Applied Cryptography and Computer Security

CSE 664 Spring 2017

Lecture 3: Perfect Secrecy, Entropy

Department of Computer Science and Engineering
University at Buffalo

Lecture Outline

- **Last lecture:**
  - classical ciphers
- **This lecture:**
  - elements of probability theory
  - perfect secrecy
  - one-time pad (Vernam's cipher)
  - entropy
  - language redundancy

Lecture Outline

- **Recall how the security of a cryptosystem is shown:**
  - **computational security**
  - **unconditional security**
- **Today we study unconditionally secure systems using probability theory**
  - **given a ciphertext, no information can be learned about the message it encrypts**
  - **ciphers we already learned about can be made unconditionally secure**

One-Time Pad

- **An example of crypto system that achieves unconditional and perfect secrecy is one-time pad (Vernam's cipher)**
  - **given a binary message m of length n**
  - **algorithm Gen produces a random binary key k of length at least n**
  - **to encrypt m with k, compute $\mathsf{Enc}_k(m) = m \oplus k$**
  - **to decrypt c with k, compute $\mathsf{Dec}_k(c) = c \oplus k$**
- **What properties does this cipher have and why is it so good?**

Elementary Probability Theory

- **A discrete random variable X consists of:**
  - **a finite set $\mathcal{X}$ of values**
  - **a probability distribution defined on $\mathcal{X}$**
- **The probability that X takes on the value x is denoted by $\mathsf{Pr}[X = x]$**
- **We must have that**
  - $\mathsf{Pr}[X = x] \geq 0$ **for all** $x \in \mathcal{X}$
  - $\sum_{x \in \mathcal{X}} \mathsf{Pr}[X = x] = 1$
- **Example: dice from homework**
  - **probability distribution is** $\mathsf{Pr}[X = 1] = \ldots = \mathsf{Pr}[X = 6] = 1/6$

Elementary Probability Theory

- **Let X and Y be random variables (defined on sets $\mathcal{X}$ and $\mathcal{Y}$, resp.)**
- **Joint probability** $\mathsf{Pr}[X = x, Y = y]$ **is the probability that X takes value x and Y takes value y**
- **Conditional probability** $\mathsf{Pr}[X = x \mid Y = y]$ **is the probability that X takes value x given that Y takes value y**
- **X and Y are independent random variables if** $\mathsf{Pr}[X = x, Y = y] = \mathsf{Pr}[X = x]\mathsf{Pr}[Y = y]$ **for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$**

Elementary Probability Theory

- **Example with two perfect dice:**
  - **Let D_1 denote the result of throwing first dice, D_2 the result of throwing the second dice, and S their sum**
  - **What is the joint probability $\mathsf{Pr}[D_1 = 2, D_2 = 5]$?**
  - **What is the conditional probability $\mathsf{Pr}[D_2 = 3 \mid D_1 = 3]$?**
  - **Are D_1 and D_2 independent?**
  - **What is the joint probability $\mathsf{Pr}[D_1 = 3, S = 5]$?**
  - **Are D_1 and S independent?**
  - **What is the conditional probability $\mathsf{Pr}[S = 8 \mid D_1 = 4]$? $\mathsf{Pr}[S = 8 \mid D_1 = 1]$? $\mathsf{Pr}[D_1 = 3 \mid S = 4]$?**

Probability Theory

- **Conditional and joint probabilities are related:**

$$\Pr[X = x, Y = y] = \Pr[X = x \mid Y = y] \cdot \Pr[Y = y] \tag{1}$$

and

$$\Pr[X = x, Y = y] = \Pr[Y = y \mid X = x] \cdot \Pr[X = x] \tag{2}$$

- **From these two expressions we obtain Bayes' Theorem:**
  - **if** $\Pr[Y = y] > 0$**, then**

$$\Pr[X = x \mid Y = y] = \frac{\Pr[X = x] \cdot \Pr[Y = y \mid X = x]}{\Pr[Y = y]} \tag{3}$$

- **How is it useful to us?**

Probability Theory

- **Corollary:** X **and** Y **are independent random variables if and only if**

$$\mathsf{Pr}[X = x \mid Y = y] = \mathsf{Pr}[X = x]$$

for all $x \in \mathcal{X}$ **and** $y \in \mathcal{Y}$

  - **follows from definition of independent random variables and equation (1)**

- **This is what we need for perfect secrecy**

What Does This Do for Us?

