Secure communications below the hearing threshold

Improved approaches for auditive steganography

Marcus Nutzinger Rainer Poisel Jürgen Wurzer

Institute of IT Security Research
St. Pölten University of Applied Sciences, Austria

27th Chaos Communication Congress
December 28, 2010

Introduction

What? – Cryptology, Cryptography and Steganography

Cryptography

- Study and practice of hiding information
- Protection of secret data
- Data transfer obvious

Steganography

- Science of covert communication
- Conceals the existence of secret information

Cryptology

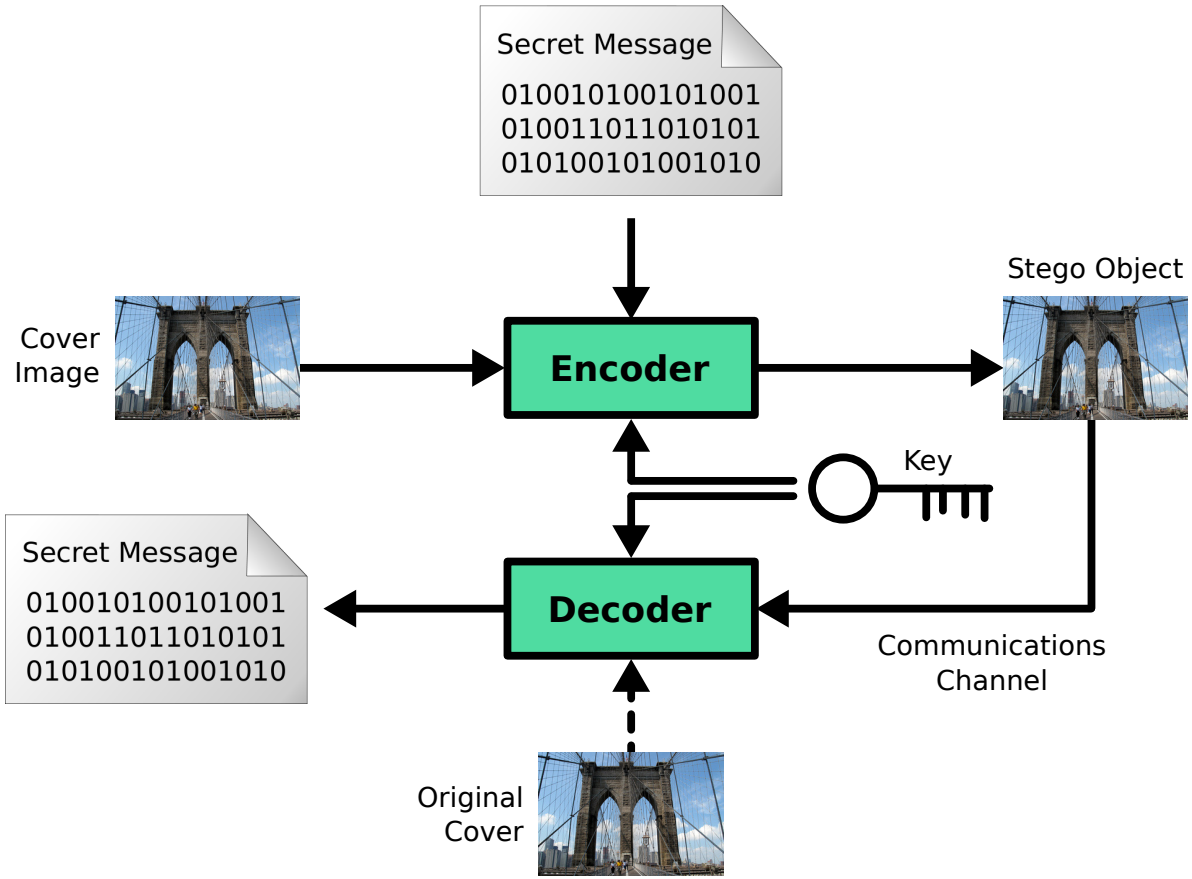
Science that incorporates both cryptography and cryptanalysis.

Introduction

What? – Steganography in Brief

Different cover media

- Technical steganography: digital images, audio and video files



Introduction

Why? – Filtering of Internet Services and Applications

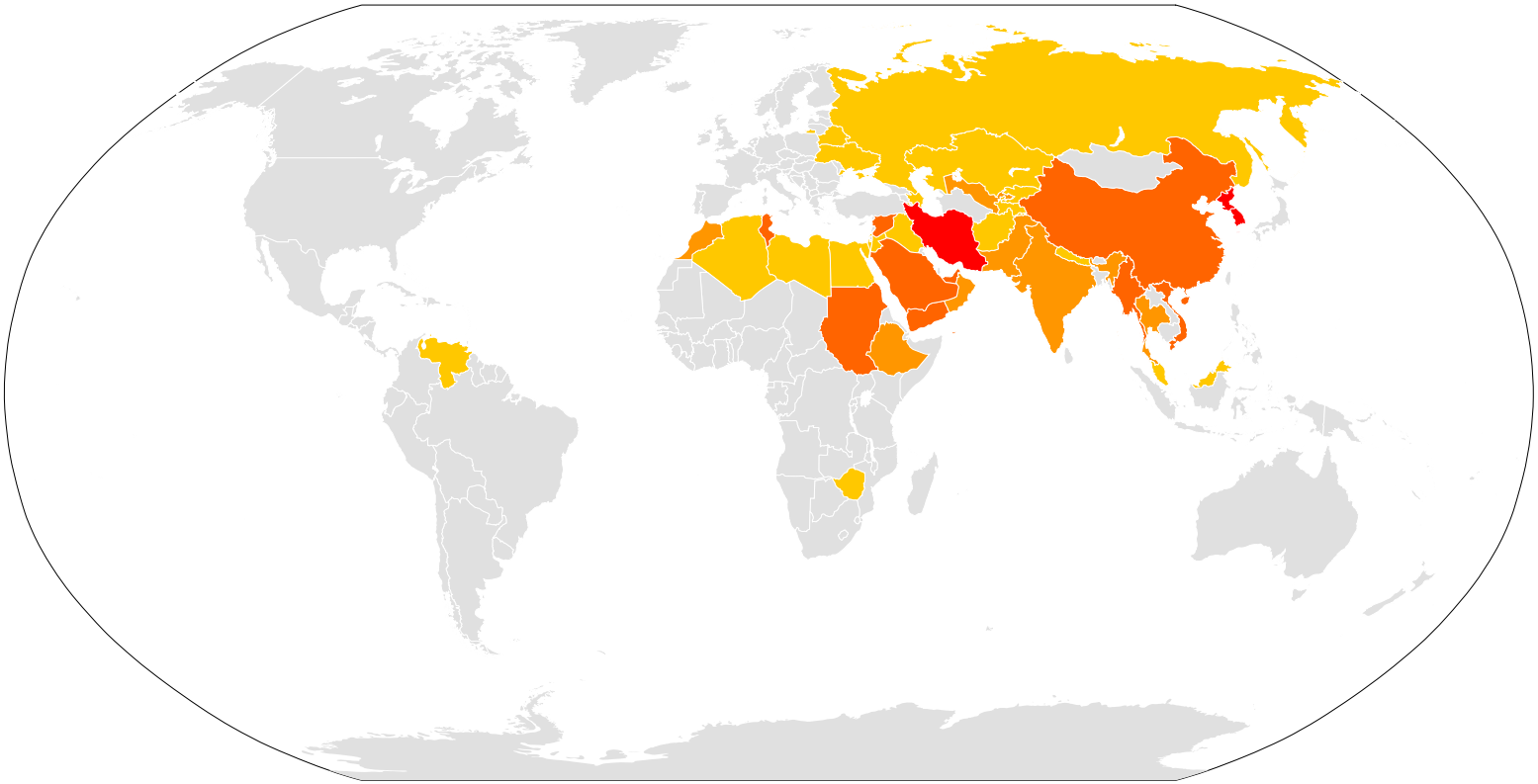
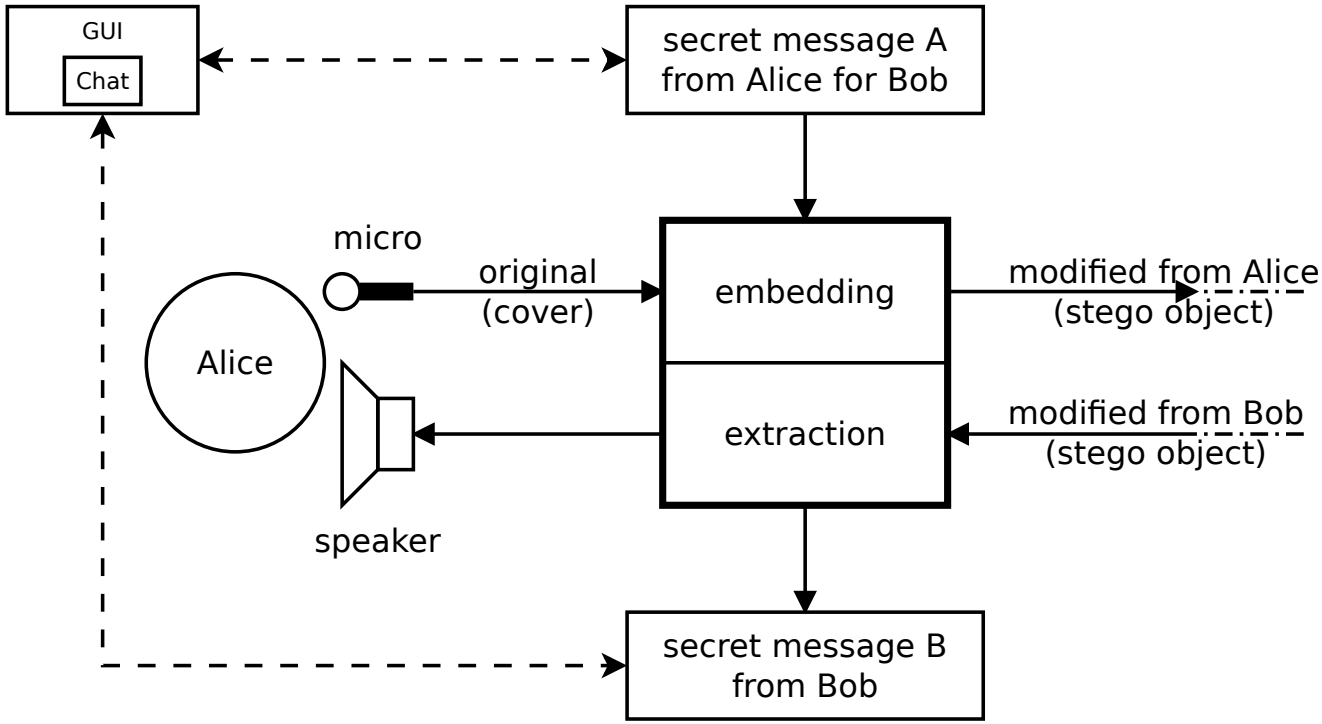


