Steganography: Techniques for Hiding Information in Digital Media

- Discriminators
- Covert Channels
- Port Knocking
- Onion Routing
- Dictionary

Steganography

Steganography is a data-hiding technique

- Linguistic steganography
- Technical Steganography

Methods used for Steganography:

- Use Noise
- Spreading out Information
- Define a Structural profile, e.g. make it look like a normal document to hide information
- Change the order, e.g. the order of a shopping list
- Split information, e.g. first word on every page
- Hide the source, e.g. Tor Onion Routing
- Define a statistical profile, e.g. trick computer for statistical analysis
- Replace randomness, e.g. add imperfections in places in a picture

JPEG Data Hiding

Typically an image is split up into 8 × 8 pixel blocks. A frequency analysis is then conducted on each of these pixel blocks using the Discrete Cosine Transform (DCT), which converts the spatial domain data into frequency components to allow for more efficient compression and data hiding.

File Carving

File carving is a forensic data recovery technique used to extract files from raw disk data (like hard drives, memory dumps, or storage images) without relying on the file system metadata.

Formats used within steganography

- Huffman encoding: A lossless data compression algorithm that assigns shorter binary codes to more frequent symbols and longer codes to less frequent ones
    - Plays a role in reducing payload size as hidden messages are compressed before embedding
- XOR operations can also be used to encrypt the message before hiding; this is a lightweight encryption to hide the meaning

Compression methods within steganography

- Lempel-Ziv (LZ): A family of more advanced dictionary-based compression algorithms, and the basis for popular formats like ZIP, PNG, and GIF. Instead of repeating characters, LZ looks for repeated patterns or sequences and replaces them with pointers to earlier occurrences
- RLE: A very simple lossless compression algorithm, compressing data by replacing repeated values with a single value and count

Zero-knowledge Proofs

A Zero-Knowledge Proof (ZKP) is a cryptographic method where:

- One party (the "prover") can prove to another party (the "verifier") that they know something, without revealing the thing itself.
- Two-billionaire problem: How can the billionaires discuss this without revealing how much they have and revealing the one with the most money?
    - Can they compare who is richer without revealing any more information than necessary?

    - Alice wants to embed a message in an image and send it to Bob, but she wantes to avoid revealing that a message even exists, especially to third party observers
    - No one should be able to prove a message was embedded, unless they have the correct key or shared knowledge

Secret Shares

Secret shares are pieces of a secret that are distributed among a group of people in such a way that no single person holds the entire secret, but together they can reconstruct it.

- Shamir’s secret sharing is an example of this technique - https://en.wikipedia.org/wiki/Shamir's_secret_sharing

Hiding the Source

A proxy can be used to ‘hide’ the original source of the message exchange, and provide anonymity to the source

- Tor Onion routing is an example of this method

Covert Channels

Two methods of establishing a covert channel: a channel where messages are exchanged in secret:

- Storage Covert channels are where one process uses direct data writing, whilst another process reads the data. It generally uses finite system resources that are shared between entities with different priveleges
- Covert timing channels use the modulation of certain resources, such as the CPU timing, in order to exhange information between processes

ChatGPT Q&A

https://chatgpt.com/share/6817060c-aaf4-8001-b664-6779336e1539