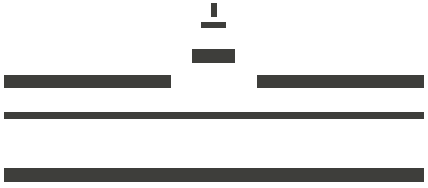


Westfälische
Wilhelms-Universität
Münster

Predictable Rain ?

Steganalysis of Public-Key Steganography using Wet Paper Codes

Matthias Carnein, Pascal Schöttle, Rainer Böhme
University of Münster

IH & MMSEC | Salzburg, Austria | June 12, 2014

Outline

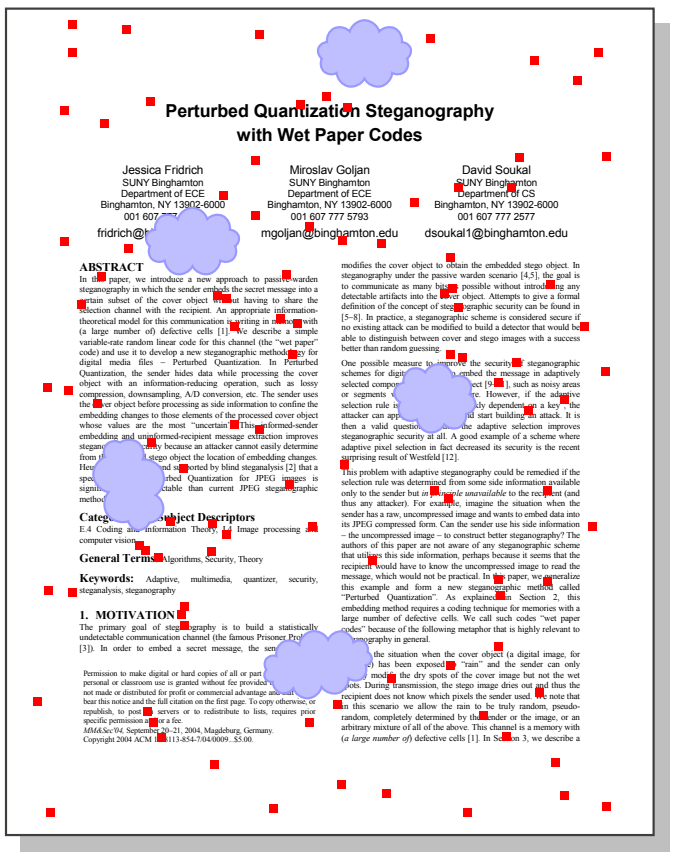
1. Related Work
2. Estimating Embedding Positions
3. Weighted Stego-Image Steganalysis
4. Empirical Results
5. Discussion & Conclusion

living.knowledge WWU Münster

Outline

living.knowledge WWU Münster

Wet Paper Codes (WPC)



Perturbed Quantization Steganography with Wet Paper Codes

Jessica Fridrich
SUNY Binghamton
Department of ECE
Binghamton, NY 13902-6000
001 607 777 ...
fridrich@...

Miroslav Goljan
SUNY Binghamton
Department of ECE
Binghamton, NY 13902-6000
001 607 777 5793
mgoljan@binghamton.edu

David Soukal
SUNY Binghamton
Department of CS
Binghamton, NY 13902-6000
001 607 777 2577
dsoukal1@binghamton.edu

ABSTRACT

In this paper, we introduce a new approach to passive-warden steganography in which the sender embeds the secret message into a certain subset of the cover object without having to share the selection channel with the recipient. An appropriate information-theoretical model for this communication is writing in memory with (a large number of) defective cells [1]. We describe a simple variable-rate random linear code for this channel (the "wet paper" code) and use it to develop a new steganographic methodology for digital media files – Perturbed Quantization. In Perturbed Quantization, the sender hides data while processing the cover object with an information-reducing operation, such as lossy compression, downsampling, A/D conversion, etc. The sender uses the cover object before processing as side information to confine the embedding changes to those elements of the processed cover object whose values are the most "uncertain". This informed-sender embedding and uninformed-recipient message extraction improves steganographic security because an attacker cannot easily determine from the stego object the location of embedding changes. Heuristic arguments supported by blind steganalysis [2] that a specific implementation of Perturbed Quantization for JPEG images is significantly more undetectable than current JPEG steganographic methods.

Categories and Subject Descriptors

E.4 Coding and Information Theory, I.4 Image processing and computer vision

General Terms: Algorithms, Security, Theory

Keywords: Adaptive, multimedia, quantizer, security, steganalysis, steganography

1. MOTIVATION

The primary goal of steganography is to build a statistically undetectable communication channel (the famous Prisoner Problem [3]). In order to embed a secret message, the sender modifies the cover object to obtain the embedded stego object. In steganography under the passive warden scenario [4,5], the goal is to communicate as many bits as possible without introducing any detectable artifacts into the cover object. Attempts to give a formal definition of the concept of steganographic security can be found in [5–8]. In practice, a steganographic scheme is considered secure if no existing attack can be modified to build a detector that would be able to distinguish between cover and stego images with a success better than random guessing.

One possible measure to improve the security of steganographic schemes for digital media is to embed the message in adaptively selected components of the cover object [9–11], such as noisy areas or segments with complex texture. However, if the adaptive selection rule is publicly known or only dependent on a key, the attacker can apply the same selection rule and start building an attack. It is then a valid question to ask whether the adaptive selection improves steganographic security at all. A good example of a scheme where adaptive pixel selection in fact decreased its security is the recent surprising result of Westfeld [12].

