Shannon's Theory, Perfect Secrecy, and the One-Time Pad

Dhananjoy Dey

Indian Institute of Information Technology, Lucknow
ddey@iiitl.ac.in

December 23, 2022

Disclaimers

1

All the pictures used in this presentation are taken from freely available websites.

Disclaimers

1

All the pictures used in this presentation are taken from freely available websites.

2

If there is a reference on a slide all of the information on that slide is attributable to that source whether quotation marks are used or not.

Outline

Outline

Approaches to Evaluating the Security of a Cryptosystem

- **Computational security:** concerns the computational effort required to break a cryptosystem. A system to be **computationally secure** if the *best algorithm* for breaking it requires at least N operations, where N very large number

Approaches to Evaluating the Security of a Cryptosystem

- **Computational security:** concerns the computational effort required to break a cryptosystem. A system to be **computationally secure** if the *best algorithm* for breaking it requires at least N operations, where N very large number $N = 2^{112}$.
- **Provable security:** is to provide evidence of security by means of a reduction. This approach only provides a proof of security relative to some other problem, not an absolute proof of security.
- **Unconditional security:** it cannot be broken, even with infinite computational resources.

Type of Attack on a Cryptosystem

The **attack model** specifies the information available to the adversary when (s)he mounts an attack.

- **Ciphertext-only Attack/Known Ciphertext Attack (KCA):** The opponent possesses a string of ciphertext y.

Type of Attack on a Cryptosystem

The **attack model** specifies the information available to the adversary when (s)he mounts an attack.

- **Ciphertext-only Attack/Known Ciphertext Attack (KCA):** The opponent possesses a string of ciphertext y.
- **Known Plaintext Attack (KPA):** The opponent possesses a string of plaintext, x, and the corresponding ciphertext, y.

Type of Attack on a Cryptosystem

The **attack model** specifies the information available to the adversary when (s)he mounts an attack.

- **Ciphertext-only Attack/Known Ciphertext Attack (KCA):** The opponent possesses a string of ciphertext y.
- **Known Plaintext Attack (KPA):** The opponent possesses a string of plaintext, x, and the corresponding ciphertext, y.
- **Chosen Plaintext Attack (CPA or CPA1):** The opponent can choose a plaintext string, x, and receives the corresponding ciphertext string, y.

Type of Attack on a Cryptosystem

The **attack model** specifies the information available to the adversary when (s)he mounts an attack.

- **Ciphertext-only Attack/Known Ciphertext Attack (KCA):** The opponent possesses a string of ciphertext y.
- **Known Plaintext Attack (KPA):** The opponent possesses a string of plaintext, x, and the corresponding ciphertext, y.
- **Chosen Plaintext Attack (CPA or CPA1):** The opponent can choose a plaintext string, x, and receives the corresponding ciphertext string, y.
- **Chosen Ciphertext Attack (CCA or CCA1):** The opponent can choose a ciphertext string, y, and receives the corresponding plaintext string, x.

Type of Attack on a Cryptosystem

- **Adaptive Chosen Plaintext Attack (ACPA or CPA2):** is a chosen plaintext attack in which the choice of plaintext may depend on the ciphertext received from previous requests.

Type of Attack on a Cryptosystem

- **Adaptive Chosen Plaintext Attack (ACPA or CPA2):** is a chosen plaintext attack in which the choice of plaintext may depend on the ciphertext received from previous requests.
- **Adaptive Chosen Ciphertext Attack (ACCA or CCA2):** is a chosen ciphertext attack where the choice of ciphertext may depend on the plaintext received from previous requests.

Outline

Perfect Secrecy

- **Assumption:** The key K is chosen using some *fixed probability distribution*

Perfect Secrecy

- **Assumption:** The key K is chosen using some *fixed probability distribution* (often a key is chosen at random)
- The key is chosen before the sender knows what the plaintext P will be. Hence, we can assume that the key and the plaintext are *independent random variables*.
- The two probability distributions on $\mathcal{P}$ and $\mathcal{K}$ induce a probability distribution on C.
- $C(K)$ denotes the set of possible ciphertexts if K is the key. Then, for every $y \in C$, we have that

$$\mathbf{Pr}[\mathbf{y} = y] = \sum_{\{K: y \in C(K)\}} \mathbf{Pr}[\mathbf{K} = K]\mathbf{Pr}[\mathbf{x} = d_K(y)].$$

Perfect Secrecy

- The conditional probability

$$\mathbf{Pr}[\mathbf{y} = y|\mathbf{x} = x] = \sum_{\{K: x = d_K(y)\}} \mathbf{Pr}[\mathbf{K} = K].$$

- The probability that x is the plaintext, given that y is the ciphertext

$$\mathbf{Pr}[\mathbf{x} = x|\mathbf{y} = y] = \frac{\mathbf{Pr}[\mathbf{x} = x] \times Pr[\mathbf{y} = y|\mathbf{x} = x]}{\mathbf{Pr}[\mathbf{y} = y]}$$

