Lecture 5 - Cryptography

CMPSC 443 - Spring 2012
Introduction Computer and Network Security
Professor Jaeger
www.cse.psu.edu/~tjaeger/cse443-s12/
A historical moment ...

The *enigma* machine was used to secure communication of German military throughout the Second World War ...

... and it changed the course of human history.
Intuition

• Cryptography is the art (and sometimes science) of secret writing
  – Less well know is that it is also used to guarantee other properties, e.g., authenticity of data
  – This is an enormously deep and important field
  – However, much of our trust in these systems is based on faith (particularly in efficient secret key algorithms)

• Cryptographers create ciphers - Cryptography
• Cryptanalyst break ciphers - Cryptanalysis

The history of cryptography is an arms race between cryptographers and cryptanalysts.
A cryptosystem is a 5-tuple consisting of

\((E, D, M, K, C)\)

Where,

- \(E\) is an \textit{encryption} algorithm
- \(D\) is an \textit{decryption} algorithm
- \(M\) is the set of \textit{plaintexts}
- \(K\) is the set of \textit{keys}
- \(C\) is the set of \textit{ciphertexts}

\[E : M \times K \rightarrow C \quad D : C \times K \rightarrow M\]
What is a key?

• A key is an input to a cryptographic algorithm used to obtain confidentiality, integrity, authenticity or other property over some data.
  – The security of the cryptosystem often depends on keeping the key secret to some set of parties.
  – The *keyspace* is the set of all possible keys
  – *Entropy* is a measure of the variance in keys
    • typically measured in bits

• Keys are often stored in some secure place:
  – passwords, on disk keyrings, ...
  – TPM, secure co-processor, smartcards, ...

• ... public keys are not, e.g., certificates, but they are different...
Transposition Ciphers

- Scrambles the symbols to produce output
- The key is the permutation of symbols
Substitution Ciphers

• Substitutes one symbol for another (codebook)
• The key defines the substitution (Sbox)
Encryption algorithm

• Algorithm used to make content unreadable by all but the intended receivers

\[ E(\text{key}, \text{plaintext}) = \text{ciphertext} \]
\[ D(\text{key}, \text{ciphertext}) = \text{plaintext} \]

• Algorithm is public, key is private

• Block vs. Stream Ciphers
  – Block: input is fixed blocks of same length
  – Stream: stream of input
Example: Caesar Cipher

- Substitution cipher
- Every character is replaced with the character three slots to the right

| 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 |

- Q: What is the key?

SECURITY AND PRIVACY
VHFXULWBDQGSULYDFB
Cyptanalyze this ....

"AVGGNALYVBAF"
Cryptanalysis of ROTx Ciphers

• Goal: to find plaintext of encoded message
• Given: ciphertext
• How: simply try all possible keys
  – Known as a brute force attack

1 T F D V S J U Z B M E Q S J W B D Z
2 U G E W T K V A C N F R T H X C E A
3 W H F X U L W B D Q G S U L Y D F B

SECURITY AND PRIVACY
Shared key cryptography

- Traditional use of cryptography
- Symmetric keys, where a single key \((k)\) is used is used for \(E\) and \(D\)

\[
D( k, E( k, p ) ) = p
\]

- All (intended) receivers have access to key
- **Note**: Management of keys determines who has access to encrypted data
  - E.g., password encrypted email
- Also known as symmetric key cryptography
The one-time pad (OTP)

• Assume you have a secret bit string $s$ of length $n$ known only to two parties, Alice and Bob
  – Alice sends a message $m$ of length of $n$ to Bob
  – Alice uses the following encryption function to generate ciphertext $c$
    \[
    \forall i = 1 \text{ to } n : c_i = m_i \oplus s_i
    \]
  – E.g., XOR the data with the secret bit string
  – An adversary Mallory cannot retrieve any part of the data

• Simple version of the proof of security:
  – Assume for simplicity that value of each bit in $m$ is equally likely, then you have no information to work with.
Data Encryption Standard (DES)

- Introduced by the US NBS (now NIST) in 1972
- Signaled the beginning of the modern area of cryptography
- Block cipher
  - Fixed sized input
- 8-byte input and a 8-byte key (56-bits+8 parity bits)
DES Round

- Initial round permutes input, then 16 rounds
- Each round key ($k_i$) is 48 bits of input key
- Function $f$ is a substitution table ($s$-boxes)
Cryptanalysis of DES

• DES has an effective 56-bit key length
  – Wiener: $1,000,000 - 3.5 hours (never built)
  – July 17, 1998, the EFF DES Cracker, which was built for less than $250,000 < 3 days
  – January 19, 1999, Distributed.Net (w/EFF), 22 hours and 15 minutes (over nearly 100,000 machines)
  – We all assume that NSA and agencies like it around the world can crack (recover key) DES in milliseconds

• What now? Give up on DES?
Variants of DES

DESX (two additional keys \(\sim= 118\)-bits)
Triple DES (three DES keys \(\sim= 112\)-bits)
Keys \(k_1, k_2, k_3\)

\[
c = E( k_3, D( k_2, E( k_1, p)))
\]
Advanced Encryption Standard (AES)

• Result of international NIST bakeoff between cryptographers
  – Rijndael (pronounced “Rhine-dall”)
  – Now supersedes DES
  – Actually AES is the standard, and the algorithm is called Rijndael, although both are often called AES
  – Used in many applications now (e.g., wireless standard 802.11i)
Hardness

• Functions
  – Plaintext $P$
  – Ciphertext $C$
  – Encryption key $k_e$
  – Decryption key $k_d$

\[ D(k_d, E(k_e, P)) = P \]

• Computing $C$ from $P$ is hard, computing $C$ from $P$ with $k_e$ is easy
• Computing $P$ from $C$ is hard, computing $P$ from $C$ with $k_d$ is easy
Key size and algorithm strength

- Key size is an oft-cited measure of the strength of an algorithm, but is strength strongly correlated (or perfectly correlated with key length)?
  - Say we have two algorithms, A and B with key sizes of 128 and 160 bits (the common measure)
  - Is A less secure than B?
  - What if A=B (for variable key-length algorithms)?
Take Away

• Cryptography
  – Secret writing

• Between parties who share a secret key
  – Should the cryptographic algorithm be secret?

• Choosing keys is “key”
  – Keyspace -- number of possible keys ($2^n$ for $n$ bits)
  – Entropy -- variance among keys (depends on generation method: want to be $n$ bits)

• Security is dependent on algorithm and key size
  – Difficult to compare between algorithms in many cases