Figure: Worldmap of internet filtering (Source: OpenNet Initiative, 2010)

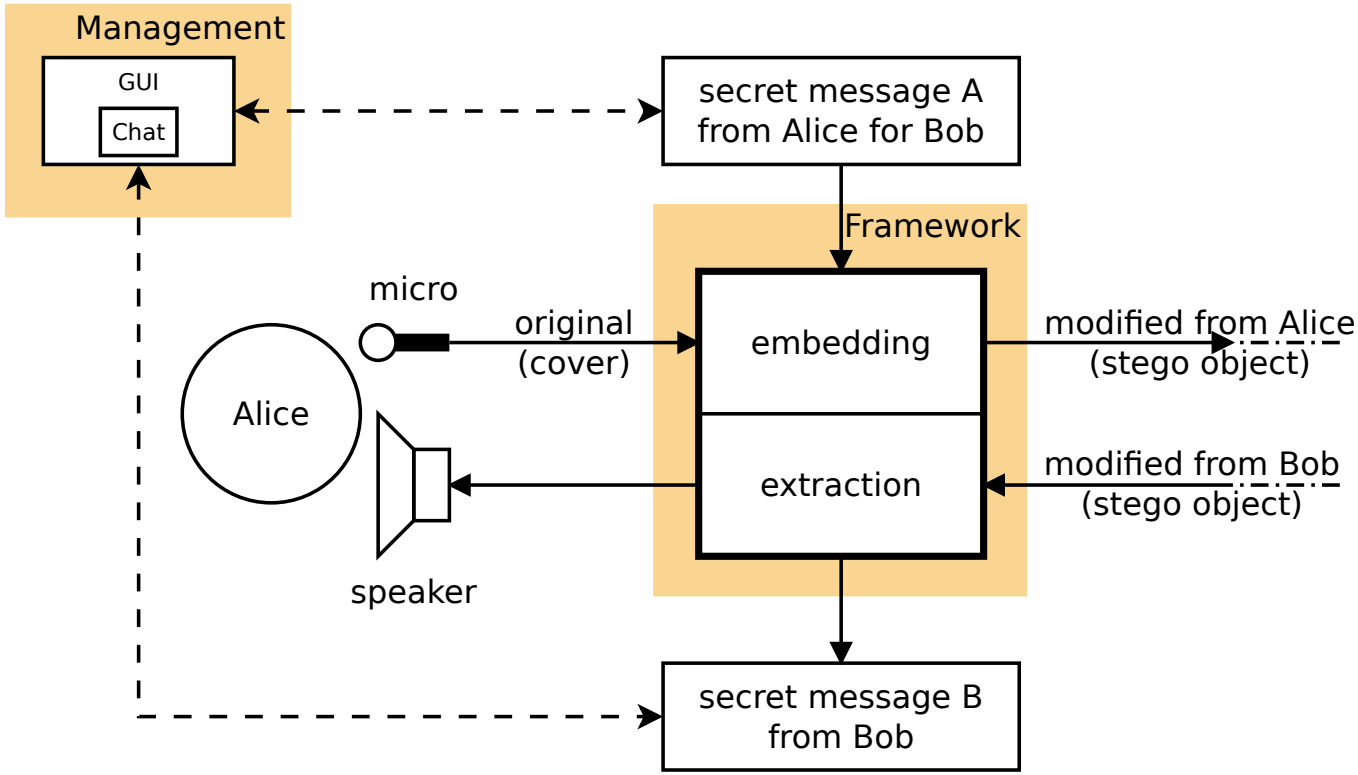
Functional Overview

How? – The Framework



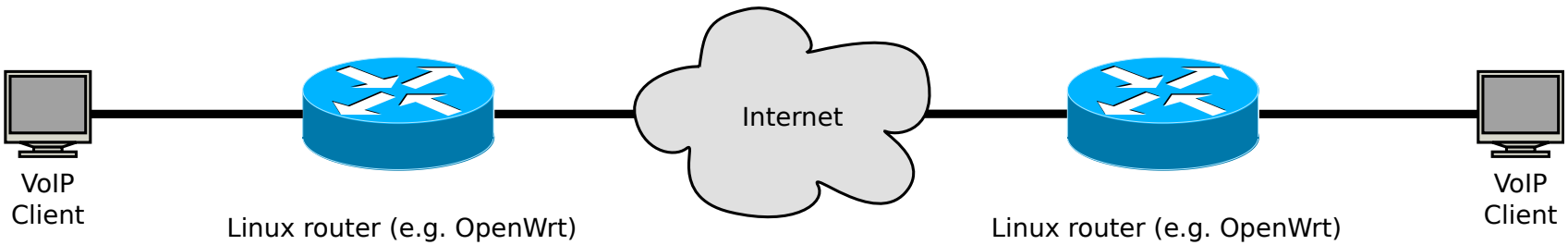
Functional Overview

How? – The Framework



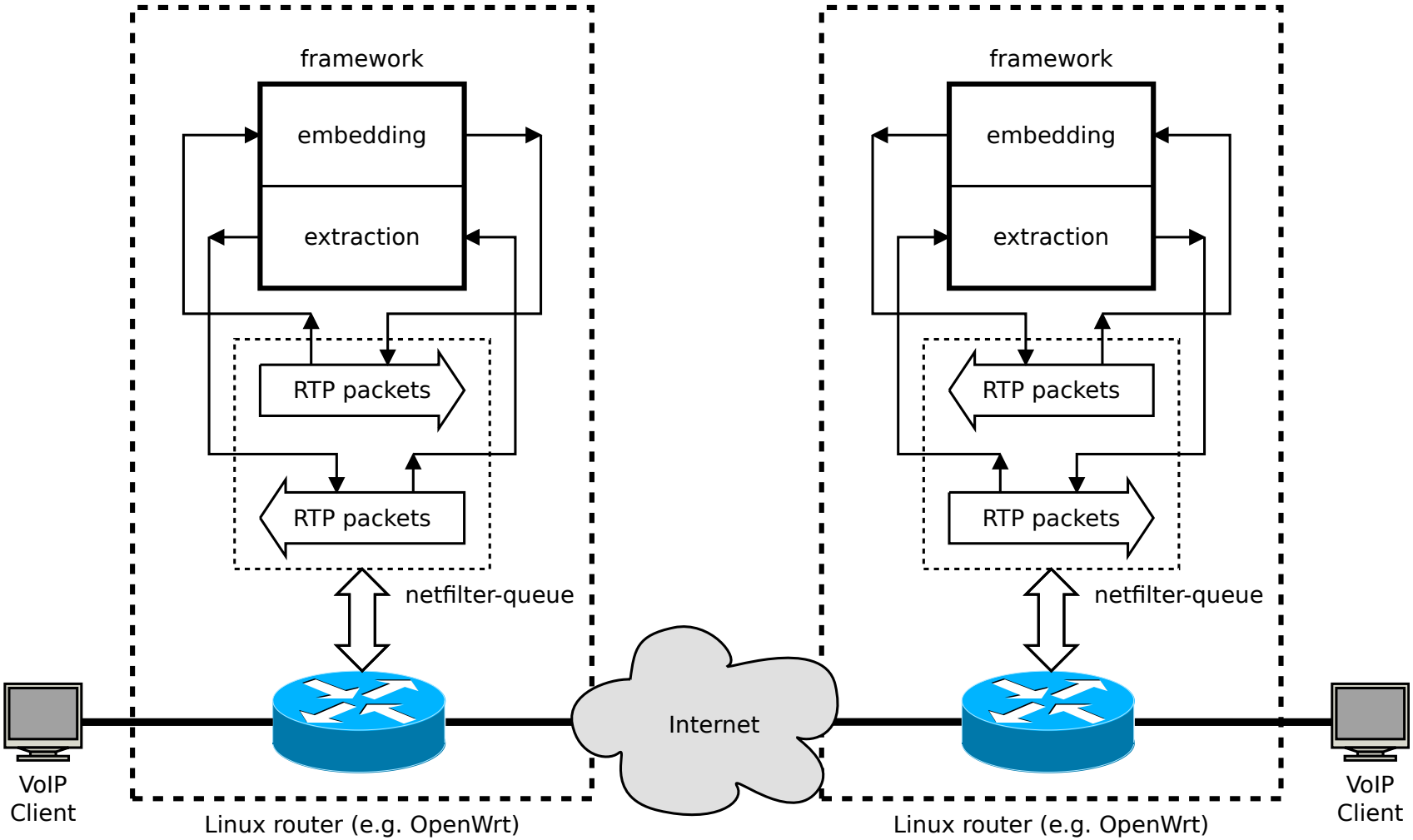
Functional Overview

Cover Media – VoIP



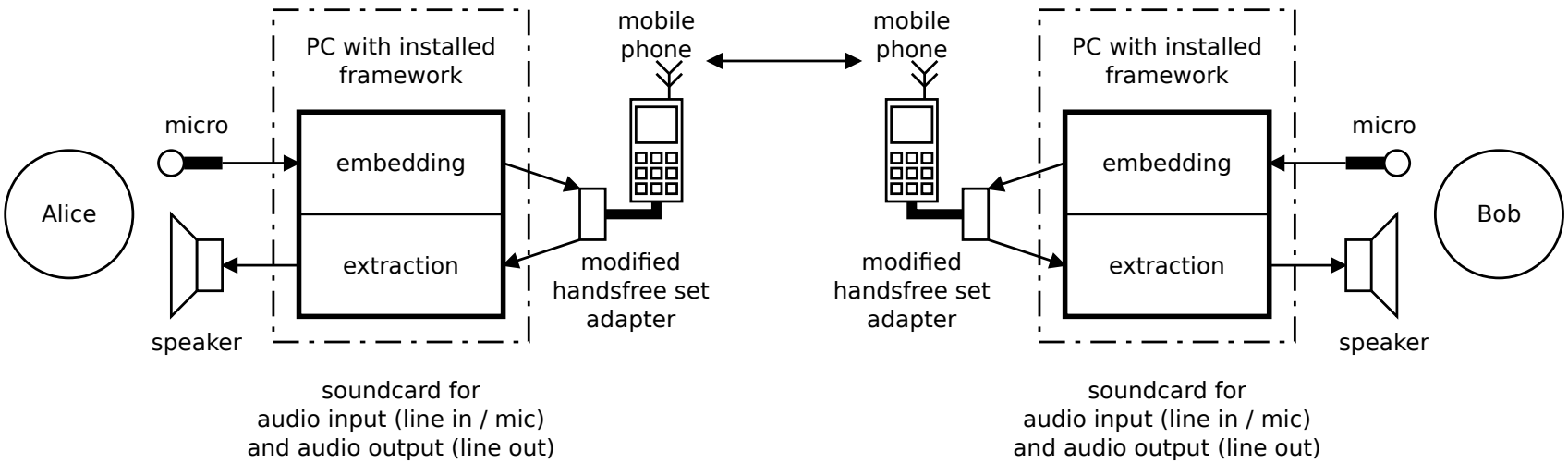
Functional Overview

