The Basic Building Blocks: Cryptography

Week 2

Stallings Ch2
Outline

- Encryption Techniques
  - Classical Techniques
  - Crash Course in Modern Crypto
Classic Techniques

- Substitution Ciphers
  - Caesar Cipher
  - Monoalphabetic Ciphers
  - Polyalphabetic Ciphers
    - Vigenere
    - Vernam
    - One Time Pad
- Transposition Ciphers
- Product Ciphers
Classic Substitution Techniques

- Letters of plaintext are replaced by
  - other letters
  - numbers
  - symbols
- If plaintext is viewed as a sequence of bits
  - Replace plaintext bits with ciphertext bits
Caesar Cipher

- Earliest known substitution cipher
- By Julius Caesar
- First attested use in military affairs
- Example:
  
  meet me after the toga party
  PHHW PH DIWHU WKH WRJD SDUWB
- Guess how?
  - Replaces each letter by 3rd letter on
Define transformation as:

\[ \begin{align*}
&\text{a b c d e f g h i j k l m n o p q r s t u v w x y z} \\
&D\ E\ F\ G\ H\ I\ J\ K\ L\ M\ N\ O\ P\ Q\ R\ S\ T\ U\ V\ W\ X\ Y\ Z\ A\ B\ C
\end{align*} \]

Give each letter a number

\[ \begin{align*}
&\text{a b c d e f g h i j k l m n o p q r s t u v w x y z} \\
&0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10 \quad 11 \quad 12 \quad 13 \quad 14 \quad 15 \quad 16 \quad 17 \quad 18 \quad 19 \quad 20 \quad 21 \quad 22 \quad 23 \quad 24 \quad 25
\end{align*} \]

Math -- Caesar cipher is:

\[ \begin{align*}
&c = E(k, p) = (p + k) \mod (26) \\
p = D(k, c) = (c - k) \mod (26)
\end{align*} \]
Cryptanalysis of Caesar Cipher

- Only have 25 possible ciphers
  - A maps to B, C, ..., Z

- Brute force search
  - Given ciphertext, just try all shifts of letters
  - Do need to recognize when have plaintext
  - E.g. break ciphertext
    - QIIX QI MR QMEQM
Monoalphabetic Cipher

- Shuffle the letters arbitrarily
  - Random permutation – how many in total?
- Each plaintext letter maps to a different random ciphertext letter
- Key is 26 letters long
- **Example:**

  Plain:  abcdefghijklmnopqrstuvwxyz
  Cipher: DKVQFIBJWPESCXHTMYAUOLRGZN

  Plaintext: ifwewishtoreplaceletters
  Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA
Monoalphabetic Cipher Security

- Key space increases from 25 (Caesar) to 26! = 4 x $10^{26}$ (English)
- [http://www.simonsingh.net/The_Black_Chamber/generalsubstitutionWithMenu.html](http://www.simonsingh.net/The_Black_Chamber/generalsubstitutionWithMenu.html)
- With so many keys, might think is secure
- But would be **WRONG**
- Problem is language characteristics
Letters are not equally used

- In English E is by far the most common letter
- Other letters like Z, J, K, Q, X are fairly rare

- Build digrams
- Trigrams ...
Use In Cryptanalysis

- Monoalphabetic substitution ciphers do not change relative letter frequencies
  - Discovered by Arabian scientists in 9\textsuperscript{th} century

1. Calculate letter frequencies for ciphertext
2. Compare counts against known values
   - peaks at: A-E-I triple, NO pair, RST triple
   - valeys at: JK, X-Z
Not even the large number of keys in a monoalphabetic cipher provides security

Idea: *encrypt multiple letters at a time*

**Example:** the **Playfair Cipher**

- Invented by Charles Wheatstone in 1854, but named after his friend Baron Playfair
# Playfair Key Matrix

- 5X5 matrix of letters based on a **keyword**
- Fill in letters of keyword (without duplicates)
- Fill rest of matrix with other letters
- **Example:** using the keyword **MONARCHY**

<table>
<thead>
<tr>
<th>M</th>
<th>O</th>
<th>N</th>
<th>A</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>H</td>
<td>Y</td>
<td>B</td>
<td>D</td>
</tr>
<tr>
<td>E</td>
<td>F</td>
<td>G</td>
<td>I/J</td>
<td>K</td>
</tr>
<tr>
<td>L</td>
<td>P</td>
<td>Q</td>
<td>S</td>
<td>T</td>
</tr>
<tr>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Z</td>
</tr>
</tbody>
</table>
Encrypting and Decrypting