This problem with adaptive steganography could be remedied if the selection rule was determined from some side information available only to the sender but *in principle unavailable* to the recipient (and thus any attacker). For example, imagine the situation when the sender has a raw, uncompressed image and wants to embed data into its JPEG compressed form. Can the sender use his side information – the uncompressed image – to construct better steganography? The authors of this paper are not aware of any steganographic scheme that utilizes this side information, perhaps because it seems that the recipient would have to know the uncompressed image to read the message, which would not be practical. In this paper, we generalize this example and form a new steganographic method called "Perturbed Quantization". As explained in Section 2, this embedding method requires a coding technique for memories with a large number of defective cells. We call such codes "wet paper codes" because of the following metaphor that is highly relevant to steganography in general.

Consider the situation when the cover object (a digital image, for example) has been exposed to "rain" and the sender can only modify the dry spots of the cover image but not the wet spots. During transmission, the stego image dries out and thus the recipient does not know which pixels the sender used. We note that in this scenario we allow the rain to be truly random, pseudo-random, completely determined by the sender or the image, or an arbitrary mixture of all of the above. This channel is a memory with (*a large number of*) defective cells [1]. In Section 3, we describe a

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
MM&Sec'04, September 20–21, 2004, Magdeburg, Germany.
Copyright 2004 ACM 1-58113-854-7/04/0009...$5.00.

- paper exposed to rain
  - dry: possible embedding position
  - wet: excluded
- dries out during transmission
- embedding positions unknown to recipient
- WPC enable extraction of message

Fridrich, Goljan and Soukal (2004)

living.knowledge WWU Münster

Wet Paper Codes (WPC)

- recipient

$$\underbrace{\mathbf{D}}_{\text{shared matrix}} \times \mathbf{b} = \mathbf{m}$$

- sender solves for change vector **v**

$$|\mathbf{m}| \overbrace{\left\{ \begin{array}{|c|} \hline \\ \mathbf{D} \\ \\ \hline \end{array} \right.}^{n} \times \begin{bmatrix} 0 \\ v_2 \\ \vdots \\ 0 \\ v_n \end{bmatrix} = \mathbf{m} - \mathbf{D} \times \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix}$$

Fridrich et al. (2005a)

living.knowledge WWU Münster

Wet Paper Codes (WPC)

- recipient

$$\underbrace{\mathbf{D}}_{\text{shared matrix}} \times \mathbf{b} = \mathbf{m}$$

- sender solves for change vector $\mathbf{v}$

$$|\mathbf{m}| \overbrace{\left\{ \mathbf{H} \right.}^{k} \times \mathbf{v} = \mathbf{m} - \mathbf{D} \times \mathbf{b}$$

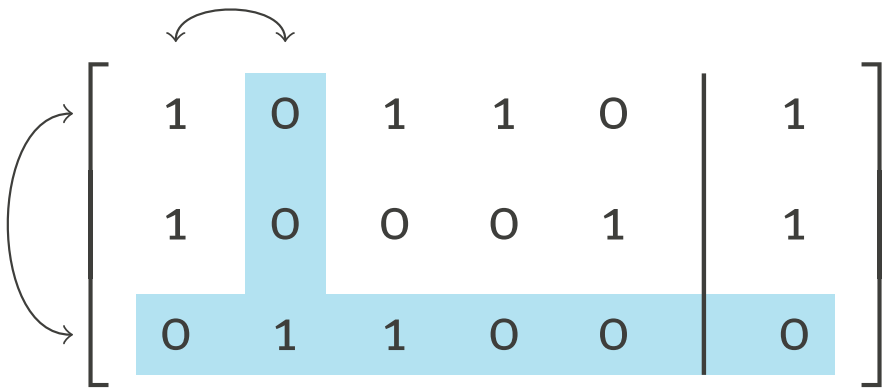
Fridrich et al. (2005a)

living.knowledge WWU Münster

Matrix LT Process

- upper triangle form by permuting rows and columns
- generate Matrix **D** following Robust Soliton Distribution (RSD)

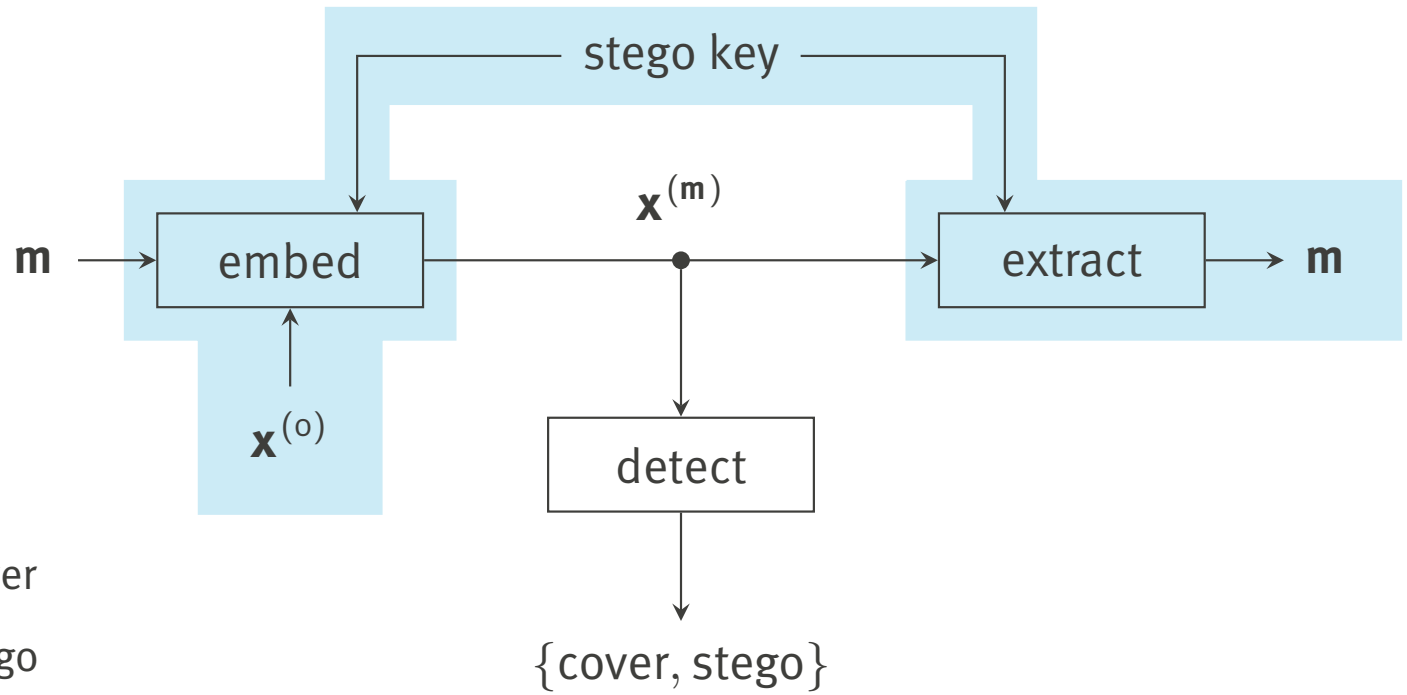
Luby (2002)



