



Real-World Steganography and Steganalysis

Martin Beneš

Guest Lecture · University of Twente, The Netherlands, 11 March 2025

Short Bio



Bc., Information Technology
Brno University of Technology



M. Sc., Statistics and Machine Learning
Linköping University



Ph. D. student, Computer Science
University of Innsbruck

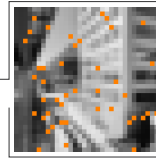


R&D Internship
Blindspot.ai, Prague

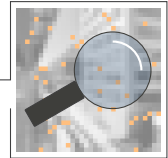
martin.benes@uibk.ac.at

Research and Interests

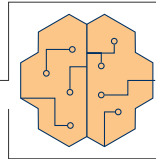
Digital image steganography



Digital image steganalysis



Machine learning

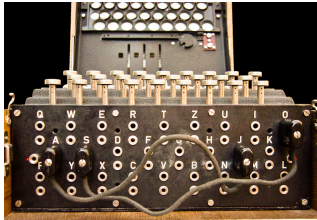


Outline

1. My bio
2. **Establishing steganography**
3. State of “real-world” steganography
4. State of steganography research

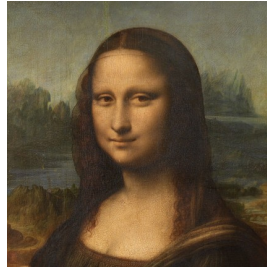
Methods of Secret Communication

Cryptography



conceals the meaning
of the message

Steganography



conceals the existence
of the secret message

Images: Wikimedia.org

Possible Applications for Steganography

Persecutive communication environments

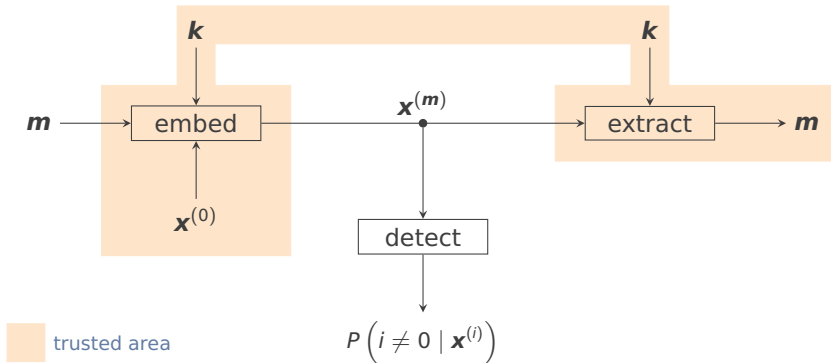
- Students during test
- Prisoners
- Organized crime
- Malware
- Intelligence agencies
- Corporate espionage
- Dissidents in repressive regimes
- Journalists under censorship
- Communication in conflict zones



Image: valorreplicas.com (with friendly permission)

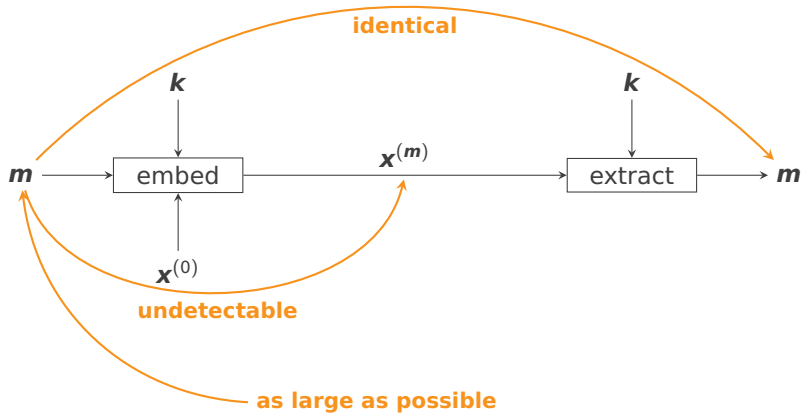
Steganographic System

m secret message $\mathbf{x}^{(0)}$ cover
 k key $\mathbf{x}^{(m)}$ stego object



Steganographic System

m secret message $x^{(0)}$ cover
 k key $x^{(m)}$ stego object



How to Embed a Message into a Cover?

