybe random, D must be deterministic

$$m = D(k, c)$$



Rohit Musti Semantic Security

• Decryption must undo Encryption

Decryption must undo Encryption

$$m = D(k, E(k, m))$$

• More formally: let K be the key space, M be the message space, and \mathcal{C} be the ciphertext space.

Decryption must undo Encryption

$$m = D(k, E(k, m))$$

• More formally: let $\mathcal K$ be the key space, M be the message space, and $\mathcal C$ be the ciphertext space.

$$E: \mathcal{K} \times \mathcal{M} \to \mathcal{C}$$

• Decryption must undo Encryption

$$m = D(k, E(k, m))$$

• More formally: let $\mathcal K$ be the key space, M be the message space, and $\mathcal C$ be the ciphertext space.

$$E: \mathcal{K} \times \mathcal{M} \to \mathcal{C}$$

$$D: \mathcal{K} \times \mathcal{C} \to \mathcal{M}$$

Decryption must undo Encryption

$$m = D(k, E(k, m))$$

• More formally: let $\mathcal K$ be the key space, M be the message space, and $\mathcal C$ be the ciphertext space.

$$E: \mathcal{K} \times \mathcal{M} \to \mathcal{C}$$

$$D: \mathcal{K} \times \mathcal{C} \to \mathcal{M}$$

ullet is defined over $(\mathcal{K},\mathcal{M},\mathcal{C})$



Rohit Musti Semi

• A one time pad is a Shannon Cipher s.t.

$$\mathcal{K} \coloneqq \mathcal{M} \coloneqq \mathcal{C} \coloneqq \{0,1\}^L$$

• A one time pad is a Shannon Cipher s.t.

$$\mathcal{K}\coloneqq\mathcal{M}\coloneqq\mathcal{C}\coloneqq\{0,1\}^L$$

• Encryption for a key $k \in \{0,1\}^L$ and a message $m \in \{0,1\}^L$:

$$E(k, m) := k \oplus m$$

• A one time pad is a Shannon Cipher s.t.

$$\mathcal{K}\coloneqq\mathcal{M}\coloneqq\mathcal{C}\coloneqq\{0,1\}^L$$

• Encryption for a key $k \in \{0,1\}^L$ and a message $m \in \{0,1\}^L$:

$$E(k, m) := k \oplus m$$

• Decryption for a key $k \in \{0,1\}^L$ and a ciphertext $c \in \{0,1\}^L$:

$$D(k,c) := k \oplus c$$

Rohit Musti

• A one time pad is a Shannon Cipher s.t.

$$\mathcal{K} \coloneqq \mathcal{M} \coloneqq \mathcal{C} \coloneqq \{0,1\}^L$$

• Encryption for a key $k \in \{0,1\}^L$ and a message $m \in \{0,1\}^L$:

$$E(k, m) := k \oplus m$$

• Decryption for a key $k \in \{0,1\}^L$ and a ciphertext $c \in \{0,1\}^L$:

$$D(k,c) := k \oplus c$$

 Future HW: Write out a mathmatical justification that the one time pad meets the earlier correctness requirement (Decryption undoes Encryption)

 \bullet key assumption: adversaries know the encryption mechanism and distribution of ${\mathcal K}$

- \bullet key assumption: adversaries know the encryption mechanism and distribution of ${\mathcal K}$
- If Alice encrypts a message m with a key k and an adversary obtains the ciphertext c, the key k needs to be hard to guess

- \bullet key assumption: adversaries know the encryption mechanism and distribution of ${\cal K}$
- If Alice encrypts a message m with a key k and an adversary obtains the ciphertext c, the key k needs to be hard to guess (if k is easy to guess, then the adversary will just guess until they discover k).

- \bullet key assumption: adversaries know the encryption mechanism and distribution of ${\mathcal K}$
- If Alice encrypts a message m with a key k and an adversary obtains the ciphertext c, the key k needs to be hard to guess (if k is easy to guess, then the adversary will just guess until they discover k).
- ullet Thus, k must be chosen uniformly & random from a large keyspace ${\cal K}$

 key assumption: adversaries may have some knowledge of the message being encrypted.

- key assumption: adversaries may have some knowledge of the message being encrypted.
- Supposed $m_0 = meet_at_lunch_time$ and $m_1 = meet_at_snack_time$,

- key assumption: adversaries may have some knowledge of the message being encrypted.
- Supposed $m_0 = meet_at_lunch_time$ and $m_1 = meet_at_snack_time$, our adversary has a 50% chance of guessing correctly.

- key assumption: adversaries may have some knowledge of the message being encrypted.
- Supposed $m_0 = meet_at_lunch_time$ and $m_1 = meet_at_snack_time$, our adversary has a 50% chance of guessing correctly.
- A secure cipher text should not increase this probability of guessing correctly.