Fridrich et al. (2005b)

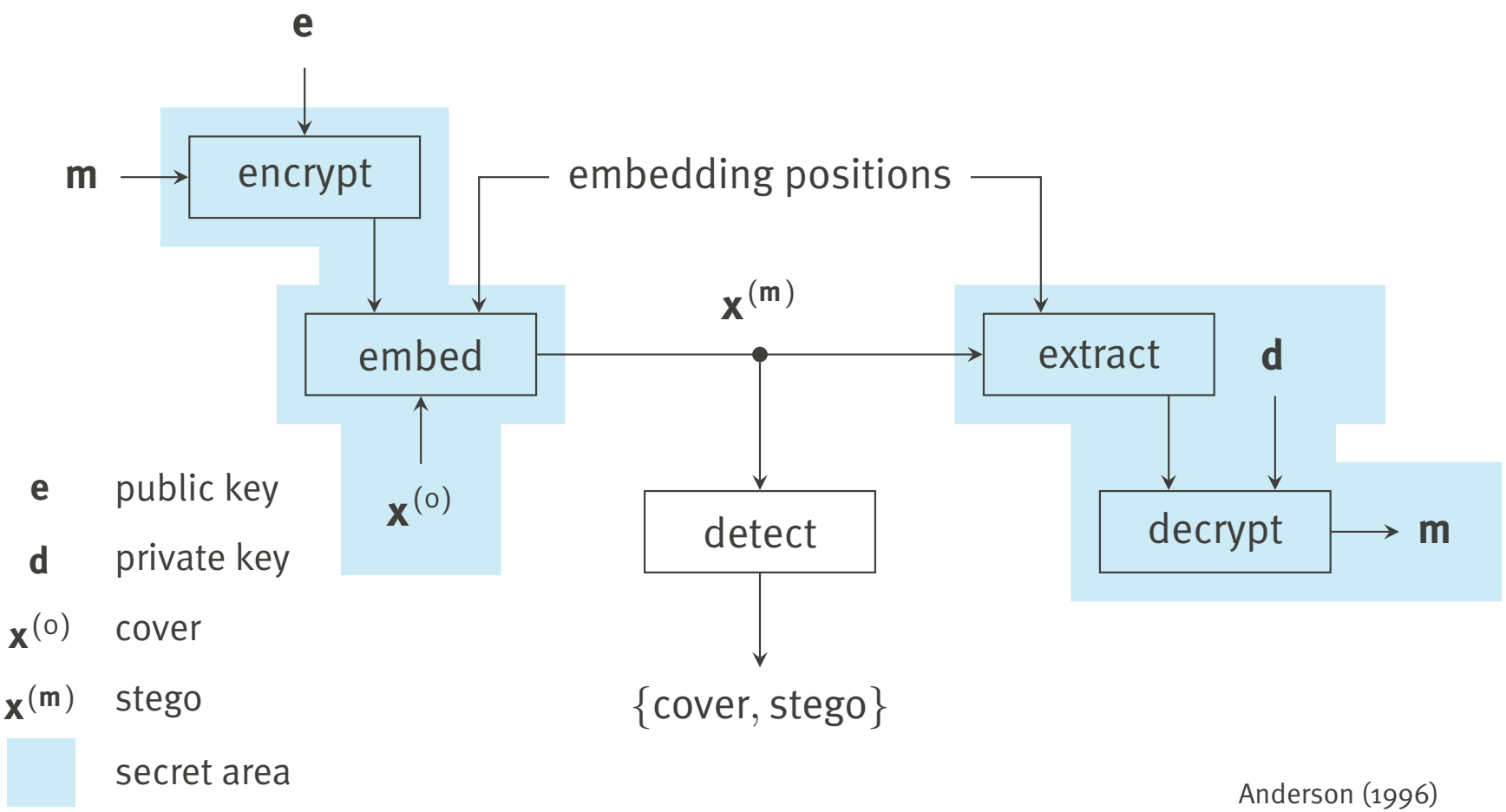
living.knowledge WWU Münster

Steganography



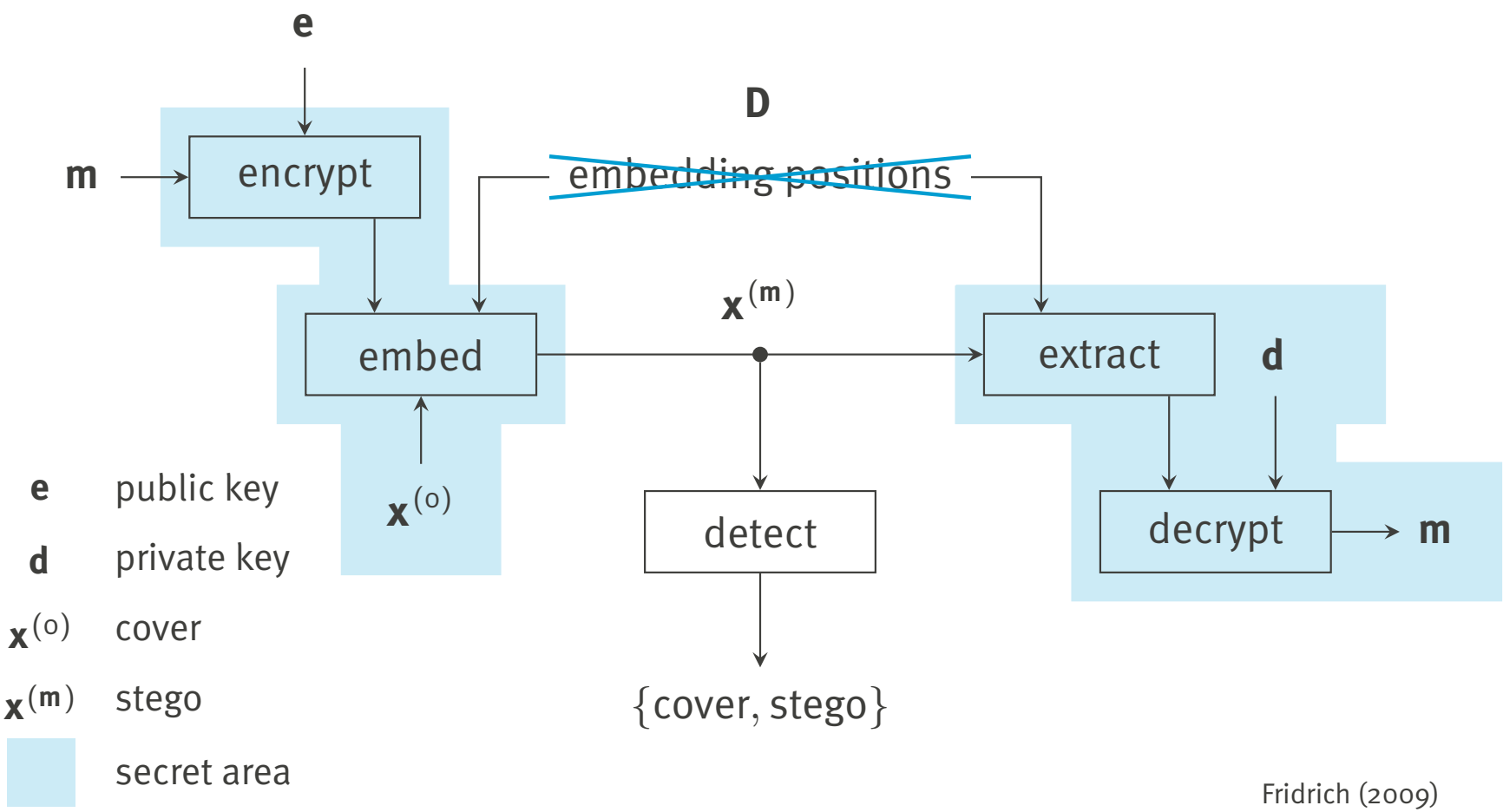