Plaintext is encrypted two letters at a time

1. If a pair is a repeated letter, insert filler like 'X'
2. If both letters fall in the same row, replace each with letter to right (wrapping back to start from end)
3. If both letters fall in the same column, replace each with the letter below it (wrapping to top from bottom)
4. Otherwise each letter is replaced by the letter in the same row and in the column of the other letter of the pair
Security of Playfair

- Security much improved over monoalphabetic
- Since have $26 \times 26 = 676$ digrams
- **Would need a 676 entry frequency table to analyse**
  - Versus 26 for a monoalphabetic
  - And correspondingly more ciphertext
- Was widely used for many years
  - E.g. by US & British military in WW1
- It *can* be broken, given a few hundred letters
  - Since still has much of plaintext structure
Classic Techniques

- Substitution Ciphers
  - Caesar Cipher
  - Monoalphabetic Ciphers
  - Polyalphabetic Ciphers
    - Vigenere
    - Vernam
    - One Time Pad
- Transposition Ciphers
- Product Ciphers
Polyalphabetic Ciphers

- Improve security using multiple cipher alphabets
- Flatten frequency distribution
- Use key to select which alphabet is used for each letter of the message
  - Use each alphabet in turn
  - Repeat from start after end of key is reached
Vigenere Ciphers

- Simplest polyalphabetic substitution cipher
- Effectively multiple Caesar ciphers
- Key is multiple letters long $K = k_1 \ k_2 \ldots \ k_m$
  - $i^{th}$ letter of key specifies $i^{th}$ alphabet to use

Plaintext $P_1, P_2 \ldots P_n$ → Succession of Caesar ciphers → Ciphertext $C_1, C_2 \ldots C_n$

$C_1 = (P_1 + k_1) \mod 26$
$C_m = (P_m + k_m) \mod 26$
$C_{m+1} = (P_{m+1} + k_1) \mod 26$

...
Example: Vigenere Cipher

- Write the plaintext out
- Write the keyword repeated above it
- Use each key letter as a Caesar cipher key
- Encrypt the corresponding plaintext letter

key #: $3424_{1519}8_{21}43424_{1519}8_{21}43424_{1519}8_{21}4$
key: deceptive deceptive deceptive
plaintext: wearediscoverededsaveyourself
ciphertext: ZICVTWQNGRZGVTWAVZHCQYGLMGJ
Security of Vigenere Cipher

- Have multiple ciphertext letters for each plaintext letter
- Hence letter frequencies are obscured
- *But not totally lost!*
Kasiski Method

- Repetitions in ciphertext give clues to period
  - Find same plaintext an exact period apart
  - Which results in the same ciphertext

  key: deceptive deceptive deceptive
  plaintext: wearediscovered saveyourself
  ciphertext: ZICVTWQNGRZGVTWAVZHCQYGLMGJ

- Suggests key size of 3 or 9
- Then attack each monoalphabetic cipher individually using same techniques as before
Vernam Cipher

- Ultimate defense is to use a key as long as the plaintext
  - With no statistical relationship to it
- Invented by AT&T engineer Gilbert Vernam in 1918
- *Originally proposed using a very long but eventually repeating key*
### Vernam Cipher

- Generalization of Vigenere
- Represent messages in binary format
  - E.g., ASCII (google it): A = 10 (decimal) = 1010 (binary)
- Key has the same length as the ciphertext

![Diagram of Vernam Cipher]

- **Encryption**
  - Key stream generator
  - Key
  - Plaintext $P_i$ → $C_i$ → Ciphertext

- **Decryption**
  - Key stream generator
  - Key
  - Ciphertext $C_i$ → $P_i$ → Plaintext
Vernam Cipher: Example
One Time Pad

- Vernam: the key was very long, but ...
  - Repeatable

- One Time Pad:
  - Generalization of Vernam
  - The key is used **ONLY ONCE**!
  - The only "Perfect security" cryptosystem

- **Problems:**
  - Production of many random keys
  - Distribution of random keys – as long as the message
Classic Techniques

- Substitution Ciphers
  - Caesar Cipher
  - Monoalphabetic Ciphers
  - Polyalphabetic Ciphers
    - Vigenere
    - Vernam
    - One Time Pad

- Transposition Ciphers

- Product Ciphers
Transposition Ciphers

- *Transposition or permutation ciphers*
- Rearrange the letter order
  - Without altering the actual letters used
- Easily recognizable
  - Have the same frequency distribution as the original text
Rail Fence Cipher

- Write message letters out diagonally over a number of rows
- Then read off cipher row by row
- *Example:* write message out as:
  
  ```
  mematrhhtgpry
  etefeteteoaat
  ```
- Resulting ciphertext
  
  `MEMATRHTGPRYETEFETEOAAT`
Row Transposition Ciphers

- Write letters of message out in rows over a specified number of columns.
- Then reorder the columns according to some key before reading off the rows.