- key assumption: adversaries may have some knowledge of the message being encrypted.
- Supposed $m_0 = meet_at_lunch_time$ and $m_1 = meet_at_snack_time$, our adversary has a 50% chance of guessing correctly.
- A secure cipher text should not increase this probability of guessing correctly.
- Suppose there are 90 keys k_0 s.t. $E(k_0, m_0) = c$ and 10 keys k_1 s.t. $E(k_1, m_1) = c$

- key assumption: adversaries may have some knowledge of the message being encrypted.
- Supposed $m_0 = meet_at_lunch_time$ and $m_1 = meet_at_snack_time$, our adversary has a 50% chance of guessing correctly.
- A secure cipher text should not increase this probability of guessing correctly.
- Suppose there are 90 keys k_0 s.t. $E(k_0, m_0) = c$ and 10 keys k_1 s.t. $E(k_1, m_1) = c$, the probability that a given message c is m_0 is 90/(90+10) = 90%, increasing our adversaries odds of guessing correctly.

8 / 16

Perfect Security

 Assumption: a key is drawn uniformly and randomly from a large key space: k.

Perfect Security

- Assumption: a key is drawn uniformly and randomly from a large key space: k.
- Let $\mathcal{E} = (E, D)$ be a Shannon Cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$

Perfect Security

- Assumption: a key is drawn uniformly and randomly from a large key space: k.
- Let $\mathcal{E} = (E, D)$ be a Shannon Cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$
- If for all $m_0, m_1 \in \mathcal{M}$, for all $c \in \mathcal{C}$, and $\mathbf{k} \xleftarrow{R} \mathcal{K}$ we have

$$Pr[E(\mathbf{k}, m_0) = c] = Pr[E(\mathbf{k}, m_1) = c]$$

One Time Pad: Perfect Security

A result of our definition of perfect security is

$$|\{k \stackrel{R}{\leftarrow} \mathcal{K} : E(k,m) = c\}| = N_c$$

One Time Pad: Perfect Security

• A result of our definition of perfect security is

$$|\{k \stackrel{R}{\leftarrow} \mathcal{K} : E(k,m) = c\}| = N_c$$

ullet For any message $m \in \{0,1\}^l$ and cipher text $c \in \{0,1\}^l$,

$$k \oplus m = c$$

Rohit Musti

One Time Pad: Perfect Security

• A result of our definition of perfect security is

$$|\{k \stackrel{R}{\leftarrow} \mathcal{K} : E(k,m) = c\}| = N_c$$

ullet For any message $m \in \{0,1\}^l$ and cipher text $c \in \{0,1\}^l$,

$$k \oplus m = c$$

thus, $N_c = 1$, satisfiying the above result.

Rohit Musti Semantic Security

• Let $\mathcal{E} = (E, D)$ be a Shannon Cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$, if this cipher is perfectly secure, then $|\mathcal{K}| \ge |\mathcal{M}|$

- Let $\mathcal{E} = (E, D)$ be a Shannon Cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$, if this cipher is perfectly secure, then $|\mathcal{K}| \ge |\mathcal{M}|$
- It is impractical to have keys that are at least as large as the size of the message we are transmitting.

- Let $\mathcal{E} = (E, D)$ be a Shannon Cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$, if this cipher is perfectly secure, then $|\mathcal{K}| \ge |\mathcal{M}|$
- It is impractical to have keys that are at least as large as the size of the message we are transmitting.
- Really, what we are concerned about are real world, computationally bounded adversaries.

- Let $\mathcal{E} = (E, D)$ be a Shannon Cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$, if this cipher is perfectly secure, then $|\mathcal{K}| \ge |\mathcal{M}|$
- It is impractical to have keys that are at least as large as the size of the message we are transmitting.
- Really, what we are concerned about are real world, computationally bounded adversaries.
- We are also interested in efficient algorithms for encrypting and decrypting. By efficient, we mean polynomial functions.

Intuition: the probability that a computationally bounded adversary can learn anything about a message m given its cipher text c is negligible.

