Classic Crypto
Overview

- We briefly consider the following classic (pen and paper) ciphers
  - Transposition ciphers
  - Substitution ciphers
  - One-time pad
  - Codebook

- These were all chosen for a reason
  - We see same principles in modern ciphers
Transposition Ciphers

- In transposition ciphers, we transpose (scramble) the plaintext letters
  - The scrambled text is the ciphertext
  - The transposition is the key

- Corresponds to Shannon’s principle of diffusion (more about this later)
  - This idea is widely used in modern ciphers
Scytale

- Spartans, circa 500 BC
- Wind strip of leather around a rod
- Write message across the rod

<table>
<thead>
<tr>
<th>T</th>
<th>H</th>
<th>E</th>
<th>T</th>
<th>I</th>
<th>M</th>
<th>E</th>
<th>H</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>C</td>
<td>O</td>
<td>M</td>
<td>E</td>
<td>T</td>
<td>H</td>
<td>E</td>
<td>W</td>
</tr>
<tr>
<td>A</td>
<td>L</td>
<td>R</td>
<td>U</td>
<td>S</td>
<td>S</td>
<td>A</td>
<td>I</td>
<td>D</td>
</tr>
<tr>
<td>T</td>
<td>O</td>
<td>T</td>
<td>A</td>
<td>L</td>
<td>K</td>
<td>O</td>
<td>F</td>
<td>M</td>
</tr>
<tr>
<td>A</td>
<td>N</td>
<td>Y</td>
<td>T</td>
<td>H</td>
<td>I</td>
<td>N</td>
<td>G</td>
<td>S</td>
</tr>
</tbody>
</table>

- When unwrapped, letters are scrambled
  TSATAHCLONEORTYTMUATIESLHMTS...
Scytale

- Suppose Alice and Bob use Scytale to encrypt a message
  - What is the key?
  - How hard is it for Trudy to break without key?

- Suppose many different rod diameters are available to Alice and Bob...
  - How hard is it for Trudy to break a message?
  - Can Trudy attack messages automatically—without manually examining each putative decrypt?
Columnar Transposition

- Put plaintext into rows of matrix then read ciphertext out of columns
- For example, suppose matrix is 3 x 4
  - Plaintext: SEETHELIGHT
    
    \[
    \begin{bmatrix}
    S & E & E & T \\
    H & E & L & I \\
    G & H & T & X \\
    \end{bmatrix}
    \]
  - Ciphertext: SHGEEHELT IX
- Same effect as Scytale
  - What is the key?
Keyword Columnar Transposition

- For example
  - Plaintext: CRYPTOISFUN
  - Matrix 3 x 4 and keyword MATH
    \[
    \begin{bmatrix}
    M & A & T & H \\
    C & R & Y & P \\
    T & O & I & S \\
    F & U & N & X
    \end{bmatrix}
    \]
  - Ciphertext: ROUPSXCTFYIN
- What is the key?
- How many keys are there?
Keyword Columnar Transposition

- How can Trudy cryptanalyze this cipher?
- Consider the ciphertext

  VOESA IVENE MRTNL EANGE WTNIM HTMLL ADLTR NISHO DWOEH

- Matrix is $n \times m$ for some $n$ and $m$
- Since 45 letters, $n \cdot m = 45$
- How many cases to try?
- How will Trudy know when she is correct?
Keyword Columnar Transposition

- The ciphertext is

VOESA IVENE MRTNL EANGW WTNIM HTMLL ADLTR NISHO DWOEH

- If encryption matrix was 9 x 5, then...

