Engineering Robust Server Software

Cryptography

Significant portions based on slides from Micah Sherr @ Georgetown
Cryptography

f(Leftover Food in HH 218)
= Al481manj417a@#1naL

f^{-1}(Al481manj417a@#1naL)
=Leftover Food in HH 218

This is an example of..
A: Confidentiality
B: Integrity
C: Authentication
D: Availability

Alice

Bob

??

Eve
• Modern cryptography: secure; advanced math
• Classical cryptography: insecure; simple math
Cryptography Terms

- **Cryptosystem**: method of disguising (encrypting) plaintext messages so that only select parties can decipher (decrypt) the ciphertext.
- **Cryptography**: the art/science of developing and using cryptosystems.
- **Cryptanalysis**: the art/science of breaking cryptosystems.
- **Cryptology**: the combined study of cryptography and cryptanalysis.

\[
\text{Encrypt: } E_k(\text{Plaintext}) \rightarrow \text{Ciphertext}
\]

\[
\text{Decrypt: } D_k(\text{Ciphertext}) \rightarrow \text{Plaintext}
\]
Kerckhoffs’ Principles

• Kerckhoffs’ principles [1883]:
  • Assume Eve knows cipher algorithm
  • Security should rely on choice of key
  • If Eve discovers the key, a new key can be chosen
• Opposite of “security by obscurity”
  • Idea of keeping algorithm secret
• Why not security by obscurity?
  • Compromised? Destroyed. (vs one key lost-> make new one)
  • Algorithms relatively easy to reverse engineer
Shannon's Principles

- Two important principles for modern/practical systems:
  - **Confusion**: each bit of cipher text depends on many key bits
  - **Diffusion**: flipping one bit of plaintext should alter many (~1/2) of ciphertext
Shannon's Principles

Two important principles for modern/practical systems:

- **Confusion**: each bit of cipher text depends on many key bits
- **Diffusion**: flipping one bit of plaintext should alter many ($\sim \frac{1}{2}$) of ciphertext
Shannon's Principles

- Two important principles for modern/practical systems:
  - **Confusion**: each bit of cipher text depends on many key bits
  - **Diffusion**: flipping one bit of plaintext should alter many ($\sim \frac{1}{2}$) of ciphertext
Classical Cryptography

FIRST LEGION ATTACK EAST FLANK

I LUVW OHJLRQ DWWDFN HDVW IODQN

• Simple/ancient classical crypto system:
  • Caesar Cipher: named after Julius Caesar
• Key: number of letters to shift by (in this case 3)
Breaking Caesar

• You may have previously written a program to crack this
  • 'e' is most common in English
  • Find most common in ciphertext -> probably 'e'
Spaces and Punctuation

FIRST LEGION ATTACK EAST FLANK

ILUVW OHJLRQ DWWDFN HDVW IODQN

• Quick side note:
  • I'm writing spaces in the plain text/cipher text (readability of examples)
  • Would not really do (makes much easier)
• Either encrypt spaces/punctuation too (computers) or
• Remove from plaintext before encrypting
Vigenère Cipher

FIRST LEGION ATTACK EAST FLANK

\[
\text{drago ndrago ndrago ndra godra}
\]

\[+3 \downarrow +1 \downarrow +7 \downarrow +0\]

\[
\text{IZRYH OVGOCIA RTZOPN EGGG WLGBX}
\]

- Key is now a vector of numbers, e.g., \((3, 17, 0, 6, 14, 13)\)
  - Usually represented by a word, e.g. "dragon"
Vigenère Security?

• Vigenère seemed unbreakable for a few centuries
  • Long enough key: smooth out frequencies
• Easy to break if you can determine key length
  • Key length 10?
    • Take letters 0, 10, 20,… frequency count
    • 1, 11, 21, 31, … frequency count. etc.
• Try many different key lengths?
  • Time consuming with pencil and paper
  • Easy with computer…
  • Vigenère broken even before computers
Vigenère

• Vigenère is what many novices make up on their own
  • Seems hard to break!
  • …but is actually easy.

• Important lesson:
  • Do not try to make up your own crypto
  • It is very hard to do correctly

• But what if…
  • Your key were as long as your message
  • And you only used it for one message?
One Time Pad

• One Time Pad
  • $E_k(M) = M \oplus K$
  • Length of $K$ is equal to Length of $M$ (same number of bytes)
  • NEVER re-use $K$
    • Re-using even once destroys guarantees

• Gives perfect secrecy
  • Without knowledge of key, guessing $M$ is just random guessing

• Difficult in practice
  • Must exchange keys securely, and cannot re-use
One Time Pad

Alice and Bob are in HQ

They generate some OTPs
One Time Pad

Alice and Bob are in HQ

They generate some OTPs

Now, Alice goes into the field
One Time Pad

Alice and Bob are in HQ
They generate some OTPs
Now, Alice goes into the field

$C_1 = M_1 \oplus K_1$

$M_2 = C_2 \oplus K_2$

$M_1 = C_1 \oplus K_1$

$C_2 = M_2 \oplus K_2$
AES

- Advanced Encryption Standard (Rijndael)
  - Symmetric key (Alice and Bob have same key)
  - Replaced DES as accepted symmetric key standard block cipher
  - "Nobody ever got fired for using AES"