Cover Media – VoIP



Functional Overview

Cover Media – GSM



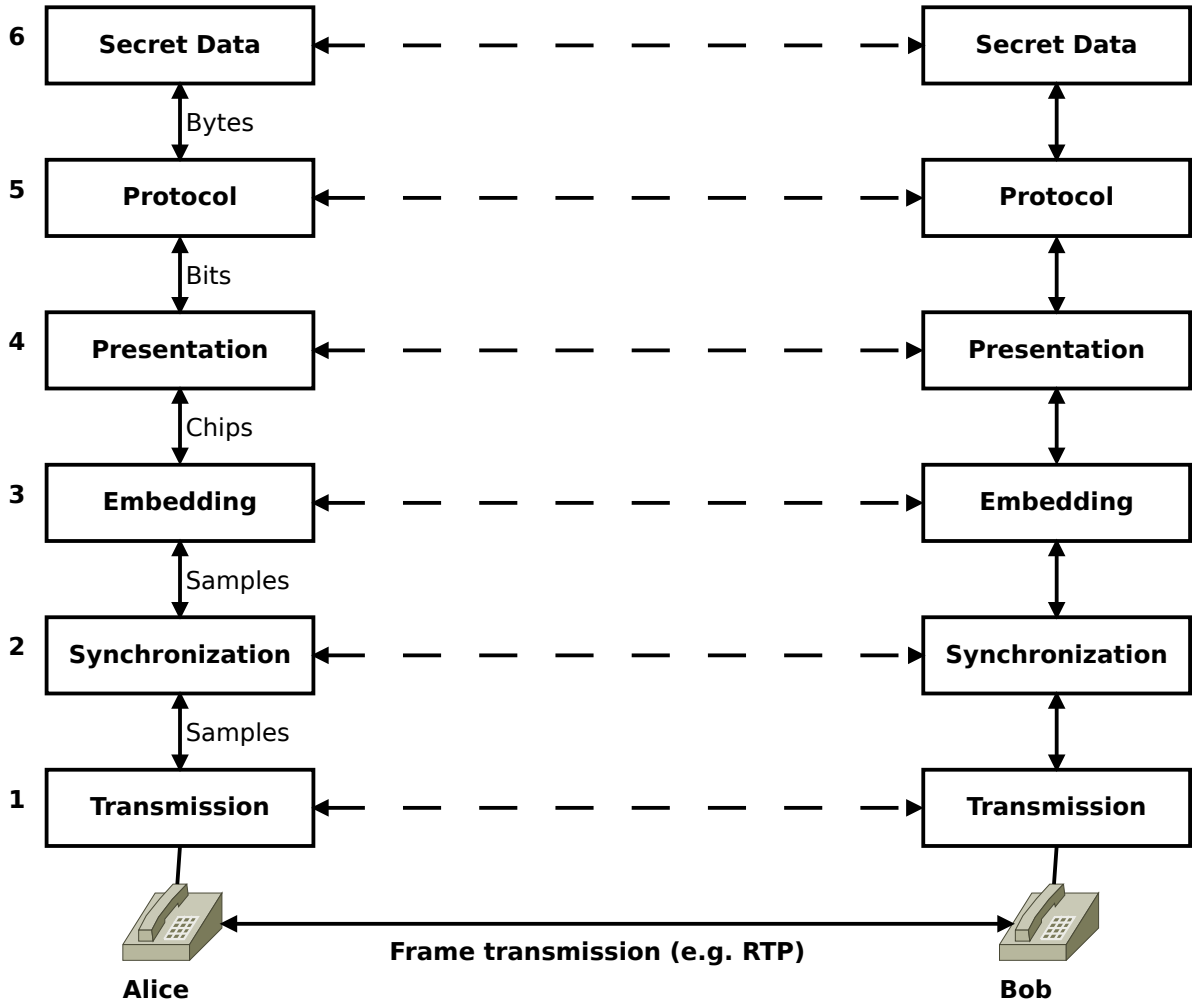
Layered Model

Tasks | Components

- Embedding secret data into the cover medium
- Protocol for data flow handling
  - Integrity checks
  - Segmentation of secret data
  - Acknowledgements
  - Sequence numbers
- Management of component instantiation
- Interfaces for third party software