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>E</td>
<td>G</td>
<td>M</td>
<td>I</td>
</tr>
<tr>
<td>O</td>
<td>M</td>
<td>E</td>
<td>E</td>
<td>S</td>
</tr>
<tr>
<td>E</td>
<td>R</td>
<td>W</td>
<td>E</td>
<td>H</td>
</tr>
<tr>
<td>S</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>O</td>
</tr>
<tr>
<td>A</td>
<td>N</td>
<td>N</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>I</td>
<td>L</td>
<td>I</td>
<td>L</td>
<td>W</td>
</tr>
<tr>
<td>V</td>
<td>E</td>
<td>M</td>
<td>T</td>
<td>O</td>
</tr>
<tr>
<td>E</td>
<td>A</td>
<td>H</td>
<td>R</td>
<td>E</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>T</td>
<td>N</td>
<td>H</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2</th>
<th>4</th>
<th>0</th>
<th>1</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>I</td>
<td>V</td>
<td>E</td>
<td>M</td>
</tr>
<tr>
<td>E</td>
<td>S</td>
<td>O</td>
<td>M</td>
<td>E</td>
</tr>
<tr>
<td>W</td>
<td>H</td>
<td>E</td>
<td>R</td>
<td>E</td>
</tr>
<tr>
<td>T</td>
<td>O</td>
<td>S</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td>N</td>
<td>D</td>
<td>A</td>
<td>N</td>
<td>D</td>
</tr>
<tr>
<td>I</td>
<td>W</td>
<td>I</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>M</td>
<td>O</td>
<td>V</td>
<td>E</td>
<td>T</td>
</tr>
<tr>
<td>H</td>
<td>E</td>
<td>E</td>
<td>A</td>
<td>R</td>
</tr>
<tr>
<td>T</td>
<td>H</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>
Cryptanalysis: Lesson I

- Exhaustive key search
  - Always an option for Trudy
- If keyspace is too large, such an attack will not succeed in a reasonable time
  - Or it will have a low probability of success
- A large keyspace is necessary for security
- But, large keyspace is not sufficient...
Double Transposition

- **Plaintext:** ATTACK AT DAWN

<table>
<thead>
<tr>
<th>columns</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>row 0</td>
<td>A</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>row 1</td>
<td>A</td>
<td>C</td>
<td>K</td>
</tr>
<tr>
<td>row 2</td>
<td>X</td>
<td>A</td>
<td>T</td>
</tr>
<tr>
<td>row 3</td>
<td>X</td>
<td>D</td>
<td>A</td>
</tr>
<tr>
<td>row 4</td>
<td>W</td>
<td>N</td>
<td>X</td>
</tr>
</tbody>
</table>

- **Ciphertext:** XTAWXNATTXADAKC

- **Key?**
  - 5 x 3 matrix, perms (2,4,0,3,1) and (0,2,1)
Double Transposition

- How can Trudy attack double transposition?
  - Spse Trudy sees 45-letter ciphertext
  - Then how many keys?
    - Size of matrix: 3 x 15, 15 x 3, 5 x 9, or 9 x 5
    - A lot of possible permutations!
      - $5! \cdot 9! > 2^{25}$ and $3! \cdot 15! > 2^{42}$
  - Size of keyspace is greater than $2^{43}$
- Is there a shortcut attack?
Double Transposition

- Shortcut attack on double transposition?
- Suppose ciphertext is
  ILILWEAHREOMEESANNDVEGMIERWEHVEMTOSTTAONNNTNH
- Suppose Trudy guesses matrix is 9 x 5
- Then Trudy has:
  \[
  \begin{array}{cccccc}
  \text{column} & 0 & 1 & 2 & 3 & 4 \\
  \text{row 0} & I & L & I & L & W \\
  \text{row 1} & E & A & H & R & E \\
  \text{row 2} & O & M & E & E & S \\
  \text{row 3} & A & N & N & D & D \\
  \text{row 4} & V & E & G & M & I \\
  \text{row 5} & E & R & W & E & H \\
  \text{row 6} & V & E & M & T & O \\
  \text{row 7} & S & T & T & A & 0 \\
  \text{row 8} & N & N & T & N & H \\
  \end{array}
  \]
- Now what?
- Try all perms?
  \[5! \cdot 9! > 2^{25}\]
- Is there a better way?
Double Transposition

- Shortcut attack on double transposition?
- Trudy tries “columns first” strategy

<table>
<thead>
<tr>
<th>column</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>row 0</td>
<td>I</td>
<td>L</td>
<td>I</td>
<td>L</td>
<td>W</td>
</tr>
<tr>
<td>row 1</td>
<td>E</td>
<td>A</td>
<td>H</td>
<td>R</td>
<td>E</td>
</tr>
<tr>
<td>row 2</td>
<td>O</td>
<td>M</td>
<td>E</td>
<td>E</td>
<td>S</td>
</tr>
<tr>
<td>row 3</td>
<td>A</td>
<td>N</td>
<td>N</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>row 4</td>
<td>V</td>
<td>E</td>
<td>G</td>
<td>M</td>
<td>I</td>
</tr>
<tr>
<td>row 5</td>
<td>E</td>
<td>R</td>
<td>W</td>
<td>E</td>
<td>H</td>
</tr>
<tr>
<td>row 6</td>
<td>V</td>
<td>E</td>
<td>M</td>
<td>T</td>
<td>O</td>
</tr>
<tr>
<td>row 7</td>
<td>S</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>O</td>
</tr>
<tr>
<td>row 8</td>
<td>N</td>
<td>N</td>
<td>T</td>
<td>N</td>
<td>H</td>
</tr>
</tbody>
</table>

