Lecture 3
One-time Pad
One-Time Pad

• Basic Idea: Extend Vigenère cipher so that the key is as long as the plaintext
  – No repeat, cannot be broken by finding key length + frequency analysis
• Key is a *random string* that is at least as long as the plaintext
• Encryption is similar to Vigenère
One-Time Pad

- Key is chosen randomly

- Plaintext $X = (x_1 \ x_2 \ ... \ x_n)$

- Key $K = (k_1 \ k_2 \ ... \ k_n)$

- Ciphertext $Y = (y_1 \ y_2 \ ... \ y_n)$

- $e_k(X) = (x_1+k_1 \ x_2+k_2 \ ... \ x_n+k_n) \ mod \ m$

- $d_k(Y) = (y_1-k_1 \ y_2-k_2 \ ... \ y_n-k_n) \ mod \ m$
One-Time Pad

- Intuitively, it is secure ...
- The key is random, so the ciphertext too will be completely random
Shannon (Information-Theoretic) Security

• Basic Idea: Ciphertext should provide no “information” about Plaintext
• We also say such a scheme has perfect secrecy.
• One-time pad has perfect secrecy
  – E.g., suppose that the ciphertext is “Hello”, can we say any plaintext is more likely than another plaintext?
    (For example “Lucky”, “Later”, “Funny” … are all equally likely)
• Result due to Shannon, 1949.

Claude Elwood Shannon (1916 - 2001), an American electrical engineer and mathematician, has been called "the father of Information Theory"
Key Randomness in One-Time Pad

- One-Time Pad uses a *very* long key, what if the key is not chosen randomly, instead, texts from, e.g., a book is used.
  - this is not One-Time Pad anymore
  - this does not have perfect secrecy
  - this can be broken

- The key in One-Time Pad should never be reused.
  - If it is reused, it is Two-Time Pad, and is insecure!
Limitations of One-Time Pad

- Perfect secrecy $\Rightarrow$ key-length $\geq$ msg-length
- Difficult to use in practice
Limitations of One-Time Pad (2)

- One-Time Pad was used in World War 2: one-time key material was printed on silk, which agents could conceal inside their clothing; whenever a key had been used, it was torn off and burnt
- Now suppose you intercepted a message from a wartime German agent which you know started with “Heil Hitler”, and the first 10 letters of cyphertext were DGTYI BWPJA
- This means that the first 10 letters of the one-time pad were wclnb tdefj since

- Plaintext: heilhitler
- Key: wclnb tdefj
- Ciphertext: DGTYIBWPJA

A spy’s message
Limitations of One-Time Pad (2)

- But once he has burnt the piece of silk with his key material, the spy can claim he’s actually a member of the anti-Nazi underground resistance, and the message actually said «Hang Hitler». This is quite possible, as the key material could just as easily have been wggsb tdefj:

- Ciphertext: DGTYIBWPJA
- Key: wggsbtdefj
- Plaintext: hanghitler

What the spy claimed he said
Limitations of One-Time Pad (2)

• Now we rarely get anything for nothing in cryptology, and the price of the perfect secrecy of the one-time pad is that it fails completely to protect message integrity. Suppose for example that you wanted to get this spy into trouble, you could change the cyphertext to DCYTI BWPJA

• Ciphertext: DCYTIBWPJA
• Key: wcInbtddefj
• Plaintext: hanghitler
The Binary Version of One-Time Pad

• Plaintext space = Ciphertext space =
  = Keyspace = \{0,1\}^n
• Key is chosen randomly
• For example:
  – Plaintext is 11011011
  – Key is 01101001
  – Then ciphertext is 10110010
Bit Operators

• Bit AND
  \[-0 \land 0 = 0\]
  \[-0 \land 1 = 0\]
  \[-1 \land 0 = 0\]
  \[-1 \land 1 = 1\]

• Bit OR
  \[-0 \lor 0 = 0\]
  \[-0 \lor 1 = 1\]
  \[-1 \lor 0 = 1\]
  \[-1 \lor 1 = 1\]

• Addition mod 2 (also known as Bit XOR)
  \[-0 \oplus 0 = 0\]
  \[-0 \oplus 1 = 1\]
  \[-1 \oplus 0 = 1\]
  \[-1 \oplus 1 = 0\]
Unconditional Security

- The adversary has *unlimited* computational resources.
- Analysis is made by using probability theory.
- Perfect secrecy: observation of the ciphertext provides *no information* to an adversary.
- Result due to Shannon, 1949.
Begin Math
Elements of Probability Theory

• A random experiment has an unpredictable outcome.