- **Recall that a cipher is associated with $\mathcal{M}$, $\mathcal{K}$, and $\mathcal{C}$**
- **Let $\mathsf{Pr}[K = k]$ denote the probability of key $k \in \mathcal{K}$ being output by Gen**
- **Let $\mathsf{Pr}[M = m]$ define the a priori probability that message m is chosen for encryption**
- **M and K are independent and define ciphertext distribution C**
- **Given M, K and Enc, we can compute $\mathsf{Pr}[M = m \mid C = c]$**
- **This takes us to the notion of perfect secrecy...**

Perfect Secrecy

- **Definition: An encryption scheme** (Gen, Enc, Dec) **has perfect secrecy if for every distribution over $\mathcal{M}$, every $m \in \mathcal{M}$ and $c \in \mathcal{C}$ s.t.** $\Pr[C = c] > 0$**:**

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

- **Interpretation: after observing ciphertext c the a posteriori probability that the message is m is identical to the a priori probability that the message is m**

Perfect Secrecy

- **Alternative definition of perfect secrecy**
  - **An encryption scheme** $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ **is perfectly secret if and only if for every distribution over** $\mathcal{M}$ **and every** $m \in \mathcal{M}$ **and** $c \in \mathcal{C}$**:**

$$\mathsf{Pr}[C = c \mid M = m] = \mathsf{Pr}[C = c]$$

  - **This means that the probability distribution of the ciphertext does not depend on the plaintext**
  - **In other words, an encryption scheme** $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ **is perfectly secret if and only if for every distribution over** $\mathcal{M}$ **and every** $m_1, m_2 \in \mathcal{M}$ **and** $c \in \mathcal{C}$**:**

$$\mathsf{Pr}[C = c \mid M = m_1] = \mathsf{Pr}[C = c \mid M = m_2]$$

Perfect Indistinguishability

- **Indistinguishability of encrypted messages allows us to formulate security requirement as an experiment or game**
  - **interactive game with adversary $\mathcal{A}$, who tries to break a cryptographic scheme**
- **Our first experiment**
  - **for eavesdropping adversaries**
  - **using private-key encryption**
  - **asks them to distinguish between encryptions of different messages**
  - **let $\mathcal{E} = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$, and we name the experiment $\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\mathcal{E}}$**

Perfect Indistinguishability

- **Experiment** $\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\mathcal{E}}$

  1. **$\mathcal{A}$ chooses two messages $m_0, m_1 \in \mathcal{M}$**
  2. **random key k is generated by Gen, and random bit $b \leftarrow \{0, 1\}$ is chosen**
  3. **ciphertext $c \leftarrow \mathsf{Enc}_k(m_b)$ is computed and given to $\mathcal{A}$**
  4. **$\mathcal{A}$ outputs bit b' as its guess for b**
  5. **experiment outputs 1 if $b' = b$ ($\mathcal{A}$ wins) and 0 otherwise**

- **Given this experiment, how should we define indistinguishability? perfect secrecy?**

Perfect Indistinguishability

- **Definition:** **An encryption scheme** $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ **over message space** $\mathcal{M}$ **is perfectly secret if for every adversary** $\mathcal{A}$ **it holds that**

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

  - **notice that is must work for every** $\mathcal{A}$

- **This definition is equivalent to our original definition of perfect secrecy**

One-Time Pad

- **One-time pad (Vernam's cipher)**
  - **for fixed integer n, let $\mathcal{M} = \mathcal{K} = \mathcal{C} = \{0,1\}^n$**
  - Gen **chooses a key k uniformly at random from $\mathcal{K}$**
    - **each key is chosen with probability 2^{-n}**
  - Enc**: given key $k \in \{0,1\}^n$ and message $m \in \{0,1\}^n$, compute** $\mathsf{Enc}_k(m) = m \oplus k$
  - Dec**: given key $k \in \{0,1\}^n$ and ciphertext $c \in \{0,1\}^n$, compute** $\mathsf{Dec}_k(c) = c \oplus k$
- **Why is it perfectly secret?**

One-Time Pad

- **Theorem: One-time pad encryption scheme achieves perfect secrecy**
- **Proof**
  - **fix distribution over $\mathcal{M}$ and message $m \in \mathcal{M}$**