Permute columns

<table>
<thead>
<tr>
<th>column</th>
<th>2</th>
<th>4</th>
<th>0</th>
<th>1</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>row 0</td>
<td>I</td>
<td>W</td>
<td>I</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>row 1</td>
<td>H</td>
<td>E</td>
<td>E</td>
<td>A</td>
<td>R</td>
</tr>
<tr>
<td>row 2</td>
<td>E</td>
<td>S</td>
<td>O</td>
<td>M</td>
<td>E</td>
</tr>
<tr>
<td>row 3</td>
<td>N</td>
<td>D</td>
<td>A</td>
<td>N</td>
<td>D</td>
</tr>
<tr>
<td>row 4</td>
<td>G</td>
<td>I</td>
<td>V</td>
<td>E</td>
<td>M</td>
</tr>
<tr>
<td>row 5</td>
<td>W</td>
<td>H</td>
<td>E</td>
<td>R</td>
<td>E</td>
</tr>
<tr>
<td>row 6</td>
<td>M</td>
<td>O</td>
<td>V</td>
<td>E</td>
<td>T</td>
</tr>
<tr>
<td>row 7</td>
<td>T</td>
<td>O</td>
<td>S</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td>row 8</td>
<td>T</td>
<td>H</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

- Now what?

Classic Crypto
Cryptanalysis: Lesson II

- Divide and conquer
  - Trudy attacks part of the keyspace
  - A great shortcut attack strategy
- Requires careful analysis of algorithm
- We will see this again and again in the attacks discussed later
- Of course, cryptographers try to prevent divide and conquer attacks
Substitution Ciphers

- In substitution ciphers, we replace the plaintext letters with other letters
  - The resulting text is the ciphertext
  - The substitution rule is the key
- Corresponds to Shannon’s principle of confusion (more on this later)
  - This idea is used in modern ciphers
Ceasar’s Cipher

- **Plaintext:**
  FOURSCOREANDSEVENYEARSAGO

- **Key:**
  
<table>
<thead>
<tr>
<th>Plaintext</th>
<th>abcdefghijklmnopqrstuvwxyz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ciphertext</td>
<td>DEFGHIJKLMNOPQRSTUVWXYZABC</td>
</tr>
</tbody>
</table>

- **Ciphertext:**
  IRXUVFRUHDAGVHYHABHDUVDIR

- More succinctly, key is “shift by 3”
Ceasar’s Cipher

- Trudy loves the Ceasar’s cipher...
- Suppose ciphertext is
  VSRQJHEREVTEXDUHSDQWU

| Plaintext | a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z |
| Ciphertext| D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z | A | B | C |

- Then plaintext is
  SPONGEBOBSQUAREPANTS
Simple Substitution

- Caesar’s cipher is trivial if we adhere to Kerckhoffs’ Principle
- We want a substitution cipher with lots of keys
- What to do?
- Generalization of Caesar’s cipher...
## Simple Substitution

- Key is some **permutation** of letters
- Need not be a shift
- For example

<table>
<thead>
<tr>
<th>Plaintext</th>
<th>a b c d e f g h i j k l m n o p q r s t u v w x y z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ciphertext</td>
<td>J I C A X S E Y V D K W B Q T Z R H F M P N U L G O</td>
</tr>
</tbody>
</table>

- Then $26! > 2^{88}$ possible keys
- That’s lots of keys!
Cryptanalysis of Simple Substitution

- Trudy knows a simple substitution is used.
- Can she find the key given ciphertext:

PBFPVYFBQXZTYFPBFEPQJHDXXQVAPTPQJKTOYQWIPBVW
LXTOXBTFXQWXBVXCXQWAXFQJVVWLEQNTOZQGGQLFXQ
WAKVWLXQWAEBIPBFXFQVXGTVJVWLBTPQWAEBFPBFH
CVLXBQUFEXWLXGDPEQVPQGVPBPBTIXPFHXZHVFA GF
OTHFEFBUFDTDHZBQPOTHXYFTODXQHFTDPTOGHFQP
BQWAQJJTOIDXQHFOQPWTDHHTIXQVAPBFZQHCFWPFH P
BFIPBQWKFABVYDYDZBOTHPBQPQJTOQTOGHFQAPBFEQ
JHDXXQVAVXEBOPEFZBVFOJWWACFCFCCFHQWAUVWFL
QHGFXVAFXQHFUFHILTAVWAFWAUTEVOTEIDHFHFQAI
TIXPFHXAFQHEFZQWGFLVWPOTFFA
Cryptanalysis of Simple Substitution

- Trudy cannot try all $2^{88}$ possible keys
- Can she be more clever?
- Statistics!
- English letter frequency counts:
Cryptanalysis of Simple Substitution

- Ciphertext:

PBFPVYFBQXZTYFPBFEOQJHDXQVAPTQJKTOYQWIPCBLXOTOXBTF
XQWAXBVCXQWAXFQJVWLEQNTQZQQGLFXQWAKVWLXQWAEBIPBF
XFQVXGTJYVWLBTQWAEFBFPBHCVXLXBUFEVWLXGDPEQVFPQVPB
PBFTIXPFHXZHVFAFOFHTEFBQURTTHDZBQPOTHXTYFTODXQHT
DPTOGFHPBPQWAQJJTODXQHFQOPWTBDHIXQVAPBFZQHCFWPFH
PBFIPBQWKFABVYDZBOTHPBPQJTTQOTOGFQAPBFEQJHDXXQV
AVXEBQPEFBVFOJWFFACFCCHQWAUVWFLQHGFVAFXQHFUFH
ILTTAVWFAFFAWEODTDHFHFQAIITIXPFHXAFQHEFZQWGFLVWPT
OFFA

- Ciphertext frequency counts:

| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| 21| 26| 6 | 10| 12| 51| 10| 25| 10| 9 | 3 | 10| 0 | 1 | 15| 28| 42| 0 | 0 | 27| 4 | 24| 22| 28| 6 | 8 |
Cryptanalysis: Lesson III

- **Statistical analysis**
  - Statistics might reveal info about key
- Ciphertext should appear random
- But randomness is not easy
  - Difficult to define random (entropy)
- Cryptographers work hard to prevent statistical attacks
Poly-Alphabetic Substitution

- Like a simple substitution, but permutation ("alphabet") changes
  - Often, a new alphabet for each letter
- Very common in classic ciphers
  - Vigenere cipher is an example
  - Discuss Vigenere later in this section
- Used in WWII-era cipher machines
Affine Cipher

- Number the letters 0 thru 25
  - A is 0, B is 1, C is 2, etc.
- Then affine cipher encryption is defined by
  \[ c_i = a p_i + b \pmod{26} \]
  - Where \( p_i \) is the \( i^{th} \) plaintext letter
  - And \( a \) and \( b \) are constants
  - Require that \( \gcd(a, 26) = 1 \) (why?)
Affine Cipher

- **Encryption:** $c_i = a p_i + b \pmod{26}$
- **Decryption:** $p_i = a^{-1}(c_i - b) \pmod{26}$
- **Keypspace size?**
  - Keyspace size is $26 \cdot \varphi(26) = 312$
  - Too small to be practical
Vigenere Cipher

- Key is of the form $K = (k_0, k_1, \ldots, k_{n-1})$
  - Where each $k_i \in \{0, 1, 2, \ldots, 25\}$
- Encryption
  $$c_i = p_i + k_i \pmod{n} \pmod{26}$$
- Decryption
  $$p_i = c_i - k_i \pmod{n} \pmod{26}$$
- Nothing tricky here!
- Just a repeating sequence of (shift by $n$) simple substitutions
Vigenere Cipher