living.knowledge WWU Münster

Public-Key Steganography (PKS)



Anderson (1996)

Public-Key Steganography (PKS)



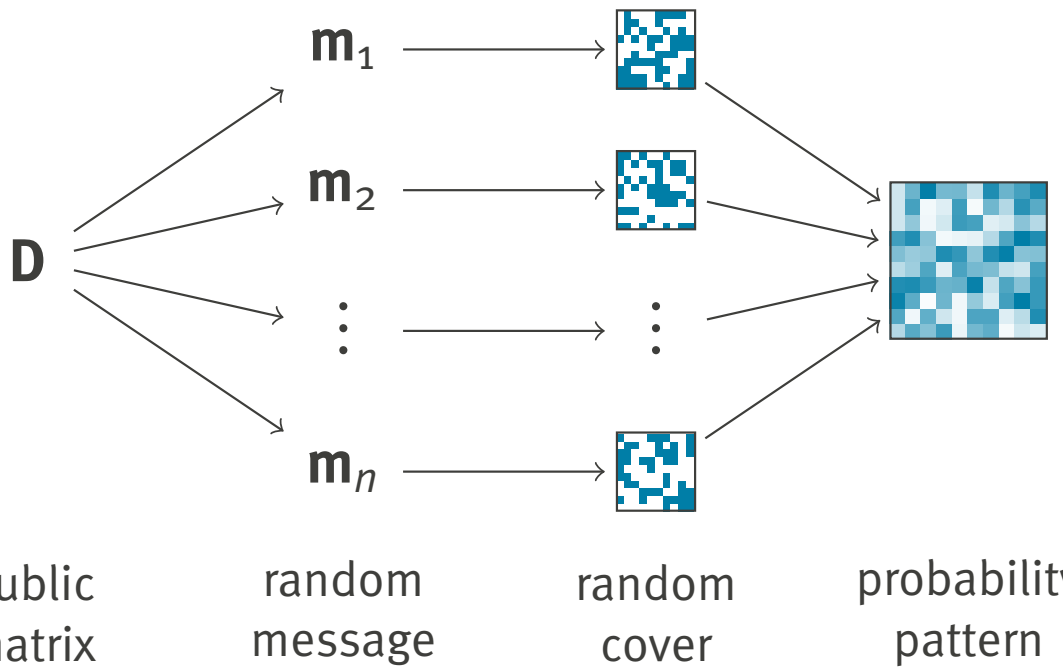
Fridrich (2009)