Intuition: the probability that a computationally bounded adversary can learn anything about a message m given its cipher text c is negligible. Experiment b

• Let $\mathcal{E} = (E, D)$ be a Cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{A} be computationally bounded

Intuition: the probability that a computationally bounded adversary can learn anything about a message m given its cipher text c is negligible. Experiment b

- Let $\mathcal{E} = (E, D)$ be a Cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{A} be computationally bounded
- ullet \mathcal{A} picks $m_0, m_1 \in \mathcal{M}$ and sends to challenger

Intuition: the probability that a computationally bounded adversary can learn anything about a message m given its cipher text c is negligible. Experiment b

- Let $\mathcal{E} = (E, D)$ be a Cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{A} be computationally bounded
- ullet \mathcal{A} picks $m_0, m_1 \in \mathcal{M}$ and sends to challenger
- Challenger selects $k \stackrel{R}{\leftarrow} \mathcal{K}$

Intuition: the probability that a computationally bounded adversary can learn anything about a message m given its cipher text c is negligible. Experiment b

- Let $\mathcal{E} = (E, D)$ be a Cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{A} be computationally bounded
- ullet \mathcal{A} picks $m_0, m_1 \in \mathcal{M}$ and sends to challenger
- Challenger selects $k \stackrel{R}{\leftarrow} \mathcal{K}$
- Challenger computes $E(k, m_b) := c$ and sends c to adversary

Rohit Musti Semantic Security

Intuition: the probability that a computationally bounded adversary can learn anything about a message m given its cipher text c is negligible. Experiment b

- Let $\mathcal{E} = (E, D)$ be a Cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{A} be computationally bounded
- ullet \mathcal{A} picks $m_0, m_1 \in \mathcal{M}$ and sends to challenger
- Challenger selects $k \stackrel{R}{\leftarrow} \mathcal{K}$
- Challenger computes $E(k, m_b) := c$ and sends c to adversary
- ullet $\mathcal A$ outputs b^* , guessing which experiment was selected by the Challenger

12/16

ullet The Semantic Security Advantage (SSA) of the adversary ${\cal A}$, is

• The Semantic Security Advantage (SSA) of the adversary A, is

$$\mathit{SSA}[\mathcal{A},\mathcal{E}] := |\mathit{Pr}[W_0] - \mathit{Pr}[W_1]|$$

where W_b is the probability that A outputs 1 in experiment b

Rohit Musti

• The Semantic Security Advantage (SSA) of the adversary A, is

$$SSA[\mathcal{A}, \mathcal{E}] := |Pr[W_0] - Pr[W_1]|$$

where W_b is the probability that A outputs 1 in experiment b

• At random, A has a 50% probability of outputting the correct bit, yielding

$$\textit{SSA}[\mathcal{A},\mathcal{E}] \coloneqq |\textit{Pr}[\textit{W}_0] = \textit{Pr}[\textit{W}_1]| = 0$$

ullet The Semantic Security Advantage (SSA) of the adversary ${\cal A}$, is

$$SSA[\mathcal{A}, \mathcal{E}] := |Pr[W_0] - Pr[W_1]|$$

where W_b is the probability that ${\mathcal A}$ outputs 1 in experiment b

 \bullet At random, ${\cal A}$ has a 50% probability of outputting the correct bit, yielding

$$SSA[\mathcal{A}, \mathcal{E}] := |Pr[W_0] = Pr[W_1]| = 0$$

• In order to be semantically secure, the advantage has to be negligible. We can think of negligible as $1+\lambda\approx 1, \lambda=2^{-100}$

13 / 16

Rohit Musti Semantic Security

Semantic Security guarantees that an adversary cannot recover a message from a cipher text.

Semantic Security guarantees that an adversary cannot recover a message from a cipher text. Here is our security game for message recovery.

Semantic Security guarantees that an adversary cannot recover a message from a cipher text. Here is our security game for message recovery.

• Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{A} be computationally bounded

Semantic Security guarantees that an adversary cannot recover a message from a cipher text. Here is our security game for message recovery.

- Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{A} be computationally bounded
- Challenger generates $k \stackrel{R}{\leftarrow} \mathcal{K}$, $m \stackrel{R}{\leftarrow} \mathcal{M}$ and computes E(k,m) = cand send c to the adversary A

Semantic Security guarantees that an adversary cannot recover a message from a cipher text. Here is our security game for message recovery.

- Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{A} be computationally bounded
- Challenger generates $k \stackrel{R}{\leftarrow} \mathcal{K}$, $m \stackrel{R}{\leftarrow} \mathcal{M}$ and computes E(k, m) = c and send c to the adversary \mathcal{A}
- ullet ${\cal A}$ outputes message m^*

Now let's use our adversary ${\cal A}$ to break semantic security, violating our original assumption.

• Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{B} be computationally bounded

- Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{B} be computationally bounded
- ullet \mathcal{B} picks $m_0, m_1 \in \mathcal{M}$ and sends to challenger

- Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{B} be computationally bounded
- ullet \mathcal{B} picks $m_0, m_1 \in \mathcal{M}$ and sends to challenger
- Challenger selects $k \stackrel{R}{\leftarrow} \mathcal{K}$

- Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{B} be computationally bounded
- ullet \mathcal{B} picks $m_0, m_1 \in \mathcal{M}$ and sends to challenger
- Challenger selects $k \stackrel{R}{\leftarrow} \mathcal{K}$
- Challenger computes $E(k, m_b) \coloneqq c$ and sends c to \mathcal{B}

- Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{B} be computationally bounded
- ullet \mathcal{B} picks $m_0, m_1 \in \mathcal{M}$ and sends to challenger
- Challenger selects $k \stackrel{R}{\leftarrow} \mathcal{K}$
- Challenger computes $E(k, m_b) := c$ and sends c to \mathcal{B}
- ullet S sends c to \mathcal{A} , simulating the message recovery game.

- Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{B} be computationally bounded
- \mathcal{B} picks $m_0, m_1 \in \mathcal{M}$ and sends to challenger
- Challenger selects $k \stackrel{R}{\leftarrow} \mathcal{K}$
- Challenger computes $E(k, m_b) := c$ and sends c to \mathcal{B}
- \mathcal{B} sends c to \mathcal{A} , simulating the message recovery game.
- \mathcal{A} returns m^* . If $m^* = m_1$, \mathcal{B} returns 1, else it returns 0.

Now let's use our adversary A to break semantic security, violating our original assumption.

- Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ and let our adversary \mathcal{B} be computationally bounded
- \mathcal{B} picks $m_0, m_1 \in \mathcal{M}$ and sends to challenger
- Challenger selects $k \stackrel{R}{\leftarrow} \mathcal{K}$
- Challenger computes $E(k, m_b) := c$ and sends c to \mathcal{B}
- \mathcal{B} sends c to \mathcal{A} , simulating the message recovery game.
- \mathcal{A} returns m^* . If $m^* = m_1$, \mathcal{B} returns 1, else it returns 0.

Thus, if we can recover the message, we have broken semantic security.

ullet Let $\mathcal{E}=(E,D)$ be a semantically secure cipher defined over $(\mathcal{K},\mathcal{M},\mathcal{C})$

- Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$
- Suppose Alice wants to send information to Bob, but doesn't want Bob to know that the message came from her.

- ullet Let $\mathcal{E}=(\mathcal{E},\mathcal{D})$ be a semantically secure cipher defined over $(\mathcal{K},\mathcal{M},\mathcal{C})$
- Suppose Alice wants to send information to Bob, but doesn't want Bob to know that the message came from her. (example: you want a radiologist to review a patient's scans without revealing the patients identity)

- ullet Let $\mathcal{E}=(E,D)$ be a semantically secure cipher defined over $(\mathcal{K},\mathcal{M},\mathcal{C})$
- Suppose Alice wants to send information to Bob, but doesn't want Bob to know that the message came from her. (example: you want a radiologist to review a patient's scans without revealing the patients identity)
- Alice must first share k_0 with router₀ and k_1 with router₁.

- ullet Let $\mathcal{E}=(E,D)$ be a semantically secure cipher defined over $(\mathcal{K},\mathcal{M},\mathcal{C})$
- Suppose Alice wants to send information to Bob, but doesn't want Bob to know that the message came from her. (example: you want a radiologist to review a patient's scans without revealing the patients identity)
- Alice must first share k₀ with router₀ and k₁ with router₁.
- Then Alice sends $E(k_0, E(k_1, m))$ to router 1.

- Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$
- Suppose Alice wants to send information to Bob, but doesn't want Bob to know that the message came from her. (example: you want a radiologist to review a patient's scans without revealing the patients identity)
- Alice must first share k_0 with router₀ and k_1 with router₁.
- Then Alice sends $E(k_0, E(k_1, m))$ to router 1.
- router₁ decrypts and sends the message $E(k_1, m)$ to router₂ in random order.

- Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$
- Suppose Alice wants to send information to Bob, but doesn't want Bob to know that the message came from her. (example: you want a radiologist to review a patient's scans without revealing the patients identity)
- Alice must first share k_0 with router₀ and k_1 with router₁.
- Then Alice sends $E(k_0, E(k_1, m))$ to router 1.
- router₁ decrypts and sends the message $E(k_1, m)$ to router₂ in random order.
- router decrypts and sends the message m to Bob in random order.

- Let $\mathcal{E} = (E, D)$ be a semantically secure cipher defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$
- Suppose Alice wants to send information to Bob, but doesn't want Bob to know that the message came from her. (example: you want a radiologist to review a patient's scans without revealing the patients identity)
- Alice must first share k_0 with router₀ and k_1 with router₁.
- Then Alice sends $E(k_0, E(k_1, m))$ to router 1.
- router₁ decrypts and sends the message $E(k_1, m)$ to router₂ in random order.
- router decrypts and sends the message m to Bob in random order.

Alice's security is 1/n where n is the number of messages sent along the network.



16 / 16

Semantic Security Rohit Musti