- For example, suppose key is MATH
  - That is, $K = (12, 0, 19, 7)$, since M is letter 12, and so on
- Plaintext: SECRETMESSAGE
- Ciphertext: EEVYQTFLESTNQ
- Encrypt:
  \[
  \begin{array}{cccccccccccccccc}
  S & E & C & R & E & T & M & E & S & S & A & G & E \\
  18 & 4 & 2 & 17 & 4 & 19 & 12 & 4 & 18 & 18 & 0 & 6 & 4 \\
  +12 & 0 & 19 & 7 & 12 & 0 & 19 & 7 & 12 & 0 & 19 & 7 & 12 \\
  \hline
  4 & 4 & 21 & 24 & 16 & 19 & 5 & 11 & 4 & 18 & 19 & 13 & 16 \\
  \end{array}
  \quad (\text{mod 26})
  \]
  
  E E V Y Q T F L E S T N Q
Vigenere Cipher

- Vigenere is just a series of $k$ simple substitution ciphers
- Should be able to do $k$ simple substitution attacks
  - Provided enough ciphertext
- But how to determine $k$ (key length)?
- Index of coincidence...
Index of Coincidence

- Assume ciphertext is English letters
- Let $n_0$ be number of As, $n_1$ number of Bs, ..., $n_{25}$ number of Zs in ciphertext
- Let $n = n_0 + n_1 + \cdots + n_{25}$
- Define index of coincidence

\[
I = \frac{(n_0/2) + (n_1/2) + \cdots + (n_{25}/2)}{\binom{n}{2}} = \frac{1}{n(n-1)} \sum_{i=0}^{25} n_i(n_i - 1)
\]

- What does this measure?
Index of Coincidence

- Gives the probability that 2 randomly selected letters are the same
- For plain English, prob. 2 letter are same:
  - $p_0^2 + p_1^2 + \ldots + p_{25}^2 \approx 0.065$, where $p_i$ is probability of $i^{th}$ letter
- Then for simple substitution, $I \approx 0.065$
- For random letters, each $p_i = 1/26$
  - Then $p_0^2 + p_1^2 + \ldots + p_{25}^2 \approx 0.03846$
- Then $I \approx 0.03846$ for poly-alphabetic substitution with a very long keyword
Index of Coincidence

- How to use this to estimate length of keyword in Vigenere cipher?
- Suppose keyword is length $k$, message is length $n$
  - Ciphertext in matrix with $k$ columns, $n/k$ rows
- Select 2 letters from same columns
  - Like selecting from simple substitution
- Select 2 letters from different columns
  - Like selecting random letters
Index of Coincidence

- Suppose $k$ columns and $n/k$ rows
- Approximate number of matching pairs from same column, but 2 different rows:
  
  \[
  0.065 \left(\frac{n}{k}\right)^2 = 0.065 \frac{1}{2} \left(\frac{n}{k}\right) \left(\frac{n}{k} - 1\right) \frac{k}{2k} = 0.065 \left(\frac{n(n - k)}{2k}\right)
  \]

- Approximate number of matching pairs from 2 different columns, and any two rows:

  \[
  0.03846 \binom{k}{2} \left(\frac{n}{k}\right)^2 = 0.03846 \frac{n^2(k - 1)}{2k}
  \]
Index of Coincidence

Approximate index of coincidence by:

\[ I \approx \frac{0.03846 \frac{n^2(k-1)}{2k} + 0.065 \left( \frac{n(n-k)}{2k} \right)}{\binom{n}{2}} \]

\[ = \frac{0.03846n(k-1) + (0.065)(n-k)}{k(n-1)} \]

Solve for \( k \) to find:

\[ k \approx \frac{0.02654n}{(0.065 - I) + n(I - 0.03846)} \]

Use \( n \) and \( I \) (known from ciphertext) to approximate length of Vigenere keyword
Index of Coincidence: Bottom Line

- A crypto breakthrough when invented
  - By William F. Friedman in 1920s
- Useful against classical and WWII-era ciphers
- Incidence of coincidence is a well-known statistical test
  - Many other statistical tests exist
Hill Cipher

- Hill cipher is not related to small mountains
- Invented by Lester Hill in 1929
  - A pre-modern block cipher
- Idea is to create a substitution cipher with a large “alphabet”
- All else being equal (which it never is) cipher should be stronger than simple substitution
Hill Cipher

- Plaintext, $p_0, p_1, p_2, \ldots$
- Each $p_i$ is block of $n$ consecutive letters
  - As a column vector
- Let $A$ be $n \times n$ invertible matrix, mod 26
- Then ciphertext block $c_i$ is given by
  - $c_i = A p_i \pmod{26}$
  - Decryption: $p_i = A^{-1} c_i \pmod{26}$
- The matrix $A$ is the key
Hill Cipher Example