Outline

living.knowledge WWU Münster

Estimating Embedding Positions

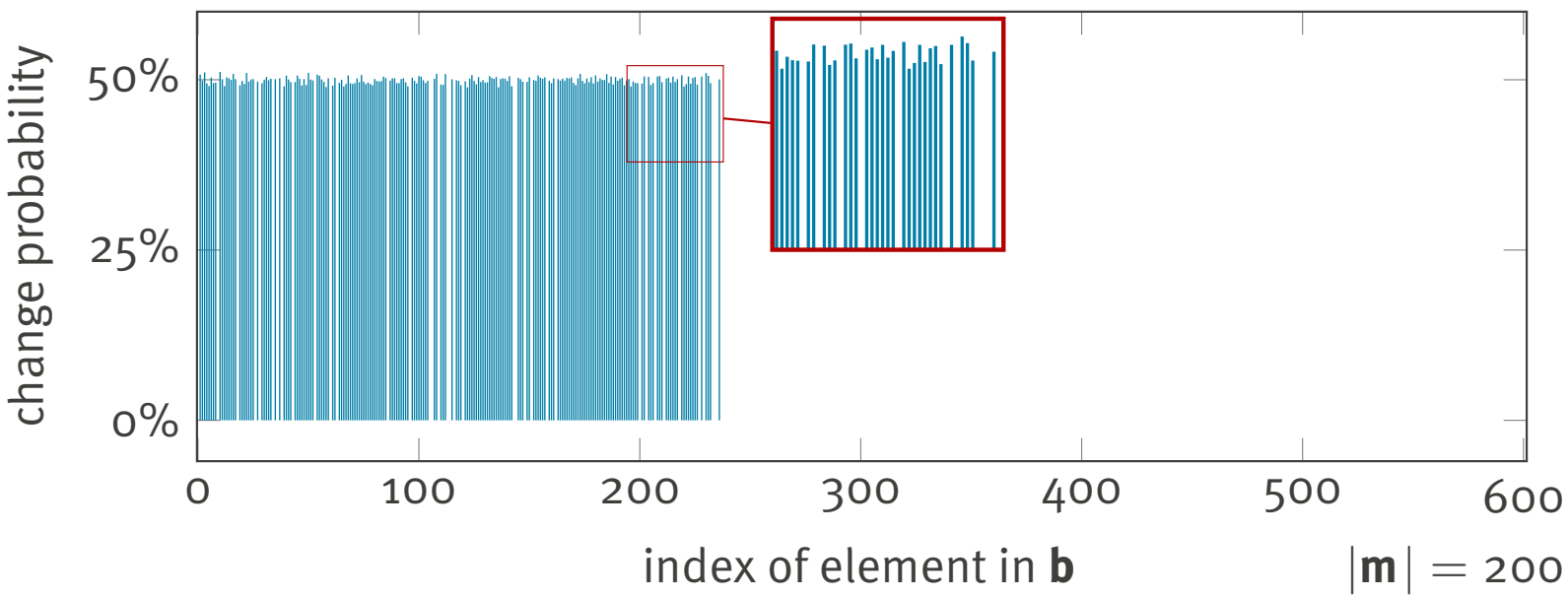
- matrix LT process favours certain columns
- large part of system of linear equations is public due to **D**
- calculate probability patterns based on **D**