File bytes

Directly modify the bytes.

Append after EOF.

File structure

Exploit the format (e.g., “hidden” marker)

Embed into the metadata.

File data

Modify the data itself.

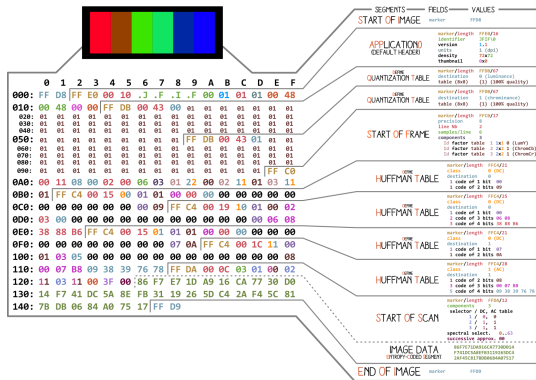
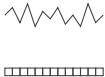


Image: Ange Albertini

Types of Covers

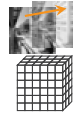
Audio



Image



Video



Network traffic

247	1.041600	172.25.95.226	142.250.185.238	QUIC
248	1.043130	138.232.17.233	172.25.95.226	TCP
249	1.043130	138.232.17.233	172.25.95.226	TLSv1...

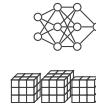


Text

The person eats a broccoli.

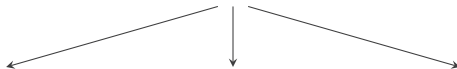


Neural network weights



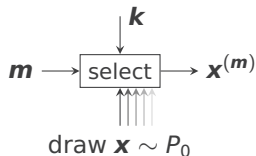
Steganographic channels are typically in the compressed domain.

Approaches to Steganography



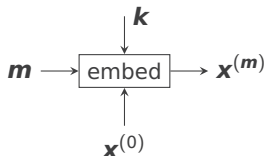
Cover selection

Select the cover that carries the desired message.



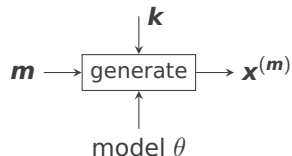
Cover modification

Modify a cover to embed the desired message.



Cover synthesis

Generate a cover that carries the desired message.



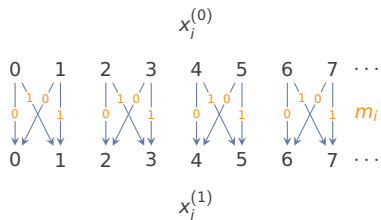
Outline

1. My bio
2. Establishing steganography
3. **State of “real-world” steganography**
4. State of steganography research

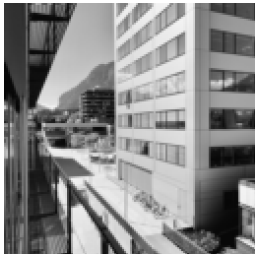
LSB Steganography

Hide the message ***m*** in the least significant bits (LSBs) of the media data.