10 rounds for 128-bit key
12 rounds for 192-bit key
14 rounds for 256-bit key
AES: Add Round Key

**Input**

<table>
<thead>
<tr>
<th>00</th>
<th>01</th>
<th>02</th>
<th>03</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>30</td>
<td>31</td>
<td>32</td>
<td>33</td>
</tr>
</tbody>
</table>

**Round key**

1F  3C  09  AB  
2C  D9  11  AA  
FC  00  99  21  
38  8E  07  4C  

**Output**

1F  3D  0B  A8  
3C  C9  03  B9  
DC  21  BB  02  
08  BF  35  7F  

**Add Round Key**

- XOR input data with round key

**What is a round key?**

- At the start, key is expanded into 11 (13, or 15) round keys
- Each round key is used once
AES: Substitute Bytes

**Substitute Bytes**
- Look up input in substitution table ("sbox").
- Substitution is 1-to-1 (each value appears once in the table).
- AES's Sbox designed with important mathematical properties

<table>
<thead>
<tr>
<th>Input</th>
<th>00</th>
<th>01</th>
<th>02</th>
<th>03</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output</th>
<th>63</th>
<th>7C</th>
<th>77</th>
<th>7B</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>82</td>
<td>C9</td>
<td>7D</td>
<td></td>
</tr>
<tr>
<td>B7</td>
<td>FD</td>
<td>93</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>04</td>
<td>C7</td>
<td>23</td>
<td>7F</td>
<td></td>
</tr>
</tbody>
</table>
**AES: Shift Rows**

- **Shift Rows**
  - Shift the (ith) row left by i positions
  - Row 0: no change
  - Row 1: shift bytes left one position

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>00 01 02 03</td>
<td>00 01 02 03</td>
</tr>
<tr>
<td>10 11 12 13</td>
<td>11 12 13 10</td>
</tr>
<tr>
<td>20 21 22 23</td>
<td>22 23 20 21</td>
</tr>
<tr>
<td>30 31 32 33</td>
<td>33 30 31 32</td>
</tr>
</tbody>
</table>

Note: “Shift” here is circular; really more of a “rotate”
AES: Mix Columns

- **Mix Columns**
  - Take each input column
  - Multiply it by a matrix as polynomial
  - Result is column in output

### Input

<table>
<thead>
<tr>
<th></th>
<th>00</th>
<th>01</th>
<th>02</th>
<th>03</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td></td>
</tr>
</tbody>
</table>

### Output

<table>
<thead>
<tr>
<th></th>
<th>XX</th>
<th>XX</th>
<th>XX</th>
</tr>
</thead>
</table>

$$\begin{bmatrix}
01 & 11 & 21 & 31
\end{bmatrix} \times \begin{bmatrix}
2 & 3 & 1 & 1 \\
1 & 2 & 3 & 1 \\
1 & 1 & 2 & 3 \\
3 & 1 & 1 & 2
\end{bmatrix} = \begin{bmatrix}
\text{XX} \\
\text{XX} \\
\text{XX}
\end{bmatrix}$$
AES: Confusion and Diffusion?

• Does AES have good confusion and diffusion?
  • A: Good confusion and good diffusion
  • B: Good confusion, poor diffusion
  • C: Poor diffusion, good confusion
  • D: Poor diffusion and poor confusion
Difficulty: Key Distribution

Bwahahah!
Diffie-Hellman Key Exchange

Here are two prime numbers: \( g \) and \( p \)

Here is the value of \( A \)

Here is the value of \( B \)

I also know \( g \) and \( p \)

I also know \( A \)

I also know \( B \)

\[
S = B^x \mod p
\]

\[
A = g^x \mod p
\]

\[
\text{Secret: } x
\]

\[
S = A^y \mod p
\]

\[
B = g^y \mod p
\]

\[
\text{Secret: } y
\]
Diffie-Hellman Key Exchange

\[ S = B^x \mod p \]
\[ A = g^x \mod p \]
Secret: \( x \)

Alice:
\[ S = (g^y \mod p)^x \mod p \]

Bob:
\[ S = (g^x \mod p)^y \mod p \]

These are equal

Eve has to solve the \textit{discrete logarithm (hard)} problem to recover \( x \) or \( y \) (and thus compute \( S \))

I also know \( g \) and \( p \)
I also know \( A \)
I also know \( B \)
Diffie-Hellman Key Exchange

This scheme works securely as long as…
A: ...Alice and Bob pre-share at least $\lg(p)$ bits
B: ...$p$ is at least 64 bits
C: ...Eve can only listen, not alter messages
D: ...Eve does not have many GPUs

$I$ also know $g$ and $p$
$I$ also know $A$
$I$ also know $B$

$A = g^x \mod p$
$B = g^y \mod p$

$S = B^x \mod p$
$S = A^y \mod p$
Diffie-Hellman Key Exchange

\[ S = B^x \mod p \]
\[ A = g^x \mod p \]
Secret: \( x \)

\[ S = A^y \mod p \]
\[ B = g^y \mod p \]
Secret: \( y \)

All of this assume Eve can only listen. What if Eve can change the messages?