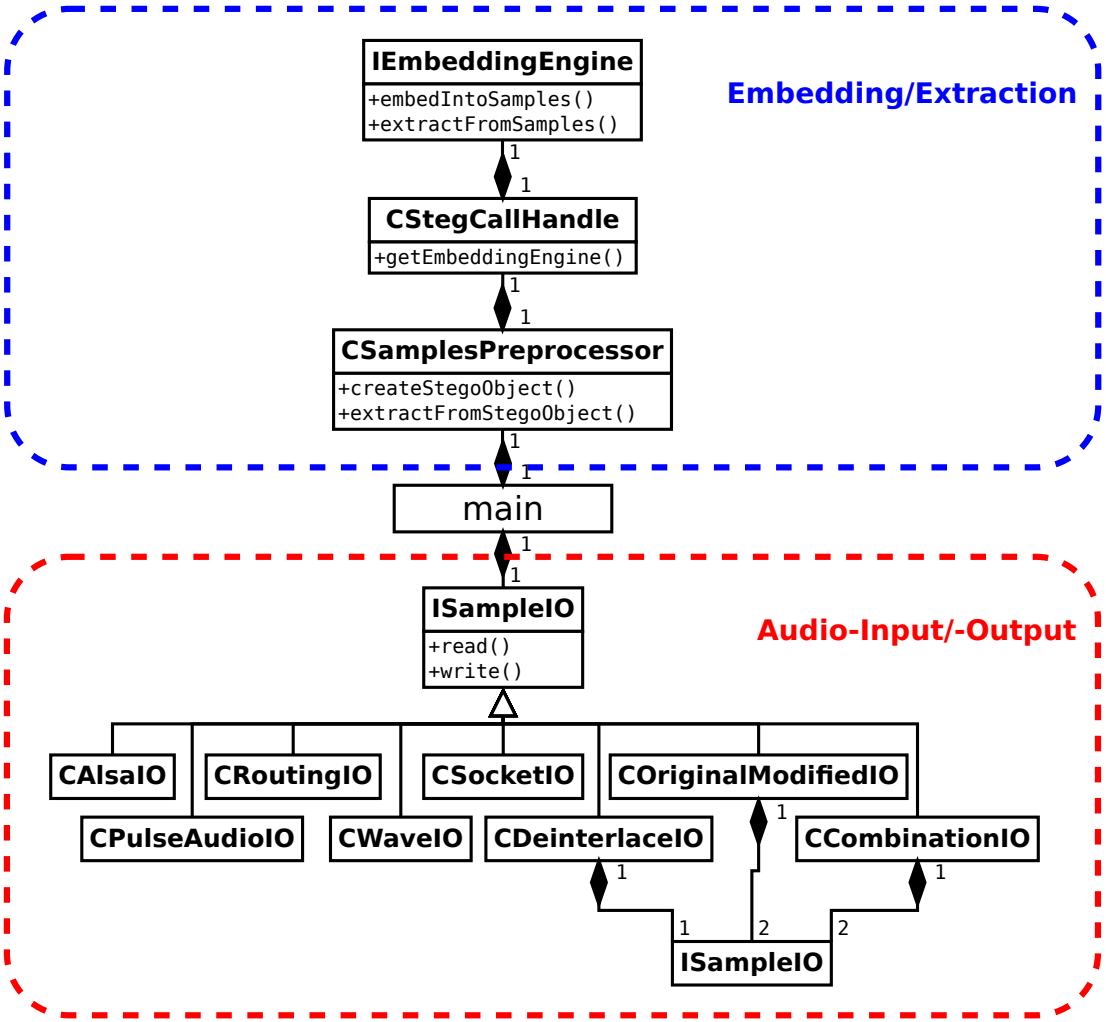
Layered Model

Steganographic Data Exchange



Layered Model

Software Architecture of the IO-components



Steganographic Algorithms

Overview

"Digital" Algorithms

- Samples are processed independently
- Codec-specific
  - LSB Hiding

"Analogue" Algorithms

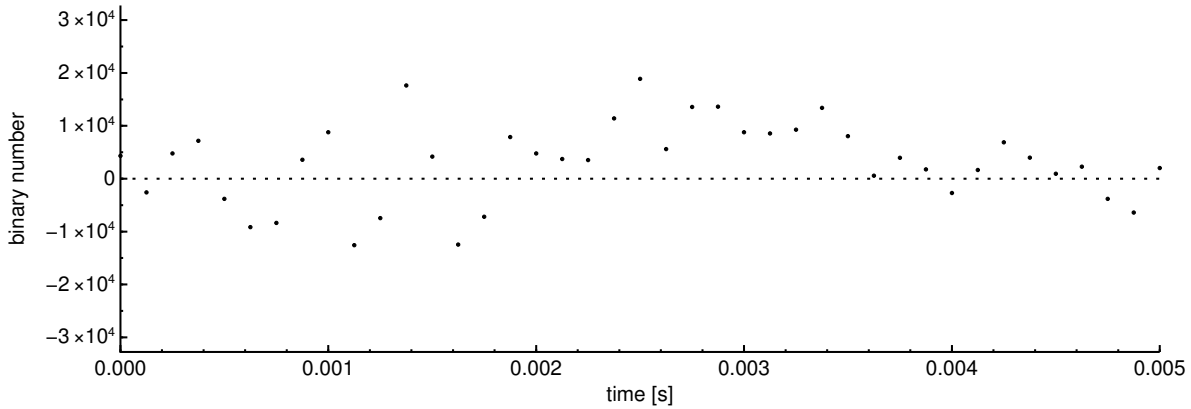
- Number of samples is seen as contiguous signal
  - Echo Hiding – time domain
  - Spread Spectrum – time or frequency domain
  - Phase Coding – frequency domain