- Let $n = 2$ and 
  \[ A = \begin{bmatrix} 22 & 13 \\ 11 & 5 \end{bmatrix} \]

- Plaintext
  \[ \text{MEETMEHERE} = (12,4,4,19,12,4,7,4,17,4) \]

- Then
  \[ p_0 = \begin{bmatrix} 12 \\ 4 \end{bmatrix}, \quad p_1 = \begin{bmatrix} 4 \\ 19 \end{bmatrix}, \quad p_2 = \begin{bmatrix} 12 \\ 4 \end{bmatrix}, \quad p_3 = \begin{bmatrix} 7 \\ 4 \end{bmatrix}, \quad p_4 = \begin{bmatrix} 17 \\ 4 \end{bmatrix} \]

- And
  \[ c_0 = \begin{bmatrix} 4 \\ 22 \end{bmatrix}, \quad c_1 = \begin{bmatrix} 23 \\ 9 \end{bmatrix}, \quad c_2 = \begin{bmatrix} 4 \\ 22 \end{bmatrix}, \quad c_3 = \begin{bmatrix} 24 \\ 19 \end{bmatrix}, \quad c_4 = \begin{bmatrix} 10 \\ 25 \end{bmatrix} \]

- Ciphertext:
  \[ (4,22,23,9,4,22,24,19,10,25) = \text{EWXJEWYTKZ} \]
Hill Cipher Cryptanalysis

- Trudy suspects Alice and Bob are using Hill cipher, with $n \times n$ matrix $A$
- Suppose Trudy knows $n$ plaintext blocks
  - Plaintext blocks $p_0, p_1, \ldots, p_{n-1}$
  - Ciphertext blocks $c_0, c_1, \ldots, c_{n-1}$
- Let $P$ be matrix with columns $p_0, p_1, \ldots, p_{n-1}$
- Let $C$ be matrix with columns $c_0, c_1, \ldots, c_{n-1}$
- Then $AP = C$ and $A = CP^{-1}$ if $P^{-1}$ exists
Cryptanalysis: Lesson IV

- **Linear** ciphers are weak
  - Since linear equations are easy to solve
- Strong cipher must have nonlinearity
  - Linear components are useful
  - But cipher cannot be entirely linear
- Cryptanalyst try to approximate nonlinear parts with linear equations
One-time Pad

- A provably secure cipher
- No other cipher we discuss is provably secure
- Why not use one-time pad for everything?
  - Impractical for most applications
  - But it does have its uses
One-time Pad Encryption

Encryption: \( \text{Plaintext} \oplus \text{Key} = \text{Ciphertext} \)

| Plaintext: | 001 000 010 100 001 010 111 100 000 101 |
| Key:       | 111 101 110 101 111 100 000 101 110 000 |
| Ciphertext:| 110 101 100 001 110 110 111 001 110 101 |

| e=000    | h=001    | i=010    | k=011    | l=100    | r=101    | s=110    | t=111    |

Classic Crypto
One-time Pad Decryption

Ciphertext: 110 101 100 001 110 110 111 001 110 101
Key: 111 101 110 101 111 100 000 101 110 000
Plaintext: 001 000 010 100 001 010 111 100 000 101

Decryption: Ciphertext $\oplus$ Key = Plaintext

h e i l h i t l e r
One-time Pad

Double agent claims sender used “key”:

\[
\begin{array}{cccccccccc}
\text{Ciphertext:} & 110 & 101 & 100 & 001 & 110 & 110 & 111 & 001 & 110 & 101 \\
\text{“key”:} & 101 & 111 & 000 & 101 & 111 & 100 & 000 & 101 & 110 & 000 \\
\text{“Plaintext”:} & 011 & 010 & 100 & 100 & 001 & 010 & 111 & 100 & 000 & 101 \\
\end{array}
\]

\[
\begin{array}{ccccccccccc}
k & i & l & l & h & i & t & l & e & r \\
\end{array}
\]

\[
\begin{array}{ccccccccccc}
e=000 & h=001 & i=010 & k=011 & l=100 & r=101 & s=110 & t=111 \\
\end{array}
\]
One-time Pad

Sender is captured and claims the key is:

\[
\begin{array}{cccccccccccc}
\text{s} & \text{r} & \text{l} & \text{h} & \text{s} & \text{s} & \text{t} & \text{h} & \text{s} & \text{r} \\
\end{array}
\]