LSB replacement (LSBR)



Selection of Embedding Positions



cover



sequential ($p = 0.25$)

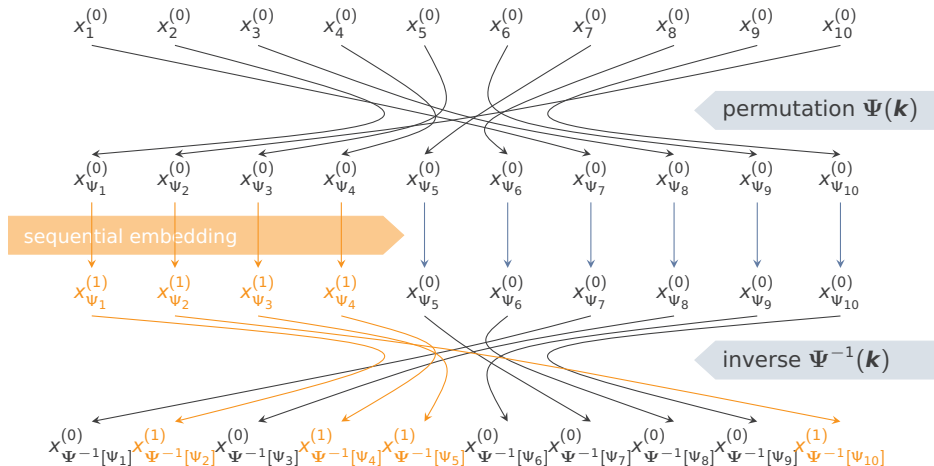


spread ($p = 0.25$)



adaptive ($p = 0.25$)

Key-Dependent Permutation



Available Steganographic Tools

Image pixels	steghide	jphide	openstego	...
JPEG coefficients	jsteg	F5	outguess	...
Audio	mp3stego	spectrology	...	

Open-source stego tools

Most public stego tools are decades **old** and use **insecure** steganography.

Publishing stego tools is ethically questionable, because it may help criminals.

Andreas Westfeld, Andreas Pfitzmann. Attacks on steganographic systems. Springer IH, 1999.

Graeme Bell, Yeuan-Kuen Lee. A Method for Automatic Identification of Signatures of Steganography Software. IEEE TIFS, 2010.

Daniel Lerch. Stego-collection. <https://github.com/daniellerch/stego-collection>

Daniel Lerch. Steganography Tools. <https://daniellerch.me/stego/intro/tools-en/>

Forensic Tools against Naïve Steganography

```
Comment      : escape at midnight
Image Width  : 1857
Image Height : 1032
```

File bytes	hexdump	binwalk	...
File structure	exiftool	strings	...
Data	visual analysis	Aperi'Solve	...

```
(venv) martin@UBK-MBP Desktop % exiftool image.jpeg
ExifTool Version Number      : 13.10
File Name                    : image.jpeg
Directory                    : .
File Size                     : 111 KB
File Modification Date/Time  : 2024:11:15 09:31:21
File Access Date/Time       : 2024:11:15 09:31:21
File Creation Date/Time     : 2024:11:15 09:31:21
Exif Version                 : 2.2
Exif Date/Time               : 2024:11:15 09:31:21
Exif Image Width             : 1857
Exif Image Height            : 1032
Profile CM Type              : Apple Computer Inc.
Profile Version              : 4.0.0
Profile Class                : Display Device Profile
Color Space Data             : RGB
Profile Connection Space     : XYZ
Profile Date Time            : 2024:11:15 09:31:21
Profile File Signature       : acsp
Primary Platform             : Apple Computer Inc.
CM Flags                     : Not Embedded, Independent
Device Manufacturer         : Apple Computer Inc.
Device Model                 : 
Device Attributes            : Reflective, Glossy, Positive, Color
Rendering Intent             : Perceptual
Connection Space Illuminant  : D50 Illuminant
Profile Creator              : Apple Computer Inc.
Profile ID                   : b7efc468c6867656f337e678c2da5df
Profile Description          : DELL U5429NE
Profile Copyright            : Copyright Apple Inc., 2024
Media White Point            : 0.9642 1 0.82491
Red Matrix Column            : 0.52797 0.23764 0.00093
Green Matrix Column          : 0.29321 0.69942 0.06573
Blue Matrix Column           : 0.14302 0.05294 0.75522
Red Tone Reproduction Curve  : (Binary data 16 bytes, use -b option to extract)
Green Tone Reproduction Curve : (Binary data 16 bytes, use -b option to extract)
Blue Tone Reproduction Curve : (Binary data 16 bytes, use -b option to extract)
Chromatic Adaptation        : 1.04576 0.02182 -0.04945 0.02797 0.99156 -0.01674 -0.0
XMP Toolkit                  : XMP Core 6.0.0
Comments                     : escape at midnight
Image Width                  : 1857
Image Height                  : 1032
Encoding Process             : Baseline DCT, Huffman coding
Bits Per Sample              : 8
Color Components              : 3
Y Cb Cr Sub Sampling         : YCbCr4:4:4 (1:1)
Image Size                   : 1857x1032
Megapixels                   : 1.9
```