• Definition
  The sample space (S) of a random phenomenon is the set of all outcomes for a given experiment.

• Definition
  The event (E) is a subset of a sample space, an event is any collection of outcomes.
Basic Axioms of Probability

• If $E$ is an event, $Pr(E)$ is the probability that event $E$ occurs, then
  – (a) $0 \leq Pr(A) \leq 1$ for any set $A$ in $S$.
  – (b) $Pr(S) = 1$, where $S$ is the sample space.
  – (c) If $E_1$, $E_2$, ..., $E_n$ is a sequence of mutually exclusive events, that is $E_i \cap E_j = 0$, for all $i \neq j$ then:

  $$Pr(E_1 \cup E_2 \cup ... \cup E_n) = \sum_{i=1}^{n} Pr(E_i)$$
Probability: More Properties

• If $E$ is an event and $Pr(E)$ is the probability that the event $E$ occurs then
  - $Pr(\hat{E}) = 1 - Pr(E)$ where $\hat{E}$ is the complimentary event of $E$
  - If outcomes in $S$ are equally like, then
    $Pr(E) = |E| / |S|$
    (where $|S|$ denotes the cardinality of the set $S$)
Random Variable

- Definition

A discrete random variable, $X$, consists of a finite set $X$, and a probability distribution defined on $X$. The probability that the random variable $X$ takes on the value $x$ is denoted $\Pr[X = x]$; sometimes, we will abbreviate this to $\Pr[x]$ if the random variable $X$ is fixed. It must be that

$$0 \leq \Pr[x] \quad \forall x \in X$$

$$\sum_{x \in X} \Pr[x] = 1$$
Relationships between Two Random Variables

• Definitions

Assume \( X \) and \( Y \) are two random variables, we define:

- joint probability: \( \Pr[x, y] \) is the probability that \( X \) takes value \( x \) and \( Y \) takes value \( y \).

- conditional probability: \( \Pr[x|y] \) is the probability that \( X \) takes on the value \( x \) given that \( Y \) takes value \( y \).

  • Note that joint probability can be related to conditional probability by the formula \( \Pr[x, y] = \Pr[x|y] \Pr[y] \)
  • Interchanging \( x \) and \( y \) we have that \( \Pr[x, y] = \Pr[y|x] \Pr[x] \)
  • This permits to obtain Bayes’ Theorem

- independent random variables: \( X \) and \( Y \) are said to be independent if \( \Pr[x,y]=\Pr[x]\Pr[y] \), for all \( x \in X \) and all \( y \in Y \)
Elements of Probability Theory

- Find the conditional probability of event \( X \) given the conditional probability of event \( Y \) and the unconditional probabilities of events \( X \) and \( Y \).

- **Bayes’ Theorem**
  If \( \Pr[y] > 0 \) then
  \[
  \Pr[x \mid y] = \frac{\Pr[y \mid x] \Pr[x]}{\Pr[y]}
  \]

- **Corollary**
  \( X \) and \( Y \) are independent random variables if and only if \( \Pr[x \mid y] = \Pr[x] \), for all \( x \in X \) and all \( y \in Y \).
End Math
Ciphers Modeled by Random Variables

• Consider a cipher \((P, C, K, E, D)\). We assume that:
  1. there is an (a-priori) probability distribution on the plaintext (message) space
  2. the key space also has a probability distribution. We assume the key is chosen before one (Alice) knows what the plaintext will be, therefore the key and the plaintext are independent random variables
  3. The two probability distributions on \(P\) and \(K\) induce a probability distribution on \(C\): the ciphertext is also a random variable
Example