Ciphertext: \[110\ 101\ 100\ 001\ 110\ 110\ 111\ 001\ 110\ 101\]

“Key”: \[111\ 101\ 000\ 011\ 101\ 110\ 001\ 011\ 101\ 101\]

“Plaintext”: \[001\ 000\ 100\ 010\ 011\ 000\ 110\ 010\ 011\ 000\]

\[
\begin{array}{cccccccccccc}
\text{h} & \text{e} & \text{l} & \text{i} & \text{k} & \text{e} & \text{s} & \text{i} & \text{k} & \text{e} \\
\end{array}
\]

e=000 \ h=001 \ i=010 \ k=011 \ l=100 \ r=101 \ s=110 \ t=111

Classic Crypto
One-time Pad Summary

- Provably secure, when used correctly
  - Ciphertext provides no info about plaintext
  - All plaintexts are equally likely
  - Pad must be random, used only once
  - Pad is known only by sender and receiver
  - Pad is same size as message
  - No assurance of message integrity

- Why not distribute message the same way as the pad?
Real-world One-time Pad

- **Project VENONA**
  - Soviet spy messages from U.S. in 1940's
  - Nuclear espionage, etc.
  - Thousands of messages
- Spy carried one-time pad into U.S.
- Spy used pad to encrypt secret messages
- Repeats within the “one-time” pads made cryptanalysis possible
[C% Ruth] learned that her husband [v] was called up by the army but he was not sent to the front. He is a mechanical engineer and is now working at the ENORMOUS [ENORMOZ] [vi] plant in SANTA FE, New Mexico. [45 groups unrecoverable]

detain VOLOK [vii] who is working in a plant on ENORMOUS. He is a FELLOWCOUNTRYMAN [ZEMLYaK] [viii]. Yesterday he learned that they had dismissed him from his work. His active work in progressive organizations in the past was cause of his dismissal. In the FELLOWCOUNTRYMAN line LIBERAL is in touch with CHESTER [ix]. They meet once a month for the payment of dues. CHESTER is interested in whether we are satisfied with the collaboration and whether there are not any misunderstandings. He does not inquire about specific items of work [KONKRETNAYa RABOTA]. In as much as CHESTER knows about the role of LIBERAL's group we beg consent to ask C. through LIBERAL about leads from among people who are working on ENOURMOUS and in other technical fields.

- “Ruth” == Ruth Greenglass
- “Liberal” == Julius Rosenberg
- “Enormous” == the atomic bomb
Codebook Cipher

- Literally, a book filled with “codes”
  - More precisely, 2 codebooks, 1 for encryption and 1 for decryption
- Key is the codebook itself
- Security of cipher requires physical security for codebook
- Codebooks widely used thru WWII
Codebook Cipher

- Literally, a book filled with “codewords”
- **Zimmerman Telegram** encrypted via codebook
  
  Februar 13605
  fest 13732
  finanzielle 13850
  folgender 13918
  Frieden 17142
  Friedenschluss 17149
  
- Modern block ciphers are codebooks!
- More on this later...
Zimmerman Telegram

- One of most famous codebook ciphers ever
- Led to US entry in WWI
- Ciphertext shown here...

Classic Crypto
Zimmerman Telegram Decrypted

- British had recovered partial codebook
- Able to fill in missing parts

Classic Crypto
Codebook Cipher

- Codebooks are susceptible to statistical analysis
  - Like simple substitution cipher, but lots of data required to attack a codebook
- Historically, codebooks very popular
- To extend useful life of a codebook, an **additive** was usually used
Codebook Additive

- Codebook **additive** is another book filled with “random” number
- Sequence of additive numbers added to codeword to yield ciphertext

plaintext  lookup in codebook  codeword  add the additive  ciphertext
Codebook Additive

- Usually, starting position in additive book selected at random by sender
- Starting additive position usually sent “in the clear” with the ciphertext
  - Part of the message indicator (MI)
  - Modern term: initialization vector (IV)
- Why does this extend the useful life of a codebook?
Cryptanalysis: Summary

- Exhaustive key search
- Divide and conquer
- Statistical analysis
- Exploit linearity
- Or any combination thereof (or anything else you can think of)
- All’s fair in love and war...
  - ...and cryptanalysis!