<https://www.aperisolve.com/>

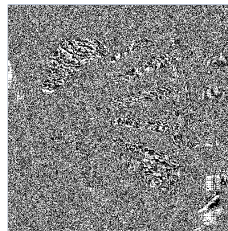
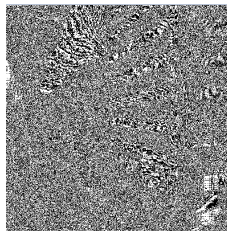
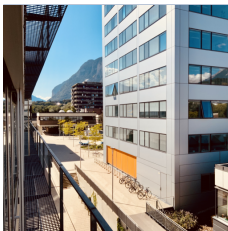
Visual Analysis

Cover

Cover LSB

Stego LSB

Natural images

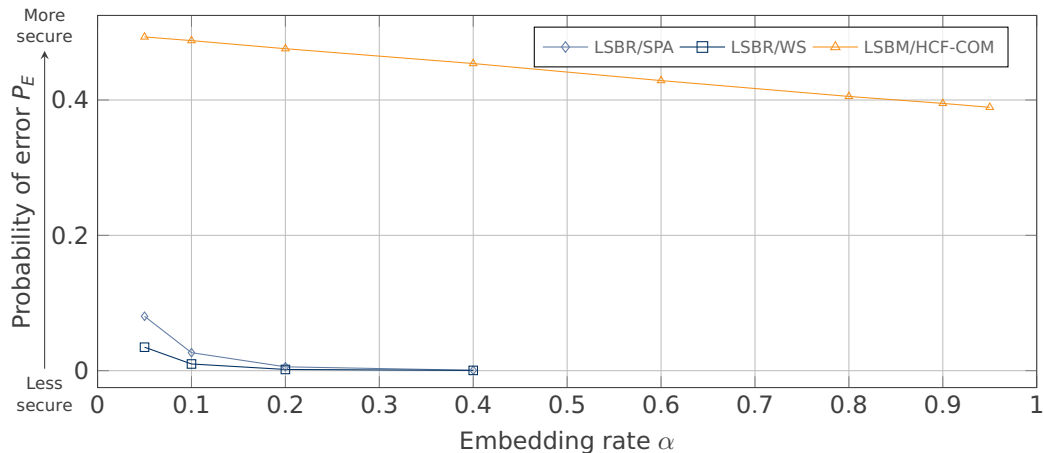


Computer graphic



Sequential embedding may be visible in the image LSB plane.

Security of LSB Steganography

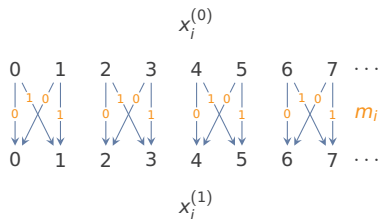


10 000 grayscale images from the BOSSBase dataset, size 512^2 , LSBR/M without coding ($e = 2$). Training set 50%.

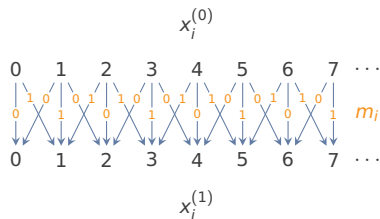
LSB Steganography

Hide the message ***m*** in the least significant bits (LSBs) of the media data.