living.knowledge WWU Münster

Probability Pattern

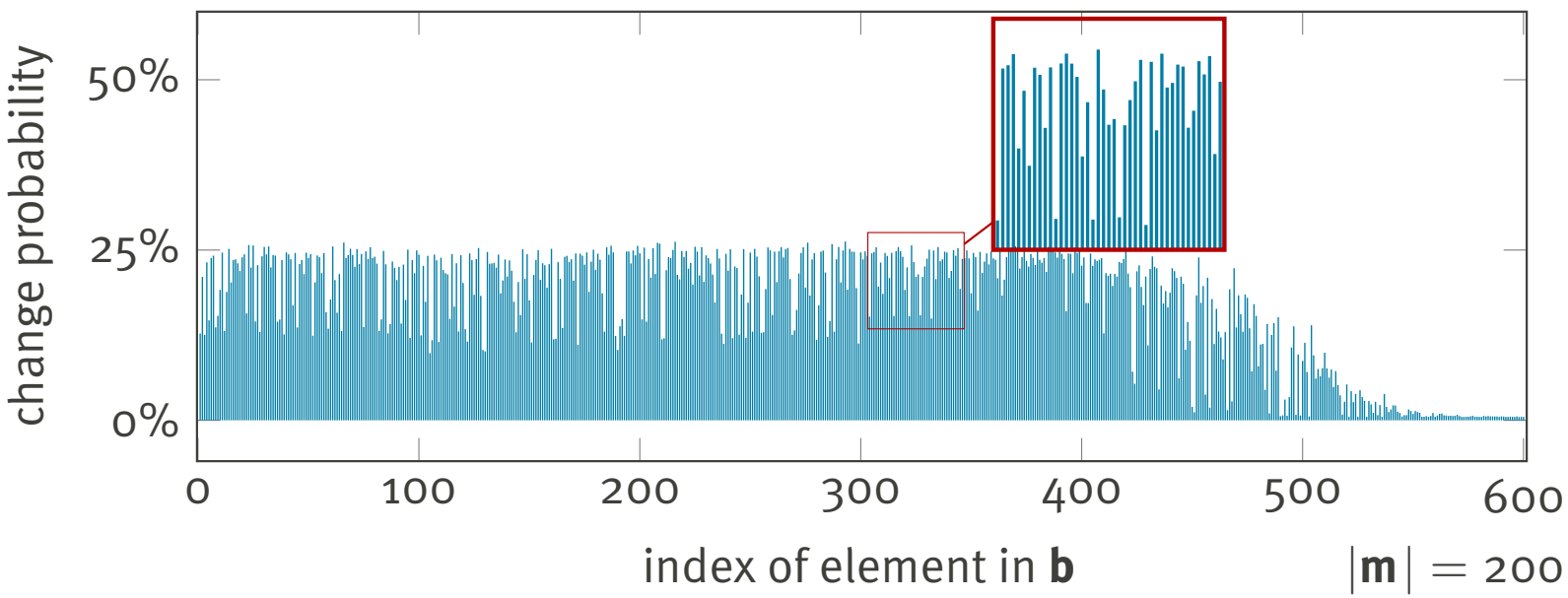
- matrix LT process
- no ‘wet’ pixels



living.knowledge WWU Münster

Probability Pattern

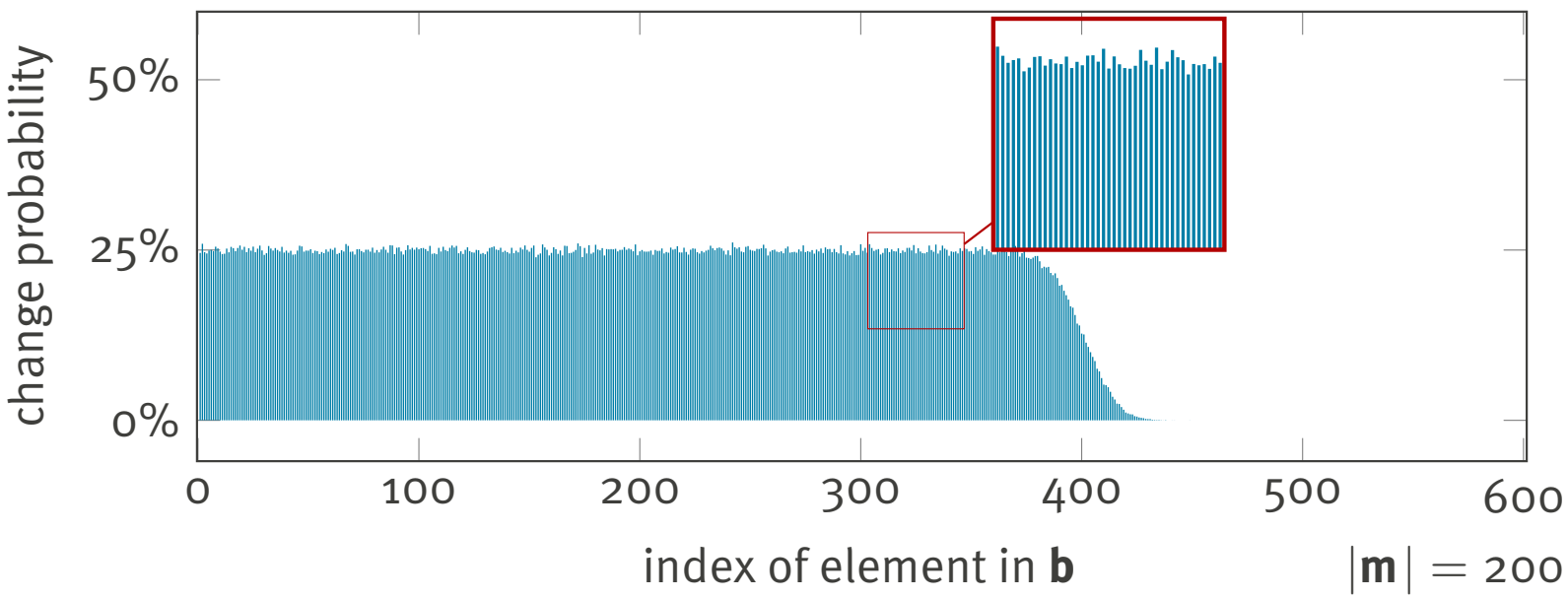
- matrix LT process
- 50% ‘wet’ pixels



living.knowledge
WWU Münster

Probability Pattern

- Gaussian elimination
- 50% 'wet' pixels

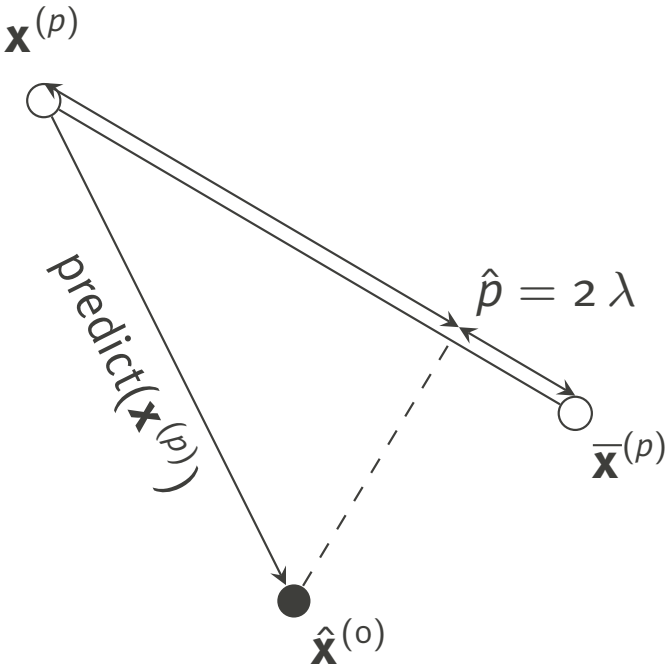


living.knowledge WWU Münster

Outline

living.knowledge
WWU Münster

Weighted Stego-Image (WS) Steganalysis



$\mathbf{x}^{(o)}$	cover
$\hat{\mathbf{x}}^{(o)}$	estimated cover
$\mathbf{x}^{(p)}$	stego
$\overline{\mathbf{x}}^{(p)}$	stego, all LSBs flipped
p	payload
λ	weighting