- $P = \{a, b\}$;
- $\Pr(a) = 1/4$; $\Pr(b) = 3/4$

- $K = \{k1, k2, k3\}$;
- $\Pr(k1) = 1/2$; $\Pr(k2) = \Pr(k3) = 1/4$

- $C = \{1, 2, 3, 4\}$;
- $e_{k1}(a) = 1$; $e_{k1}(b) = 2$;
- $e_{k2}(a) = 2$; $e_{k2}(b) = 3$;
- $e_{k3}(a) = 3$; $e_{k3}(b) = 4$

$P =$ Plaintext

$C =$ Ciphertext

$K =$ Key

**Encryption Matrix**

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k1$</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$k2$</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>$k3$</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Example

- $P = \{a, b\}; \quad \Pr(a) = \frac{1}{4}; \quad \Pr(b) = \frac{3}{4}$
- $K = \{k1, k2, k3\}; \quad \Pr(k1) = \frac{1}{2}; \quad \Pr(k2) = \Pr(k3) = \frac{1}{4}$
- $C = \{1, 2, 3, 4\};$
  - $e_{k1}(a) = 1; \quad e_{k1}(b) = 2;$
  - $e_{k2}(a) = 2; \quad e_{k2}(b) = 3;$
  - $e_{k3}(a) = 3; \quad e_{k3}(b) = 4;$
- We now compute the probability distribution of the ciphertext:
  - $\Pr(1) = \Pr(k1) \Pr(a) = \frac{1}{2} \times 1 = \frac{1}{8}$
  - $\Pr(2) = \Pr(k1) \Pr(b) + \Pr(k2) \Pr(a) = \frac{1}{2} \times \frac{3}{4} + \frac{1}{4} \times 1 = \frac{7}{16}$
  - $\Pr(3) = \frac{1}{4}$
  - $\Pr(4) = \frac{3}{16}$
Example

- $P = \{a, b\}; \ Pr(a) = 1/4; \ Pr(b) = 3/4$
- $K = \{k_1, k_2, k_3\}; \ Pr(k_1) = 1/2; \ Pr(k_2) = Pr(k_3) = 1/4$
- $C = \{1, 2, 3, 4\}$
- Distribution of the ciphertext:
  - $Pr(1) = 1/8, Pr(2) = 7/16, Pr(3) = 1/4, Pr(4) = 3/16$

... Conditional probability distribution on the Plaintext, given that a certain ciphertext has been observed (we use Bayes)

$$Pr[a | 1] = \frac{Pr[1 | a]Pr[a]}{Pr[1]} = \frac{1 \cdot 1}{\frac{1}{8}} = 1$$

<table>
<thead>
<tr>
<th>Encryption Matrix</th>
<th>a</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>k1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>k2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>k3</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

| Pp(a | 1) | Pp(b | 1) |
|-------|-------|
| 1     | 0     |
| Pp(a | 2) | Pp(b | 2) |
| 1/7   | 6/7   |
| Pp(a | 3) | Pp(b | 3) |
| 1/4   | 3/4   |
| Pp(a | 4) | Pp(b | 4) |
| 0     | 1     |

DOES THIS CRYPTO SYSTEM HAVE PERFECT SECRECY?
Perfect Secrecy

• Definition
Informally, perfect secrecy means that an attacker cannot obtain any information about the plaintext, by observing the ciphertext.

What type of attack is this?

• Definition
A cryptosystem has perfect secrecy if \( \Pr[x|y] = \Pr[x] \), for all \( x \in P \) and \( y \in C \), where \( P \) is the set of plaintext and \( C \) is the set of ciphertext.
Perfect Secrecy

- What can I say about $\Pr[x|y]$ and $\Pr[x]$, for all $x \in P$ and $y \in C$?
- From Bayes’ Theorem

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

- Given
- Don’t know it, but can be computed
- Don’t know it, but can be computed
Perfect Secrecy

- **KNOWN, Pr[x], Pr[k]**

  $C(k)$: the set of all possible ciphertexts if key is $k$.

  \[
  \Pr[y | x] = \sum_{k : x = d_k(y)} \Pr[k] \\
  \Pr[y] = \sum_{k : y \in C(x)} \Pr[k] \Pr[x] \\
  \Pr[x | y] = \frac{\Pr[x] \cdot \sum_{k : x = d_k(y)} \Pr[k]}{\sum_{k : y \in C(x)} \Pr[k] \Pr[x]}
  \]
Example