I also know \( g \) and \( p \)
I also know \( A \)
I also know \( B \)
Man In the Middle (MITM) Attack

Here are two numbers: \( g \) and \( p \)

Here is the value of \( A \)

(I also know \( g \) and \( p \))

I also make: \( z \)

\( C = g^z \mod p \)

\( S_{Alice} = A^z \mod p \)

\( S_{Bob} = B^z \mod p \)

Here is the value of \( B \)

(Replace \( A \) with \( C \))

(Replace \( B \) with \( C \))

I also know \( g \) and \( p \)

\( S = C^y \mod p \)

\( A = g^x \mod p \)

Secret: \( x \)

Secret: \( y \)
Man In the Middle (MITM) Attack

At this point, Eve has exchanged (different) keys with Alice and Bob.

Eve can now decrypt, view (and alter) a message, then encrypt it and send it along.

I also know \( g \) and \( p \)
I also make: \( z \)

\( S_{\text{Alice}} = A^z \mod p \)
\( S_{\text{Bob}} = B^z \mod p \)
Man In The Middle Attack

• Alice needs to know that she is receiving Bob's message unchanged
  • Which security principles are these?
    • A: Confidentiality and Integrity
    • B: Integrity and Authentication
    • C: Authentication and Availability
    • D: Integrity and Confidentiality
Man In The Middle Attack

- Alice needs to know that she is receiving Bob's message unchanged
  - Which security principles are these?
- **Integrity**: don't let Eve tamper with things
- **Authentication**: message actually came from Bob (not someone else)
- Cryptographic solution: **signatures**
  - Bob will generate a cryptographic validation of the message
  - (and that it was from him)
- For this, we need **public key** cryptography: e.g., RSA
  - Also called **asymmetric key** cryptography
Public Key Cryptography

Bob picks two random primes: $p$ and $q$

Bob computes $n = pq$

Number of bits in $n$ is key length

Bob computes $\lambda(n) = \text{lcm}(p-1,q-1)$

Bob picks $e$ st. $1 < e < \lambda(n)$
- Where $e$ and $\lambda(n)$ are coprime
Bob solves for $d = e^{-1} \mod \lambda(n)$
- That is $ed = 1 \mod \lambda(n)$

Bob publishes his **public key** $(e,n)$
Bob keeps private key $d$, secret (he also keeps $n$)
Encryption

Bob's public key: \((e,n)\)

Alice can send a message \((M)\) to Bob:
She computes \(C = M^e \mod n\)
and sends \(C\) to Bob

Bob computes \(M = C^d \mod n\)
Private key: \((d,n)\)

Eve does not have \(d\).
She cannot recover the message from \((e,n)\)
Alice computes $M' = S^e \mod n$
checks that $M = M'$

Bob's public key: $(e,n)$

Bob wants Alice to know message $M$ is from him. He computes
$S = M^d \mod n$
and sends $M$ and $S$ to Alice

Eve cannot fake this signature because she does not have $d$

Bob's public key: $(e,n)$

Private key: $(d,n)$
...But Did We Fix Anything?

Great if Alice can get Bob's public key in a trusted way (then again, she could get an AES key that way)... but if not...?

Bob publishes his **public key** \((e, n)\)

What if Eve intercepts and convinces Alice of a fake key?
...But Did We Fix Anything?

Suppose Alice already trusts Ted.

Bob can take his public key (and proof that he is Bob) to Ted

Ted's public key

Now Bob can send the signed key to Alice

Ted can sign Bob's key: he can make the message "Bob's key is (e,n)" and cryptographically sign it with his key

Bob's key is (e,n)

— Ted
Certificates

- Certificates: electronic documents attesting to ownership of a key
  - Cryptographically signed by Certificate Authority (CA)
    - To be meaningful, CA needs to be trusted
    - Trust may be done in several steps: A signs B, B signs C.
  - Generally contains expiration date
- https uses certificates
  - Your computer trusts certain CAs
Side Channels

• AES + RSA: hard to break algorithmically
  • VERY Difficult to recover key, or decipher message without key
• Can be attacked by side channels
  • Information leaked from physical characteristics of execution
  • E.g., power, temperature, memory access pattern, instruction timing…
Side Channel Example

- AES: some steps sped up with 4KB lookup table
  - Indexed by input to that stage
  - Tell which cache block \(\rightarrow\) gain much information \(\rightarrow\) recover key
- Attacker runs code on same core
  - Measures time to perform loads
  - Determines hits/misses in cache
  - Figures out "victim"'s memory access pattern
- Similar attacks on RSA based on multiplication patterns
  - Timing, power, …
Cryptography Wrap Up

• Quick introduction to basics of cryptography
  • Classical systems: weak
  • AES: symmetric key
  • RSA: public key (asymmetric key)

• A few attacks:
  • MITM
  • Side channels

• Idea of signing + certificates