Weighted Stego-Image (WS):

$$\mathbf{x}^{(p,\lambda)} = \lambda \overline{\mathbf{x}}^{(p)} + (1-\lambda)\,\mathbf{x}^{(p)}$$

Estimator of embedding rate:

$$\hat{p} = 2 \arg\min_{\lambda} \sum_{i=1}^{n} \left(x_i^{(p,\lambda)} - \hat{x}_i^{(o)} \right)^2$$

Theorem 1 of Fridrich and Goljan (2004)

living.knowledge WWU Münster

Adding Predictability Weights

x_1	x_2	x_3	x_4	x_5
x_6	x_7	$\hat{x}_8$	x_9	x_{10}
x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
x_{16}	x_{17}	x_{18}	x_{19}	x_{20}
x_{21}	x_{22}	x_{23}	x_{24}	x_{25}

Weighted WS Steganalysis

$$\hat{p} = 2 \arg\min_{\lambda} \sum_{i=1}^{n} w_i \left(x_i^{(p,\lambda)} - \hat{x}_i^{(0)} \right)^2$$

- weighting takes predictability into account

Theorem 3 of Fridrich and Goljan (2004)

living.knowledge WWU Münster

Adding Predictability Weights

x_1	x_2	x_3	x_4	x_5
x_6	x_7	x_8	x_9	x_{10}
x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
x_{16}	x_{17}	x_{18}	x_{19}	x_{20}
x_{21}	x_{22}	x_{23}	x_{24}	x_{25}

Probability WS Steganalysis

$$\hat{p} = 2 \arg\min_{\lambda} \sum_{i=1}^{n} \omega_i \left(x_i^{(p,\lambda)} - \hat{x}_i^{(0)} \right)^2$$

- weighting takes predictability into account
- replace weight with normalized probabilities:

$$\omega_i = \frac{p_i}{\sum_{j=1}^{n} p_j}$$

Theorem 3 of Fridrich and Goljan (2004)

living.knowledge WWU Münster

Sorting the Stego-Image

p_2	p_{25}	p_1	p_4	p_{22}
p_7	p_8	p_{17}	p_{14}	p_{15}
p_{18}	p_5	p_{16}	p_{11}	p_{24}
p_{10}	p_9	p_{13}	p_{21}	p_{19}
p_3	p_6	p_{12}	p_{23}	p_{20}

Sequential WS Steganalysis

$$E(l) = \sum_{i=1}^{l} \left(\frac{1}{2} \left(x_i^{(p)} + \bar{x}_i^{(p)} \right) - \hat{x}_i^{(o)} \right)^2 + \sum_{i=l+1}^{n} \left(x_i^{(p)} - \hat{x}_i^{(o)} \right)^2$$

- assumes initial sequential embedding

Ker (2007)

living.knowledge WWU Münster

Sorting the Stego-Image

x_2	x_{25}	x_1	x_4	x_{22}
x_7	x_8	x_{17}	x_{14}	x_{15}
x_{18}	x_5	x_{16}	x_{11}	x_{24}
x_{10}	x_9	x_{13}	x_{21}	x_{19}
x_3	x_6	x_{12}	x_{23}	x_{20}

Sequential WS Steganalysis

$$E(l) = \sum_{i=1}^{l} \left(\frac{1}{2} \left(x_i^{(p)} + \bar{x}_i^{(p)} \right) - \hat{x}_i^{(o)} \right)^2 + \sum_{i=l+1}^{n} \left(x_i^{(p)} - \hat{x}_i^{(o)} \right)^2$$

- assumes initial sequential embedding
- reorder stego-image according to probabilities

Ker (2007)

living.knowledge WWU Münster

Sorting the Stego-Image

y_1	y_2	y_3	y_4	y_5
y_6	y_7	y_8	y_9	y_{10}
y_{11}	y_{12}	y_{13}	y_{14}	y_{15}
y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
y_{21}	y_{22}	y_{23}	y_{24}	y_{25}

Sorted WS Steganalysis

$$E(l) = \sum_{i=1}^{l} \left(\frac{1}{2} \left(y_i^{(p)} + \overline{y}_i^{(p)} \right) - \hat{y}_i^{(o)} \right)^2 + \sum_{i=l+1}^{n} \left(y_i^{(p)} - \hat{y}_i^{(o)} \right)^2$$

- assumes initial sequential embedding
- reorder stego-image according to probabilities

Ker (2007)