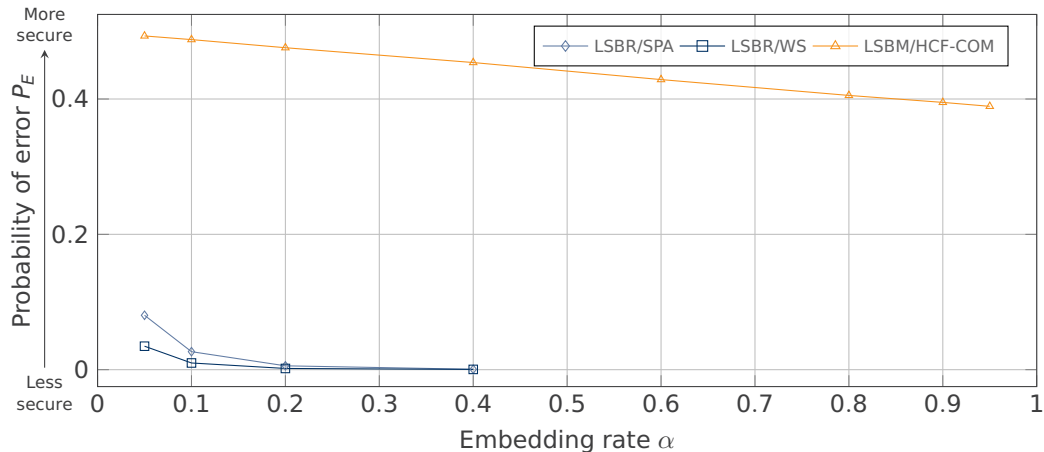
LSB replacement (LSBR)



LSB matching (LSBM)

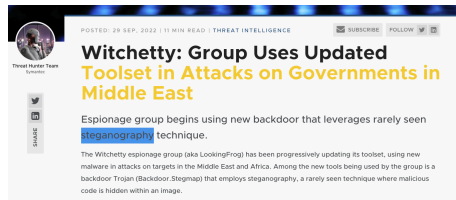


Security of LSB Steganography



10 000 grayscale images from the BOSSBase dataset, size 512^2 , LSBR/M without coding ($e = 2$). Training set 50%.

Real World Cases of Steganography



There is little evidence of steganography being used in the digital communication.
(perhaps due to its nature to remain hidden)

<https://www.welivesecurity.com/2022/09/06/worok-big-picture/>

<https://www.security.com/threat-intelligence/witchetty-steganography-espionage>

Mukesh Dalal, Mamta Juneja: A survey on information hiding using video steganography. Springer AIR, 2021.

Real World Cases of Steganalysis

Common task during challenges, competitions, and CTFs.

A relevant competence for

- intelligence services and
- law-enforcement agencies.

UNCOVER project (EU Horizon 2020)



V. Leask, R. Coganne, D. Borghys, H. Bruyninckx. UNCOVER: Development of an efficient steganalysis framework for uncovering hidden data in digital media. ACM ARES, 2022.

<https://alaska.utt.fr/>

<https://www.uncoverproject.eu/>

Outline

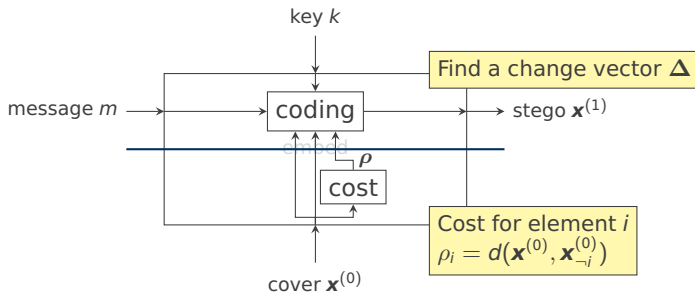
1. My bio
2. Establishing steganography
3. State of “real-world” steganography
4. **State of steganography research**

