Introduction to Cryptography

Christian Engman

Big O Theory Club

September 15, 2023

Table of Contents

Introduction

Symmetric-Key Ciphers

Public-Key Cryptosystems

Hashing Algorithms

Signatures

Introduction

- Cryptography is the science of secure communication. From military communications to bank information to private text messages, our modern world depends on the existence of secure protocols that ensure that transmitted information is only seen by its intended recipient.
- Secret codes have been used since ancient times, but the field of cryptography has improved exponentially in security and utility alongside advancements in mathematics (specifically number theory and linear algebra), the birth of computer science, and the wide-scale availability of comping hardware.

Goals of Cryptography

- Ensuring Confidentiality.
- Maintaining Data Integrity.
- Authenticating individuals and messages.
- Authorizing actions by certain users.
- Certifying information.
- Allowing ease of communication.

Symmetric-Key Ciphers

- Symmetric-key ciphers have a single 'key' that can be used to encrypt and decrypt information.
- The oldest type of cryptosystems are built on this model.
- Basic examples include substitution ciphers (including the Caeser Cipher allegedly used by the Romans) and Vigenere Ciphers.
- These types of schemes are widely used today.

One-Time Pads

- Simple and unbreakable in theory.
- uses a plaintext P and a key K that is the same length as C, and obtains a secret ciphertext $C = P + K$.
- Recipient can simply decrypt by calculating $C - K$, if they know what K is.
- If K is sufficiently random (From a uniform or other high-entropy distribution), C appears completely random and has no association with P.
- Drawbacks: K needs to be agreed on beforehand, and is no longer secure if it is reused.

Block Ciphers

- Block ciphers take in fixed size plaintext (e.x. 64 bits) and use a key K to obtain a ciphertext of the same size.
- The same key can be used to go from the Ciphertext to the plaintext.
- Ideal security: the only way to break the cipher is to go through every possible key, even if you are able to encrypt and decrypt text arbitrarily, i.e. no statistical dependence between input and output.
- Common modern block ciphers: AES ($n = 128$, $k = 128, 192, 256$) and DES ($n = 64$, $k = 56$) / 3DES.

AES

- Most widely used symmetric key cipher today, with no known practical attacks when implemented correctly.
- Block size is 128, keys can be 128, 192, or 256 bits.
- Repeats a series of steps 10-14 times, depending on the key size:
- SubBytes, ShiftRows, MixColumns (skipped on last round), AddRoundKey
- Reversible by going through rounds in reverse order.

Modes of Operation

- What happens when we want to encrypt something larger than the size of a block?
- The first step is to break the plaintext up into chunks that are the same as the block size, but applying cipher to each block individually is not the best approach.
- Common methods: Cipher Block-Chaining, CCM, Galois Counter Mode.
- Prevents common blocks from always being encrypted as the same thing.

Cipher Block Chaining (CBC)

- CBC: Start with an n bit IV, along with a t block long plaintext $p_1, \ldots, p_t$
- Encryption:

$$c_0 \leftarrow IV, \; c_j \leftarrow E_K(c_{j-1} \oplus p_j)$$

- Decryption:

$$c_o \leftarrow IV \; p_j \leftarrow c_{j-1} \oplus E_K^{-1}(c_j)$$

Stream Ciphers: An alternative to Block Ciphers

- A key K is used to generate a long stream of pseudorandom bits called a keystream, which is them combined with the plaintext to obtain the ciphertext.
- Common examples include Salsa/ChaCha20 and RC4.
- Certain Block Cipher operating modes behave like stream ciphers.

Public-Key Cryptosystems

- Modern symmetric-key schemes are secure and fast, however they do not allow for communication betweeen two parties that have not agreed on a key beforehand.
- The modern internet requires our ability to establish a secure connection, without transmitting unencrypted keys over a network or having to agree on keys physically.
- Public Key Cryptosystems use a public-private key pair, where any inidividual can encrpyt a message with the public key, but it can only be decrpyted with the private key.
- This allows for commucation between users that have not pre-autheticated.
- Common examples include RSA and ECC

RSA

- RSA uses a public key pair (e, n) and a private key pair (d, n) to enable asymmetric encryption.
- Pick 2 large prime numbers (usually 512-1024 bits each), p and q.
- Then, let $n = pq$
- pick e as a mid-sized prime number ($e = 65537$ is common)
- then, d is calculated as $e^{-1} \mod \phi(n) = (p-1)(q-1)$

RSA

- Then suppose we want to encypt a plaintext m. The first step is to encode m as a number smaller than n. This may require breaking up m into smaller pieces.
- Then, we encrpyt m as $c = m^e \mod n$
- given c, the recipient can decrypt c as $m = c^d \mod n$.

RSA

- RSA is still considered secure and is widely used to this day. Howvever, the implementation details matter a lot.
- RSA security is dependent on the fact that n cannot be factored into p and q. This is generally true, however, the choice of p and q matters a lot in this situation, and generating large primes can be quite difficult.
- RSA also requires a large amount of computation to encrypt, decrpyt, and generate keys, especially relative to symmetric-key ciphers like AES.
- In modern network protocols, variants of RSA are used to agree on a key between two clients, and then the key is used for AES-encrpyted communication.

Other Public-Key Cryptosystems

- Elliptic Curve methods are popular in modern times, as they require smaller keys than RSA. They are much more algebraically complicated than RSA, however, and still depend on the hardness of factoring a number into large primes.
- Diffie-Hellman and its variants popular for key exchange, and its derivative ElGamal is used as a public-key cryptosystems.
- Non-number theoretic cryptosystems like LWE, which depends on the hardness of a problem called lattice reduction, have appeared in response to the apparent insecurity of RSA and ECC against quantum computing.

Hashing

- Cryptographic hashes are widely used for data integrity verification and authentication.
- Used to store things like passwords.
- Model: given an arbitrary-size plaintext m, output a hash $h = E(m)$, such that E will always output h when m is given to it.
- E should give pseudorandom outputs to reduce the possibility of collisions.
- Common examples include SHA-256 and BCrypt.
- Can be used to generate MACs (Message Authentication Codes) that are sent along with an encrypted message to prevent it from being tampered with.

Signing Messages With RSA

- Suppose Alice wants to send a message to Bob, so that Bob knows its from her.
- Alice first computes a hash of her message $h = E(M)$. Then, she takes her private RSA key pair (d, n), and sends bob $s = h^d \mod n$
- Then, when Bob receives her message, he can use her public key pair (e, n) to calculate $s^e \mod n = h$, and then verify that the hash of m is the same as h.

Other Topics

- TLS (Transport Layer Security) specifies the way HTTPS and other protcols communicate securely. The specification includes the full suite of cryptographic tools that are used on the modern internet.
- Other cryptographic tools include error correcting codes (Reed-Solomon, BCH) and compression algorithms.