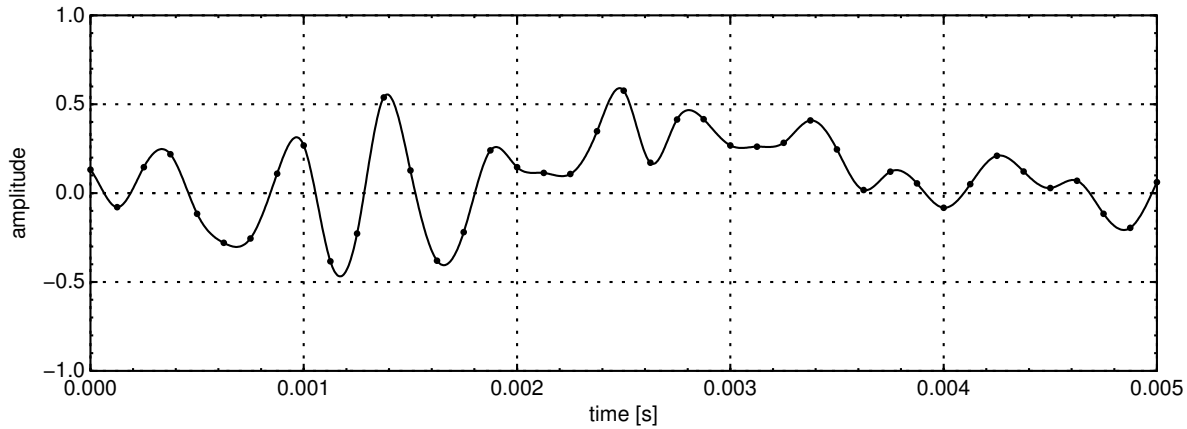
Steganographic Algorithms

Analogue Representation

From audio format (PCM, MP3, . . .) to floating-point

- Decoding of coded sample values before modification



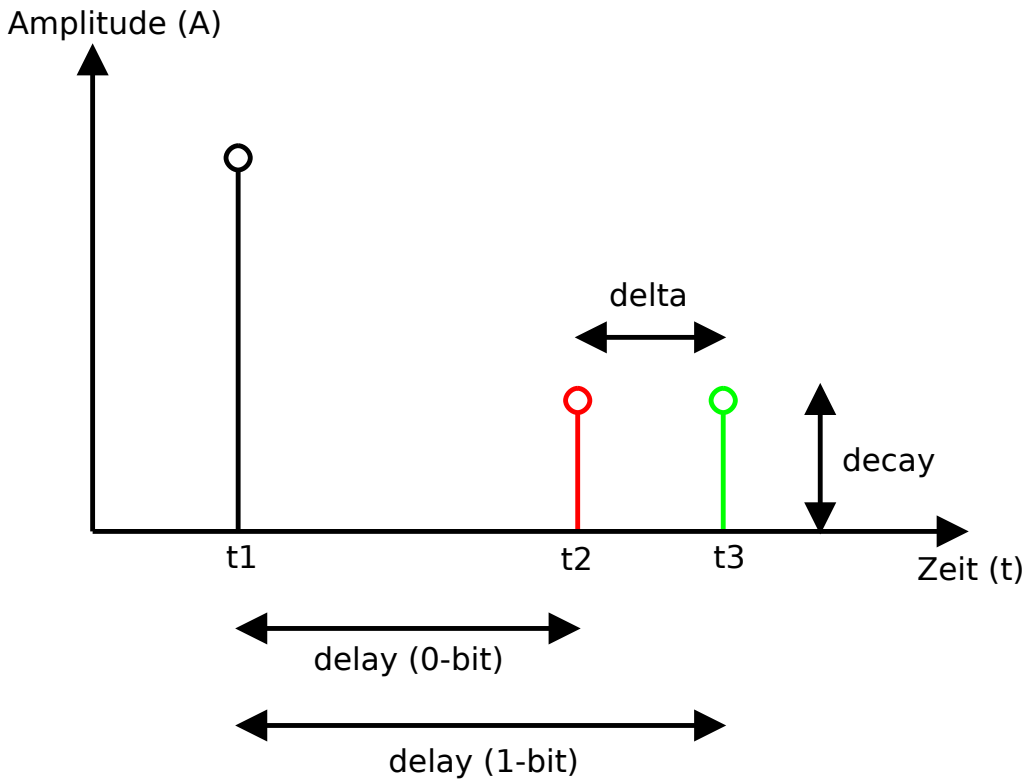


Steganographic Algorithms

Echo Hiding

Three important parameters

- Delay, decay and space between chosen delays
- Further: blocksize



Steganographic Algorithms

Echo Hiding

Echo generation

- Addition of delayed and decayed part of signal

Echo detection

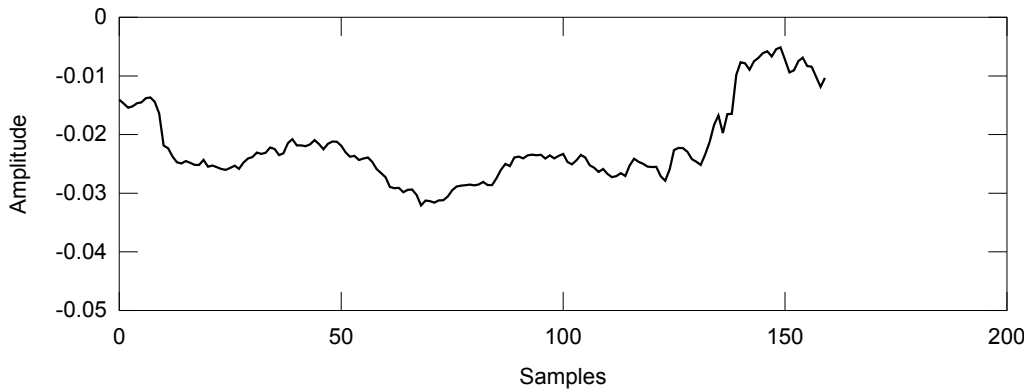
- Calculation of *Cepstrum*
- FFT and log
- Seperates periodic parts (i.e. echos) and signal
- Idea: Periodic parts shown as peaks

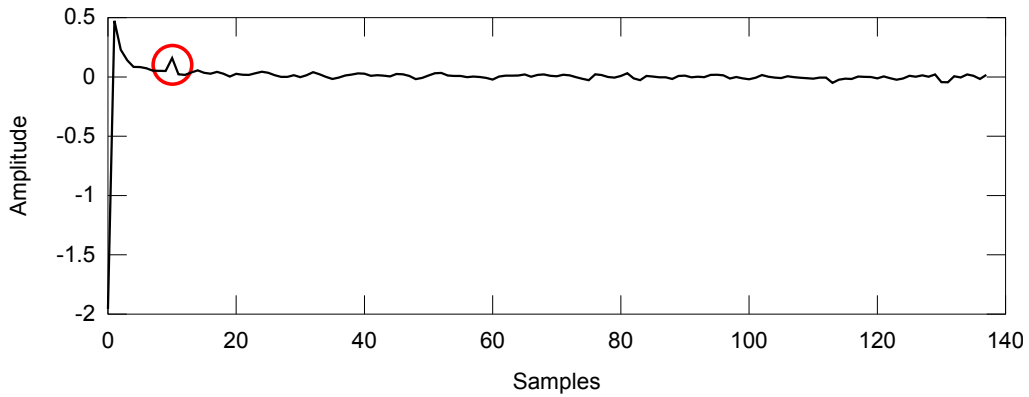
Steganographic Algorithms

Echo Hiding

Example

- $f_s = 8kHz$, $delay = 1.25ms$
- Not always as reliable (silent parts, natural echos, . . .)



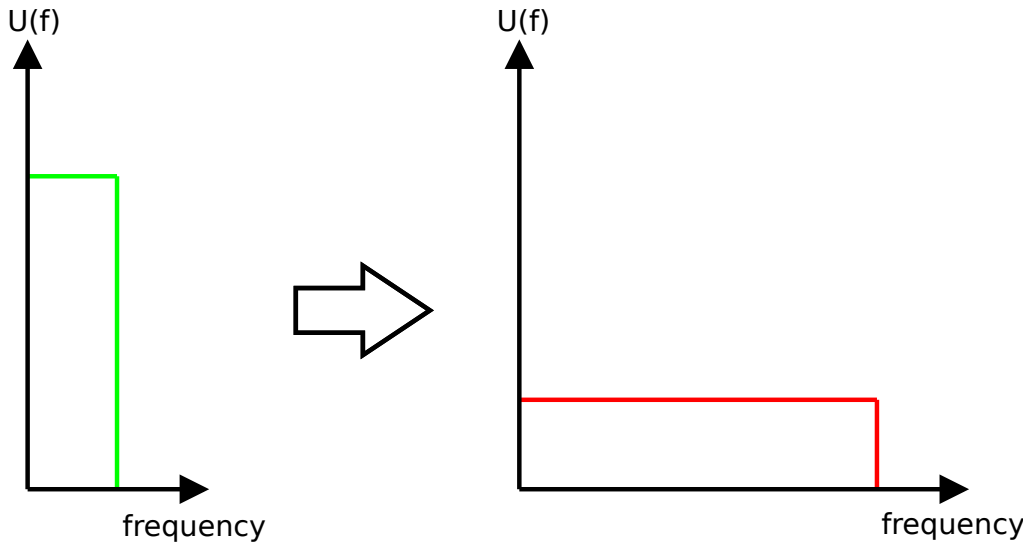


Steganographic Algorithms

Spread Spectrum

Overview

- Narrow band signal is spread over a large bandwidth
- Signal vanishes in noise floor
- Spreading code is needed for retrieval
  - Has to be exchanged before communication
  - Can be generated via PN-sequence generator (LFSR)