Paradigms of Modern Steganography

Separation principle

Two step embedding

- cost calculation
- message coding

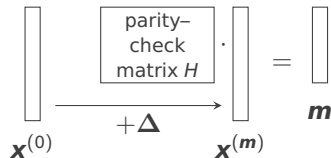


$$\mathbf{x}_{\neg i}^{(0)} = \{x_0^{(0)}, \dots, x_{i-1}^{(0)}, x_i^{(1)}, x_{i+1}^{(0)}, \dots, x_N^{(0)}\}$$

Steganographic Syndrome Coding

Posed as a constrained optimization problem.

$$\underbrace{\min_{\Delta} \sum_{i=1}^N \Delta_i \cdot \rho_i}_{\text{optimization}}, \text{ s. t. }, \underbrace{H \cdot x^{(m)} = m}_{\text{constraint}}$$



Objectives of coding methods

- Reduce # of changes (Hamming codes)
- Avoid saturation (wet paper codes)
- Minimize cost (trellis codes)

Example:

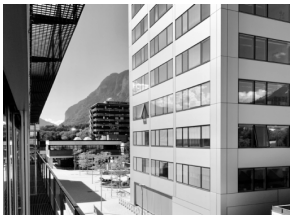
i	1	2	3	4	...
$x_i^{(0)}$	42	26	14	98	...
ρ_i	0.05	0.14	0.02	1.10	...
Δ_i	1	0	1	0	...
$x_i^{(m)}$	43	26	13	98	...

$m = 01 \dots$

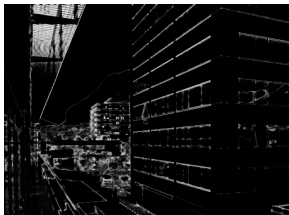
$m_2 = 1$

Comparison of Cost Functions

Cover



HUGO 2010



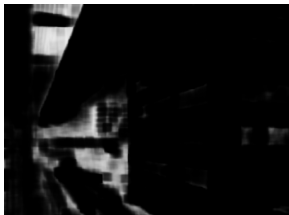
WOW 2012



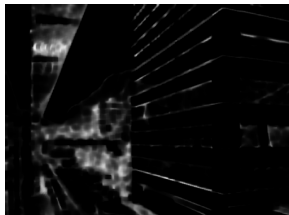
1
embedding probability β_i at $\alpha = 0.2$ bpp
0



S-UNIWARD 2014



HILL 2014



MiPOD 2015

Modern Steganalysis

Targeted attacks

= against a stego method

Compatibility attacks

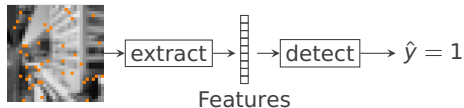
= against a cover

Handcrafted features

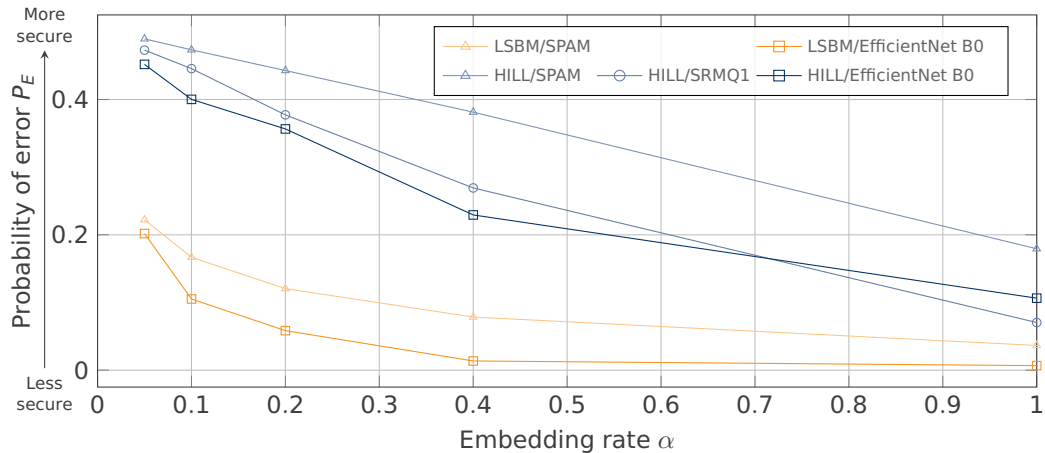
Project the image to a feature space and apply a learning-based detector.