Example

Example

- Let $\mathcal{P} = \{a, b\}$ with
$$Pr[a] = 1/4,\ Pr[b] = 3/4.$$
- Let $\mathcal{K} = \{K_1, K_2, K_3\}$ with
$$Pr[K_1] = 1/2,\ Pr[K_2] = Pr[K_3] = 1/4.$$
- Let $\mathcal{C} = \{1, 2, 3, 4\}$, and suppose the encryption functions are defined to be
$e_{K_1}(a) = 1,\ e_{K_1}(b) = 2; \quad e_{K_2}(a) = 2,\ e_{K_2}(b) = 3; \quad e_{K_3}(a) = 3,\ e_{K_3}(b) = 4.$

Example

Example

- This cryptosystem can be represented by the following encryption matrix:

	a	b
K_1	1	2
K_2	2	3
K_3	3	4

- Compute the probability distribution on $\mathcal{C}$:

$$Pr[1] = Pr[K_1].Pr[a]$$

Example

Example

- This cryptosystem can be represented by the following encryption matrix:

	a	b
K_1	1	2
K_2	2	3
K_3	3	4

- Compute the probability distribution on $\mathcal{C}$:

$$\begin{aligned} Pr[1] &= Pr[K_1].Pr[a] = \frac{1}{2} \times \frac{1}{4} = \frac{1}{8} \\ Pr[2] &= \end{aligned}$$

Example

Example

- This cryptosystem can be represented by the following encryption matrix:

	a	b
K_1	1	2
K_2	2	3
K_3	3	4

- Compute the probability distribution on $\mathcal{C}$:

$$\begin{aligned}
Pr[1] &= Pr[K_1].Pr[a] = \frac{1}{2} \times \frac{1}{4} = \frac{1}{8} \\
Pr[2] &= Pr[K_1].Pr[b] + Pr[K_2].Pr[a] = \frac{1}{2} \times \frac{3}{4} + \frac{1}{4} \times \frac{1}{4} = \frac{7}{16} \\
Pr[3] &= Pr[K_2].Pr[b] + Pr[K_3].Pr[a] = \frac{1}{4} \times \frac{3}{4} + \frac{1}{4} \times \frac{1}{4} = \frac{1}{4} \\
Pr[4] &= Pr[K_3].Pr[b] = \frac{1}{4} \times \frac{3}{4} = \frac{3}{16}
\end{aligned}$$

Example

Example

- Now, compute the conditional probability distributions on the plaintext

$$Pr[a|1] =$$

Example

Example

- Now, compute the conditional probability distributions on the plaintext

$$
\begin{aligned}
Pr[a|1] &= \frac{Pr[a].Pr[K_1]}{Pr[1]} = 1 & Pr[b|1] &= 0^{b} \\
Pr[a|2] &= \frac{1}{7} & Pr[b|2] &= \frac{6}{7} \\
Pr[a|3] &= \frac{1}{4} & Pr[b|3] &= \frac{3}{4} \\
Pr[a|4] &= 0^{a} & Pr[b|4] &= 1
\end{aligned}
$$

[a] There does not exist any key for which a is mapped to 4

[b] There does not exist any key for which b is mapped to 1

Perfect Secrecy

Definition

*A cryptosystem has **perfect secrecy** if*

$$\mathbf{Pr}[x|y] = \mathbf{Pr}[x]\ \forall x \in \mathcal{P}, y \in \mathcal{C}.$$

Perfect Secrecy

Definition

A cryptosystem has ***perfect secrecy*** *if*

$$\mathbf{Pr}[x|y] = \mathbf{Pr}[x] \;\forall x \in \mathcal{P}, y \in \mathcal{C}.$$

Theorem

Suppose $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ *is a cryptosystem where* $|\mathcal{K}| = |\mathcal{C}| = |\mathcal{P}|$. *Then the cryptosystem provides* ***perfect secrecy*** *iff every key is used with equal probability* $\frac{1}{|\mathcal{K}|}$*, and for every* $x \in \mathcal{P}$ *and every* $y \in \mathcal{C}$*,* $\exists\,!\,K \;: e_K(x) = y$.

One-time Pad

Definition

Let $\mathcal{P} = \mathcal{C} = \mathcal{K} = (\mathbb{Z}_2)^n$ *for* $n \geq 1$. *For* $K \in (\mathbb{Z}_2)^n$ *, define* $e_K(x)$

$$e_K(x) = (x_1 + K_1, \ldots, x_n + K_n) \mod 2,$$

where $x = (x_1, \ldots, x_n)$ *and* $K = (K_1, \ldots, K_n)$*.*

Decryption is identical to encryption. If $y = (y_1, \ldots, y_n)$*, then*

$$d_K(y) = (y_1 + K_1, \ldots, y_n + K_n) \mod 2.$$

Outline

Uncertainly and Information

- Tomorrow, the sun will rise from the East
- The phone will ring before the class is over.
- It will snow in Lucknow by the end of January 2023

Uncertainly and Information

- Tomorrow, the sun will rise from the East
- The phone will ring before the class is over.
- It will snow in Lucknow by the end of January 2023

Note: a high probability event conveys less information than a low probability event.