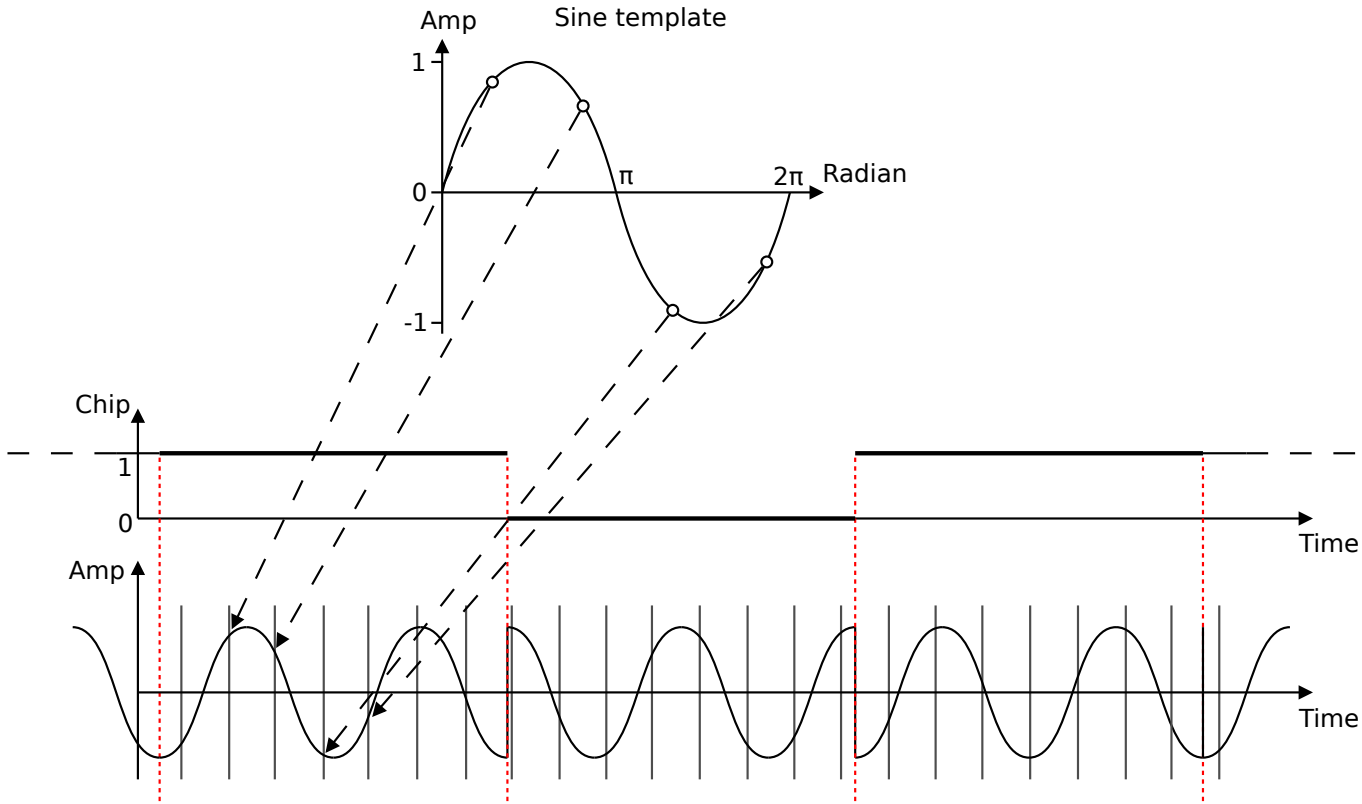


Steganographic Algorithms

Spread Spectrum

Spreading process

- BPSK modulation with sine carrier (variable frequency)
- Specific samples are calculated on-the-fly from template

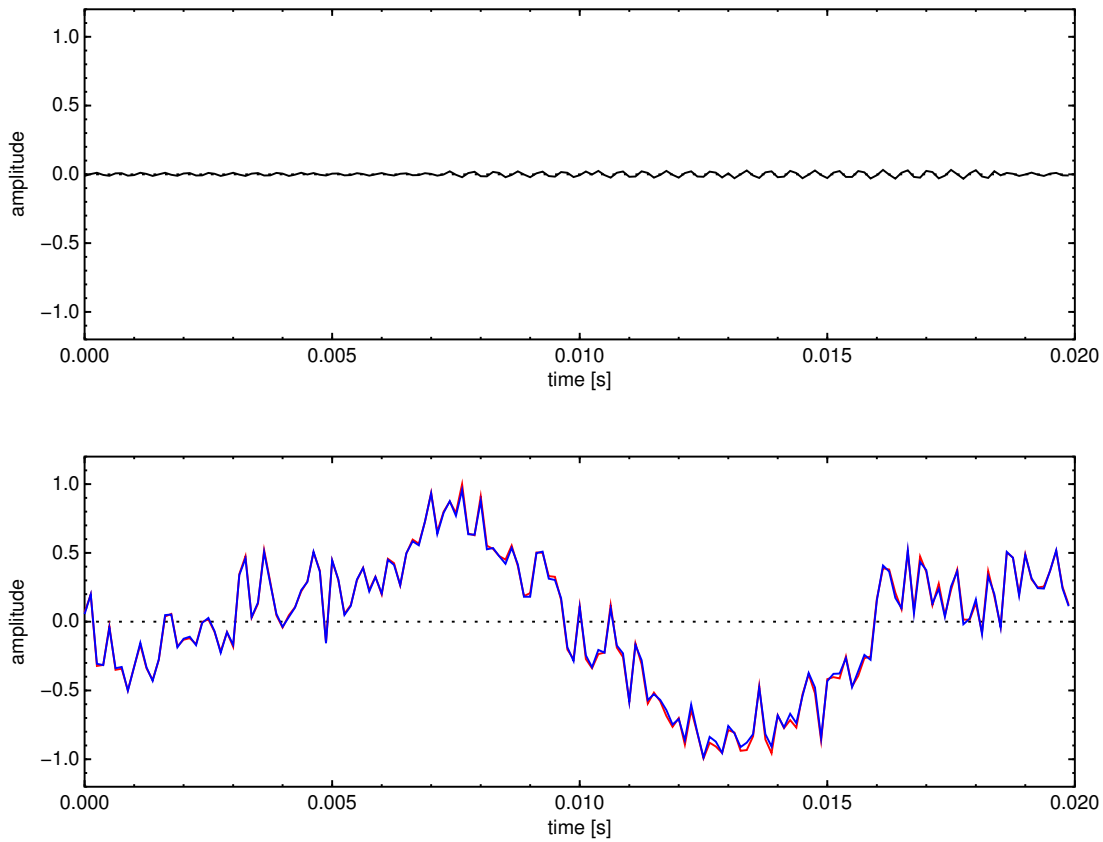


Steganographic Algorithms

Spread Spectrum

Example – Embedding

- Add modulated chip sequence to audio signal
- Red = original signal, blue = modified signal

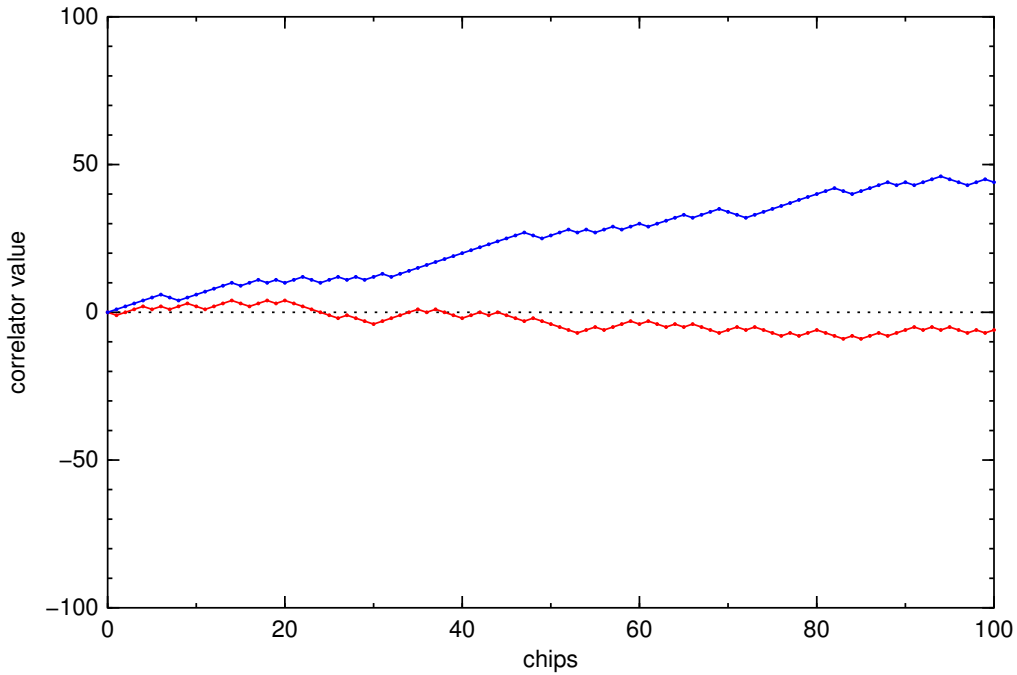


Steganographic Algorithms

Spread Spectrum

Extraction

- Multiplication of received signal with sine carrier
- Mean value over chip time as input for correlator
- Correlation against both chips sequences
- Example: blue = matching sequence