- $P = \{a, b\}; \quad \Pr(a) = 1/4; \quad \Pr(b) = 3/4$
- $K = \{k1, k2, k3\}; \quad \Pr(k1) = 1/2; \quad \Pr(k2) = \Pr(k3) = 1/4$
- $C = \{1, 2, 3, 4\};$
  - $e_{k1}(a) = 1; \quad e_{k1}(b) = 2; $
  - $e_{k2}(a) = 2; \quad e_{k2}(b) = 3; $
  - $e_{k3}(a) = 3; \quad e_{k3}(b) = 4;$

- Distribution of the ciphertext:
  - $\Pr(1) = \Pr(k1) \Pr(a) = 1/2 * 1/4 = 1/8$
  - $\Pr(2) = \Pr(k1) \Pr(b) + \Pr(k2) \Pr(a) = 1/2 * 3/4 + 1/4 *1/4 = 7/16$
  - Similarly: $\Pr(3) = 1/4; \quad \Pr(4) = 3/16;$

- Conditional probability distribution of the ciphertext (we use Bayes)
  - $\Pr(a|1) = \Pr(1|a)\Pr(a)/\Pr(1) = 1/2*1/4/(1/8) = 1$
  - Similarly: $\Pr(a|2) = 1/7; \quad \Pr(a|3) = 1/4; \quad \Pr(a|4) = 0;$
  - $\Pr(b|1) = 0; \quad \Pr(b|2) = 6/7; \quad \Pr(b|3) = 3/4; \quad \Pr(b|4) = 1$

DOES THIS CRYPTOSYSTEM HAVE PERFECT SECRECY?
Names connected with OTP

• Co-inventors of One-time-pad
  – **Joseph Mauborgne** (1881-1971) became a Major General in the United States Army
  – **Gilbert Sandford Vernam** (1890 - 1960) was AT&T Bell Labs engineer

• Security of OTP
  – **Claude Elwood Shannon** (1916 - 2001), American electronic engineer and mathematician, was "the father of information theory."
Perfect secrecy of One-Time Pad
One-Time Pad has Perfect Secrecy

- $P = C = K = \{0,1\}^n$, the key is chosen randomly, the key used only once per message

- Proof: We need to show that for any probability of the plaintext, $\forall x \forall y$, $\Pr[x \mid y] = \Pr[x]$

\[
\Pr[x \mid y] = \frac{\Pr[x] \Pr[y \mid x]}{\Pr[y]} = \frac{\Pr[x] \Pr[k]}{\sum_{x \in X} \Pr[x] \Pr[k]} = \frac{\Pr[x]^{1/2^n}}{\sum_{x \in X} \Pr[x]^{1/2^n}} = \frac{\Pr[x]}{\sum_{x \in X} \Pr[x]} = \Pr[x]
\]
Modern Cryptography

• One-time pad requires the length of the key to be the length of the plaintext and the key to be used only once. Difficult to manage.

• Alternative: design cryptosystems where a key is used more than once.

• What about the attacker? Resource constrained, make it infeasible for adversary to break the cipher.
Stream Ciphers

• In OTP, a key is described by a random bit string of length n
• Stream ciphers:
• Idea: replace “rand” by “pseudo rand”
• Use Pseudo Random Number Generator (PRNG)
• PRNG: \{0, 1\}^s \rightarrow \{0, 1\}^n
  – expand a short (e.g., 128-bit) random seed into a long (e.g., 10^6 bit) string that “looks random”
  – Secret key is the seed
  – \(E_{\text{seed}}[M] = M \oplus \text{PRNG}(\text{seed})\)
Properties of Stream Ciphers

• Does not have perfect secrecy
  – security depends on PRNG
• PRNG must be “unpredictable”
  – given consecutive sequence of bits output (but not seed), next bit must be hard to predict
• Typical stream ciphers are very fast
• Used in many places, often incorrectly
  – SSL( Rivest Cipher 4, or RC4), DVD (LFSR), WEP (RC4), etc.
Fundamental Weaknesses of Stream Ciphers

• If the same key-stream is used twice ever, then easy to break.

• Highly malleable
  – easy to change ciphertext so that plaintext changes in predictable, e.g., flip bits

• Weaknesses exist even if the PRNG is strong