Uncertainly and Information

- Tomorrow, the sun will rise from the East
- The phone will ring before the class is over.
- It will snow in Lucknow by the end of January 2023

Note: a high probability event conveys less information than a low probability event.

Definition

*The **self information** of the event $X = x_i$ for $1 \leq i \leq n$ is defined as*

$$I(x_i) = \log\left(\frac{1}{P(x_i)}\right) = -\log(P(x_i))$$

Entropy

- Entropy can be thought of as a mathematical measure of information or uncertainty, and is computed as a function of a probability distribution.

Definition

Suppose $\mathbf{X}$ is a discrete random variable. Then, the entropy or average self information of the random variable $\mathbf{X}$ is defined as

$$H(\mathbf{X}) = -\sum_{x \in X} \mathbf{Pr}[x] \log_2 \mathbf{Pr}[x].$$

Properties of Entropy

Theorem

Suppose $\mathbf{X}$ *is a random variable having a probability distribution that takes on the values* $p_1, p_2, \ldots, p_n$,*where* $p_i > 0,\ 1 \leq i \leq n$. *Then* $H(\mathbf{X}) \leq \log_2 n$,

Properties of Entropy

Theorem

Suppose $\mathbf{X}$ *is a random variable having a probability distribution that takes on the values* $p_1, p_2, \ldots, p_n$,*where* $p_i > 0,\ 1 \leq i \leq n$. *Then* $H(\mathbf{X}) \leq \log_2 n$,*with equality iff* $p_i = 1/n,\ 1 \leq i \leq n$.

Theorem

$H(\mathbf{X}, \mathbf{Y}) \leq H(\mathbf{X}) + H(\mathbf{Y})$, *with equality if and only if* $\mathbf{X}$ *and* $\mathbf{Y}$ *are independent random variables.*

Conditional Entropy

Definition

The conditional entropy $H(\mathbf{X}|\mathbf{Y})$ *is defined by the weighted average over all possible values* y*. It is computed as*

$$\begin{aligned} H(\mathbf{X}|\mathbf{Y}) &= \sum_y \mathbf{Pr}[y].H(\mathbf{X}|y) \\ &= -\sum_y \sum_x \mathbf{Pr}[y]\mathbf{Pr}[x|y] \log_2 \mathbf{Pr}[x|y]. \end{aligned}$$

Theorem

$$H(\mathbf{X}, \mathbf{Y}) = H(\mathbf{Y}) + H(\mathbf{X}|\mathbf{Y}).$$

Corollary

$H(\mathbf{X}|\mathbf{Y}) \leq H(\mathbf{X})$*, with equality iff* $\mathbf{X}$ *and* $\mathbf{Y}$ *are independent.*

Spurious Keys

Theorem

Let $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ be a cryptosystem. Then

$$H(\mathbf{K}|\mathbf{C}) = H(\mathbf{K}) + H(\mathbf{P}) - H(\mathbf{C}).$$

Definition

- *Attacker to guess the key from the ciphertext shall guess the key and decrypt the cipher.*
- *He checks whether the plaintext obtained is 'meaningful' English. If not, he rules out the key.*
- *But due to the redundancy of language more than one key will pass this test.*
- *Those keys, apart from the correct key, are called **spurious**.*

Entropy of Plain Text

- H_L: measure of the amount of information per letter of '*meaningful*' strings of plaintext.
- A random string of plaintext formed using English letter has an entropy of $\log_2(26) \approx 4.76$ bits
- A first order entropy of the English text is $H(P) \approx 4.14$ bits
- A second order entropy of the English text is $\frac{H(P^2)}{2} \approx 3.56$ bits
- The entropy of a natural language L denoted by H_L and is defined by

$$H_L = \lim_{n\to\infty} \frac{H(P^n)}{n}$$

Redundancy

Definition

The redundancy of L is defined as

$$R_L = 1 - \frac{H_L}{\log_2 |\mathcal{P}|}$$

Redundancy

Definition

The redundancy of L is defined as

$$R_L = 1 - \frac{H_L}{\log_2 |\mathcal{P}|}$$

- For English Language, $1 \leq H_L \leq 1.5$. Let's take $H_L = 1.25$
- $|\mathcal{P}| = 26$
- $R_L = 0.75$

English Language is 75% redundant

Unicity Distance

Definition

The ***unicity distance*** *of a cryptosystem is defined to be the value of n, denoted by n_0, at which the expected number of spurious keys becomes zero i.e., the average amount of ciphertext required for an opponent to be able to uniquely compute the key, given enough computing time.*

References

C. E. Shannon,
A Mathematical Theory of Communication, Bell Systems Technical Journal, 27 (1948), 623-656.
http://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf

C. E. Shannon,
Communication Theory of Secrecy Systems Bell Systems Technical Journal, 28 (1949), 656-715.
http://netlab.cs.ucla.edu/wiki/files/shannon1949.pdf

D. R. Stinson & M. B. Paterson,
Cryptography – Theory and Practice, CRC, 2019.

The End

Thanks a lot for your attention!