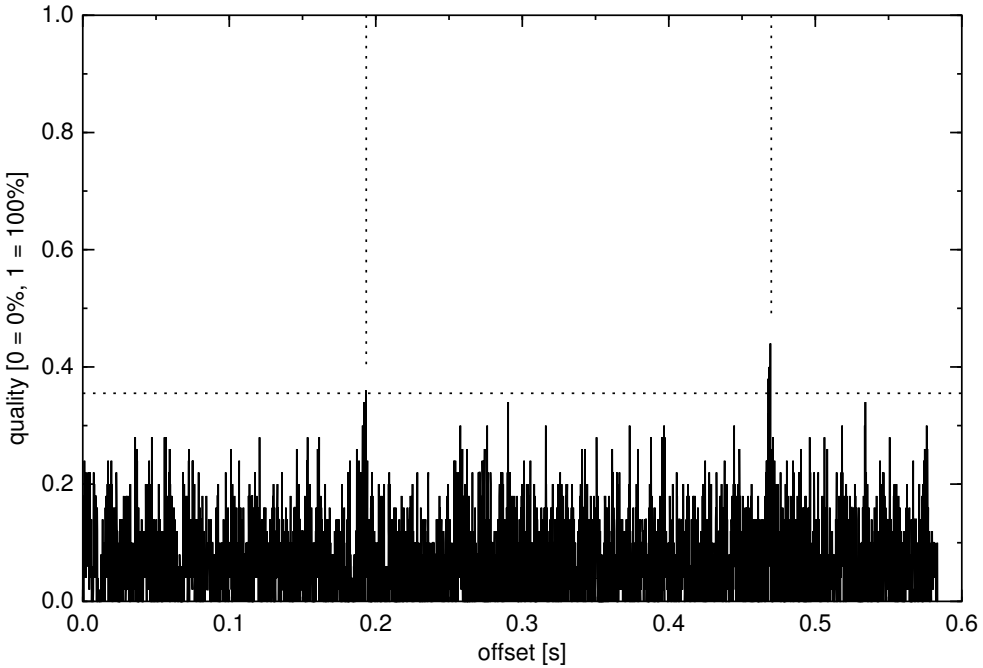


Steganographic Algorithms

Spread Spectrum

Extraction – Synchronization

- Important issue for spread spectrum communication
- Searching the chip sequence start offset
- Compute correlator value for various possible offsets
- Example: histogram of bit quality by time offset

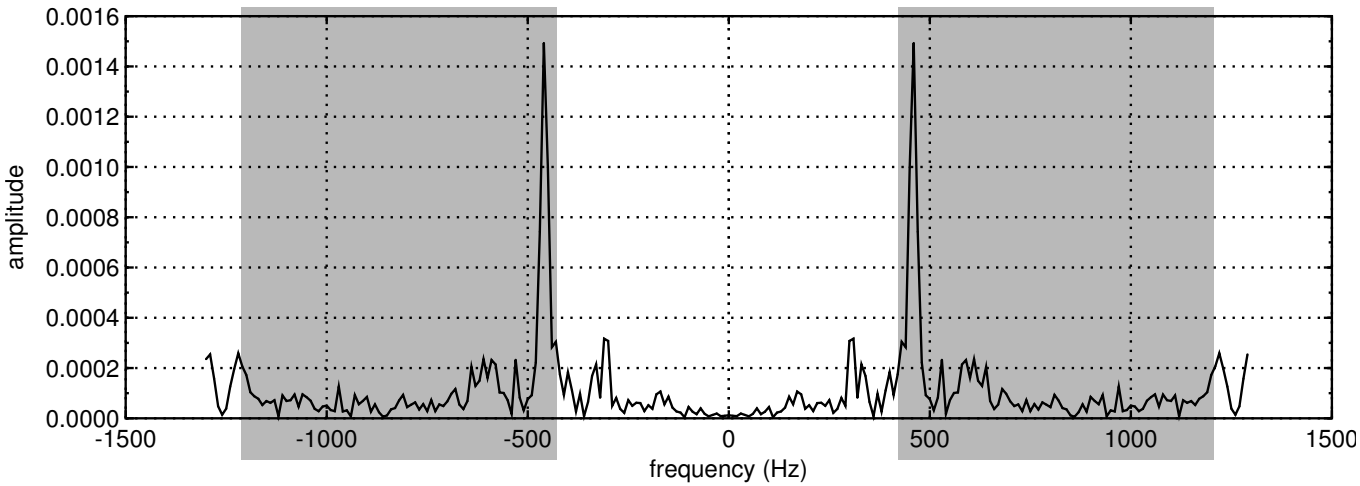


Steganographic Algorithms

Phase Coding

Representation of the audio signal

- FFT of configured time duration ⇒ embedding delay
- For embedding, selection of affected frequency range possible
- Example:
  - chosen frequency interval [400,1200]
  - negative frequencies are modified accordingly