Key: 4 3 1 2 5 6 7

attack
postpone
duntill
woamxyz

Ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ
Product Ciphers

- Ciphers using substitutions or transpositions are not secure because of language characteristics
- **Use several ciphers in succession to make harder:**
  - Two substitutions make a more complex substitution
  - Two transpositions make more complex transposition
  - **Product cipher:** a substitution followed by a transposition makes a new much harder cipher
- **This is bridge from classical to modern ciphers**
**Rotor Machines**

- Before modern ciphers, rotor machines were most common complex ciphers in use
- Widely used in WW II
  - German Enigma, Allied Hagelin, Japanese Purple
- Implemented a very complex, varying substitution cipher
- Used a series of cylinders, each giving one substitution, which rotated and changed after each letter was encrypted
- With 3 cylinders have $26^3=17576$ alphabets
Hagelin Rotor Machine
Enigma Rotor Machine
Steganography

- An alternative to encryption
- *Hides existence of message*

```plaintext
qANQR1DBwU4D/TIT68XXuiUQCADfj2o4b4aFYBcWumA7hR1Wvz9rbv2BR6WbEUsyZBIEFtjqCd96qF38sp9IQiJKlNaZfx2GLRWikPZwchUXxB+AA5+lqsG/ELBvRa
c9XefaYpbbAZ6z6LkOQ+eE0XASE7aEEPfdxvZZT37dVyiyyxBRRYNLN8Bphdr2zv/z/9Ak4/OLnLiJRk05/2UNE5Z0a+3IcvITMmfGajvRhkXqocavPOKiin3hv7+Vx88
uLLem2/fQHZhGcQvkqzVqXx8SmNw5gzuuvwJV1WHj9muDGBY0MkjiZIRI7azWnoU93KCnmpR60VO4rDRAS5uGl9fioSvze+q8XqxubaNsgdKkoD+tB/4u4c4tznLfw1L2
YBS+dzFDw5desMFS07JkecAS4NB9jAu9K+f7PTAsesCBNETDd49BTOFFTWuavAfEGLYcPrcn4s3EriUgvL3OzPR4P1chNu6sa3ZJkJTBriDoA3VpnqG3hxqfNyO1qAka
mJJuQ53Ob9ThaFH8YCe/VqUFdw+bQtrAJ6NpjIx/i/x0FfOInhC/bBw7pDLXBFNaXHdILQRQPdmmWskKznOSarq4GjpRTQo4hpCRJJ5aU7tZO9HPTZXF6iRIT0wa47
AR5nvkEKOIAjW5HaDKiJriuWLdtN4OXecWvxFsjR32ebz76U8aLpAK87GZEyTzBx
dV+iH0hwyT/y1cZQ/E5USEp4oKWF4uqquPee1OPeFMBo4CvuGyhZXD/18Ft/53YWIebvdiCqsOoabK3jEfdGEexe63zDf0=MPrf
```
Instead, consider

PRESIDENT'S EMBARGO RULING SHOULD HAVE IMMEDIATE NOTICE. GRAVE SITUATION AFFECTING INTERNATIONAL LAW. STATEMENT FORESHADOWS RUIN OF MANY NEUTRALS. YELLOW JOURNALS UNIFYING NATIONAL EXCITEMENT IMMENSELY

Becomes

PERSHINGSAILS FROM NY JUNE 1
Steganography: How?

- Using invisible ink
- Hiding in LSB in graphic image or sound file

- Hidden image is revealed by removing all but the two least significant bits of each color component and a subsequent normalization [wikipedia]
Steganography

- Has drawbacks
  - High overhead to hide relatively few bits

- Advantage:
  - Obscure the use of encryption
Conclusions

- Secrecy in most of the presented techniques (except one time pads) is based on
  - Knowledge of key
  - Knowledge of the algorithm

- *Staple of Modern Cryptography:*
  - Security should not depend on secrecy of algorithm
  - Algorithms should be public!