Neural networks

Architectures: *EfficientNet B0*, *SRNet*



Security of Modern Steganography



10 000 grayscale images from the BOSSBase dataset, size 512^2 , LSBM without coding ($e = 2$), HILL with optimal coding. Training set 50%.

The Source of Covers

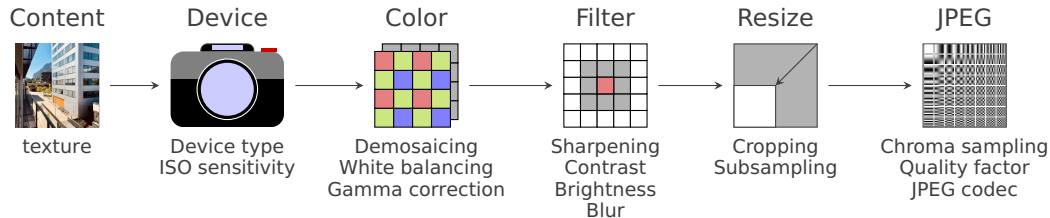
Cryptography needs a good random number generator.



Steganography needs a good source of covers.

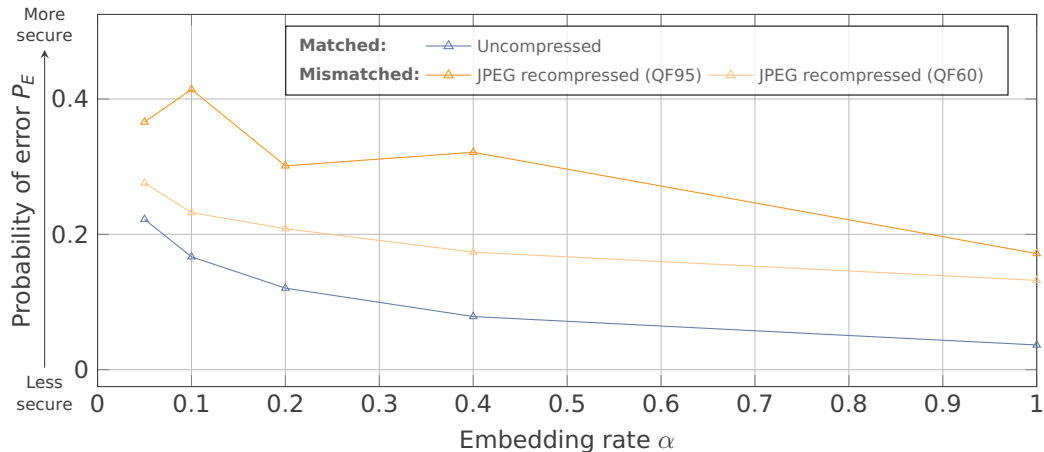


Cover source



A. Mallet, M. Beneš, R. Cogranne. Cover-source mismatch in steganalysis: systematic review. EURASIP JIS, 26, 2024.

Cover-Source Mismatch



10 000 grayscale images from the BOSSBase dataset, size 512^2 , libjpeg-turbo, LSBM. Training set 50%. SPAM+FLD.



Bedankt !

Real-World Steganography and Steganalysis

martin.benes@uibk.ac.at



Guest Lecture · University of Twente, The Netherlands, 11 March 2025