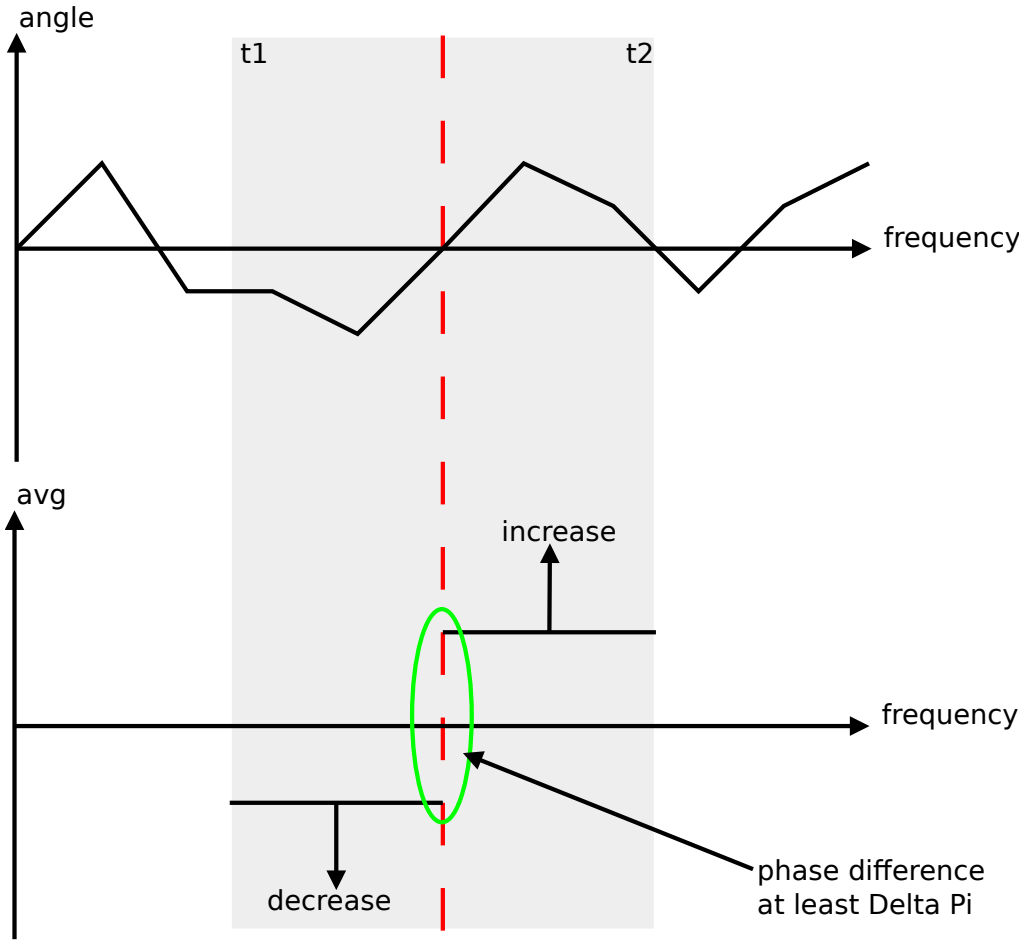


Steganographic Algorithms

Phase Coding

Embedding via phase difference

- Mean value of adjacent parts of signal is adjusted

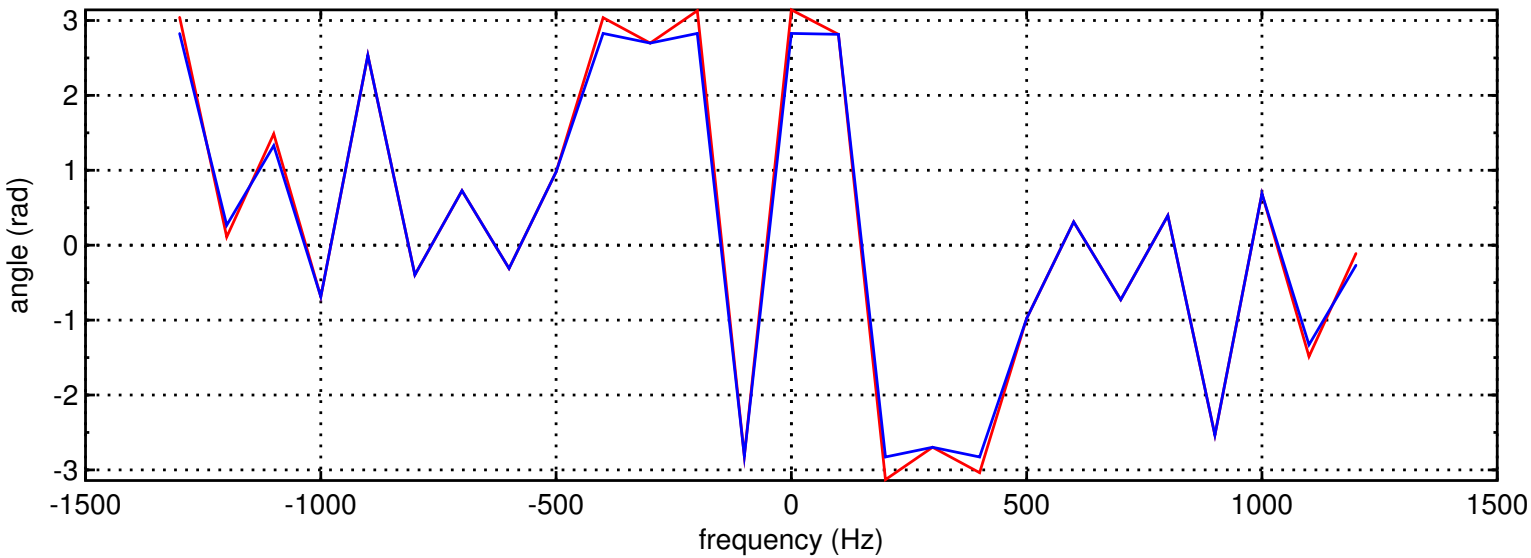


Steganographic Algorithms

Phase Coding

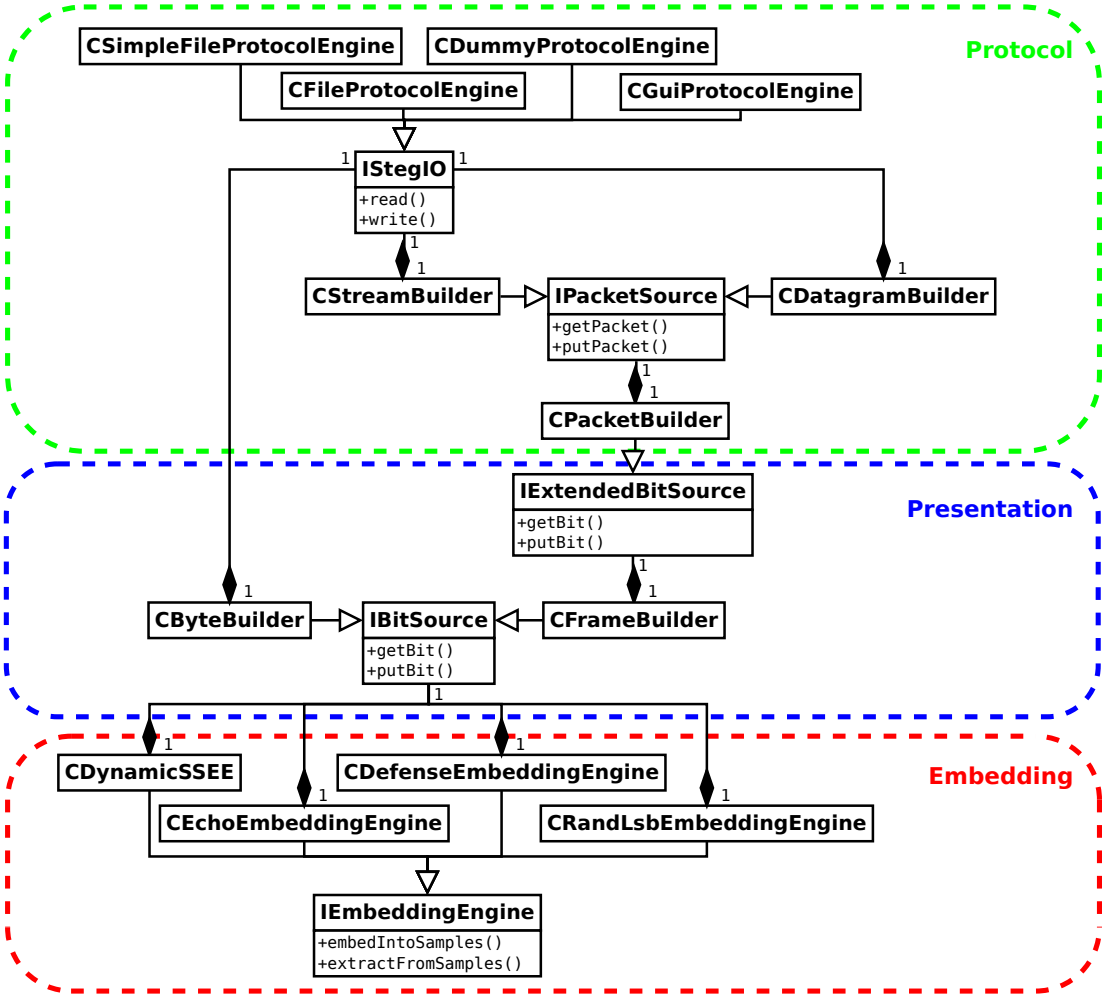
Example

- Minimum phase difference $= \pi/10$
- Red $=$ original phase, blue $=$ modified phase



Steganographic Algorithms

Software Architecture of the Embedding Components



Defense

Analyzing Robustness

- No steganalysis
- Different approaches
  - Noise
  - Jitter
  - Frequency shifting (semitone)
  - Signal cancelling

Implementation

Platforms - Commodity Hardware

Mipsel

- Asus WL-500g Premium v1, based on Broadcom 4704
- (modified) OpenWrt SDK
- Port to other platforms:
  - Routing-Interface
  - Audio-Interface

OpenWrt SDK: Adaptions

- Support for NFQUEUE
- Additional packages added
- Customized firewall settings
- Customized start scripts

Implementation

Platforms - Mobile- and Smartphones

Possibly usage on smartphones

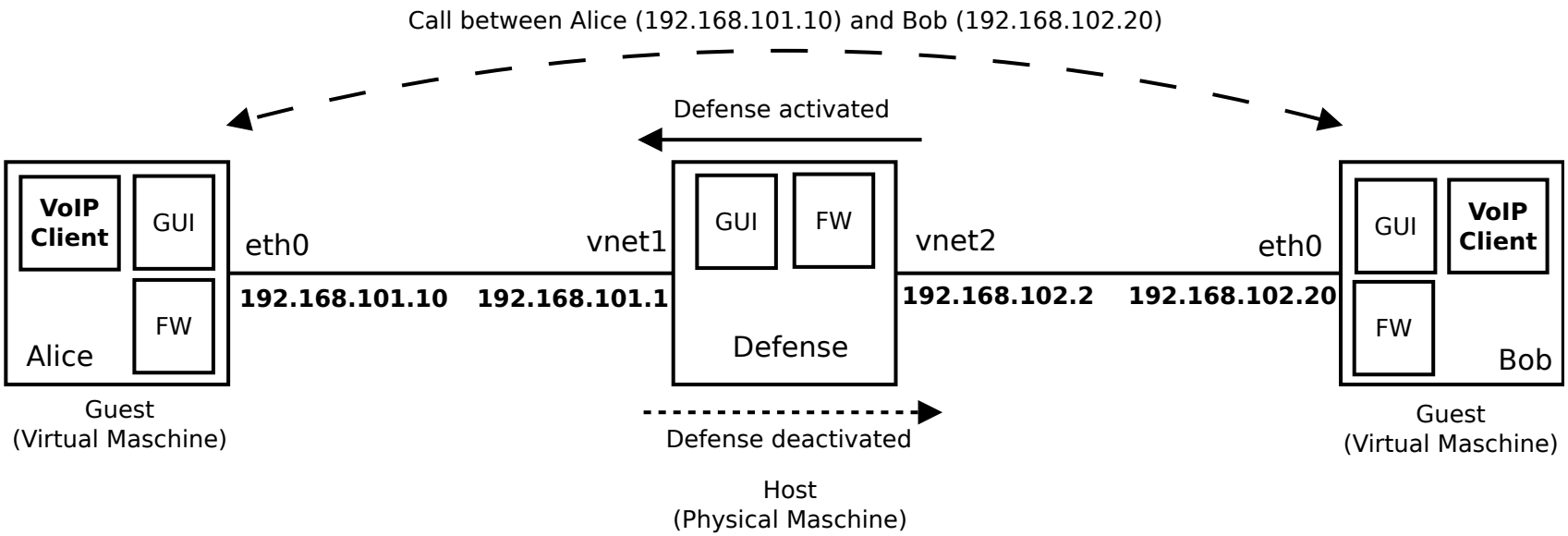
- currently only Linux based phones considered
- e. g. Android powered smartphones
  - NDK allows for reuse of our C++ codebase

Scenarios

- VoIP
- Raw Voice-Data

Demonstration

Setup



Outlook

Future scenarios

- Video streams as cover medium
- Windows-Port
- Better usability
- Improved data throughput
- Smaller, more powerful devices
- Use of steganographic loaders

Thank you for your attention!
Any questions?