    $\mathsf{Pr}[C = c \mid M = m] =$

  - **this works for all distributions and all m, so for all distributions over $\mathcal{M}$, all $m_1, m_2 \in \mathcal{M}$, and all $c \in \mathcal{C}$:**

$$\mathsf{Pr}[C = c \mid M = m_1] = \mathsf{Pr}[C = c \mid M = m_2] = \frac{1}{2^n}$$

  - **by definition of perfect secrecy, this encryption is perfectly secret**

More on One-Time Pad

- **One-time pad can be defined on units larger than bits (e.g., letters)**
- **One-time pad questions:**
  - **Since the key must be long, what if we use text from a book as our key?**
  - **What if we reuse the key on different messages?**
  - **Can we securely encrypt using a short/reusable key?**
    - **no encryption scheme with smaller key space than message space can be perfectly secret**

Perfect Secrecy

- It can be shown that
  - Shift cipher has perfect secrecy if
    - the key is chosen randomly
    - it is used to encrypt a single letter
  - Similarly, Vigenère cipher has perfect secrecy if
    - each letter in the key is chosen randomly
    - the message has the same length as the key
- (Shannon's theorem) In general, to achieve perfect secrecy:
  - every key must be chosen with equal probability
  - for every message m and every ciphertext c, there is a unique key k such that $\mathsf{Enc}_k(m) = c$

Entropy

- **Entropy** H **measures the amount of information (or amount of uncertainty)**
- **The larger** H **of a message distribution is, the harder it is to predict that message**
- H **is measured in bits as the minimum number of bits required to encode all possible messages**

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

- **Examples**

Entropy

- **If there are n messages and they are all equally probable, then**

$$H(X) = -\sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{1}{n} = -\log_2 \frac{1}{n} = \log_2 n$$

- **Entropy is commonly used in security to measure information leakage**
  - **compute entropy before and after transmitting a ciphertext**
  - **if entropy associated with messages changes, leakage of information about transmitted message takes place**
  - **similarly, if uncertainty associated with the keys changes after transmission, leakage of key information takes place**

Entropy

- **Entropy after transmission is captured using conditional entropy** $H(X|Y)$
  - $H(M) - H(M|C)$ **defines information leakage about messages**
  - $H(K) - (K|C)$ **defines information leakage about keys**
- **Perfect secrecy is achieved if (and only if)** $H(M) = H(M|C)$
  - **that is, it is required that** M **and** C **are independent variables**

Entropy

- **Conditional entropy $H(X|Y)$ is defined as follows:**
  - **for each value y of Y, we get a conditional probability distribution on X, denoted by $X|y$**

$$H(X|y) = -\sum_{x \in \mathcal{X}} \Pr[X = x|Y = y] \cdot \log_2 \Pr[X = x|Y = y]$$

  - **conditional entropy $H(X|Y)$ is defined as the weighted average (w.r.t. probabilities $\Pr[Y = y]$) of entropies $H(X|y)$ over all possible y**

$$H(X|Y) = -\sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} (\Pr[Y = y] \cdot \Pr[X = x|Y = y] \cdot \log_2 \Pr[X = x|Y = y])$$

Language Redundancy

- **Absolute rate of a language**
  - **is the maximum number of bits that can be encoded in each character**
  - **assuming that each character sequence is equally likely**
- **In an alphabet of ℓ letters:**
  - **there are ℓ^n possible strings of size n**
  - **if all of them are equiprobable, the entropy of a string is $\log_2 \ell^n$**
  - **then the absolute language rate**

$$r_a = \frac{\log_2 \ell^n}{n} = \frac{n \log_2 \ell}{n} = \log_2 \ell$$

- **For English with $\ell = 26$, $r_a = 4.7$ bits**

Language Redundancy

- **Now compare that rate with the amount of information each English letter actually encodes**

- **Entropy of a language L is defined as**

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

  - **it measures the amount of entropy per letter and represents the average number of bits of information per character**

- **For English, $1 \leq H_L \leq 1.5$ bits per character**

- **Redundancy of English**

$$R_L = 1 - \frac{H_L}{r_a} = 1 - \frac{1.25}{4.7} \approx 0.75$$

Summary

- **Probabilities** are used to evaluate security of a cipher
- **Perfect secrecy** achieves unconditional security
- **One-time pad** is a provably unbreakable cipher but is hard to use in practice
- **Entropy** is used to measure the amount of uncertainty of the encryption key given a ciphertext
- **Next time:**
  - private-key encryption
  - computational security