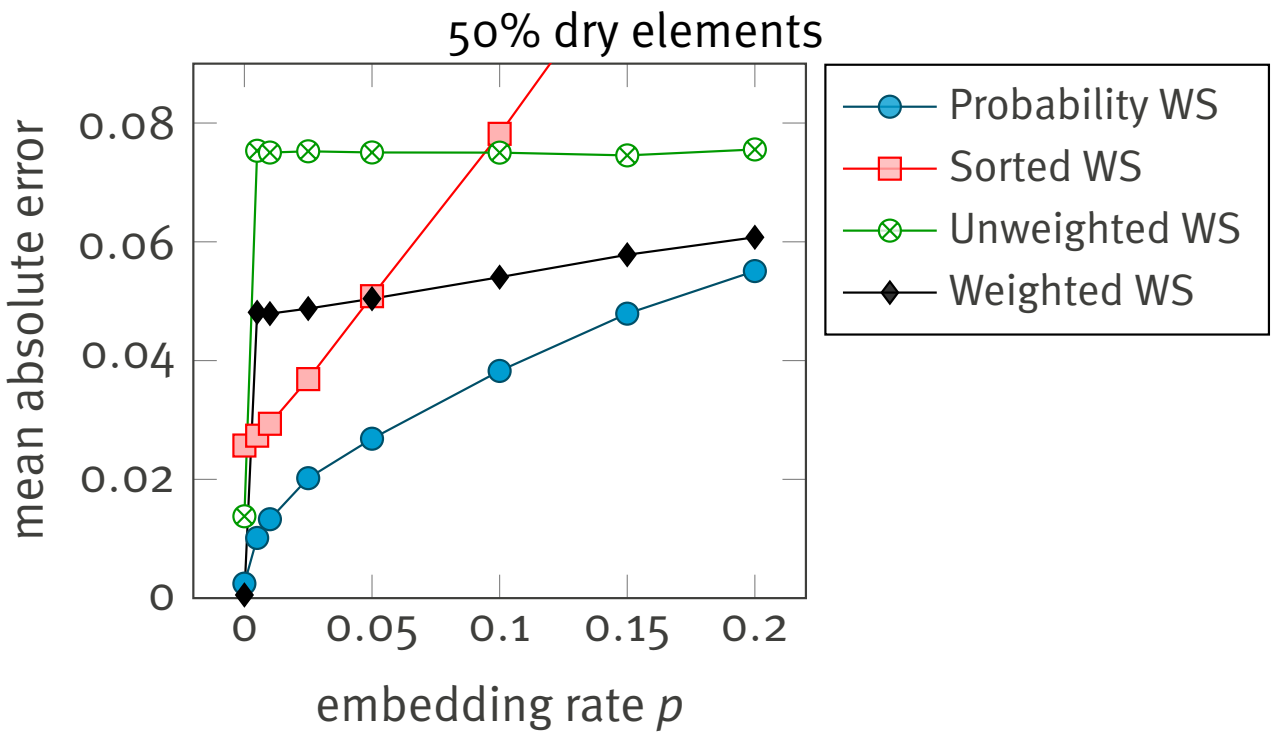
living.knowledge WWU Münster

Outline

living.knowledge WWU Münster

Empirical Results

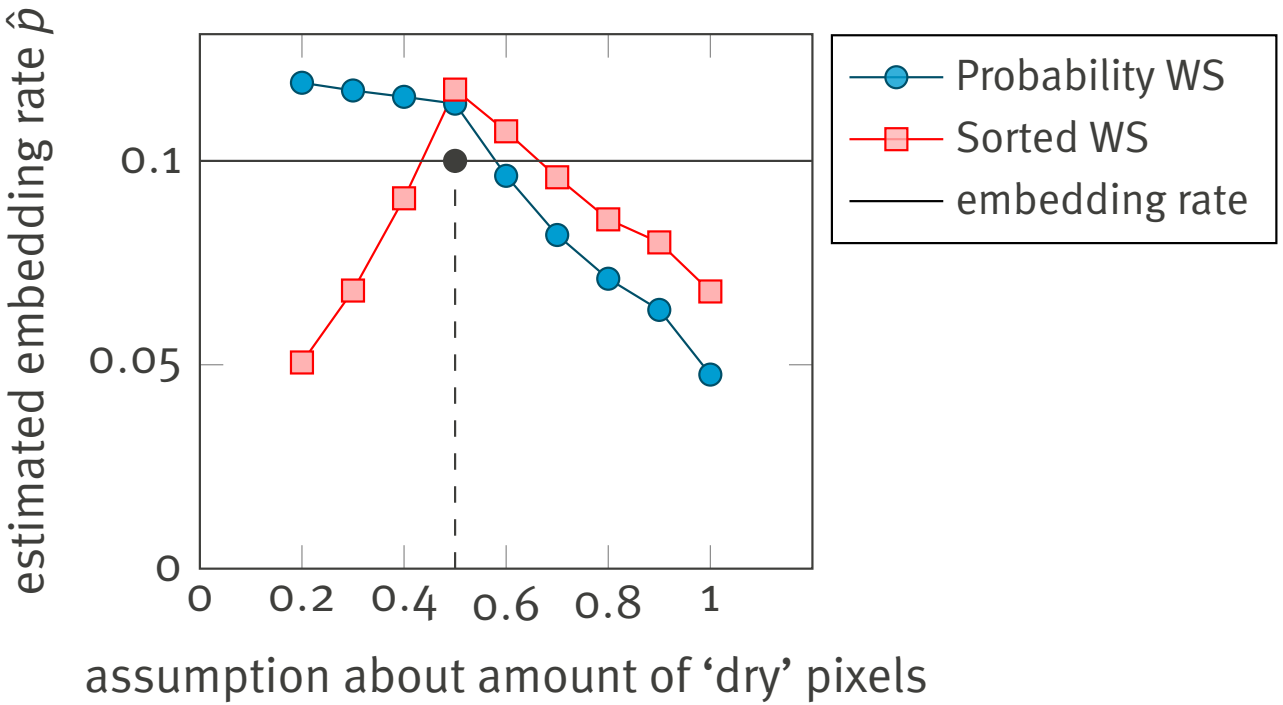
- 10,000 BOSSBase images, cropped to 128 × 128 pixels Bas et al. (2011)
- amount of dry elements unknown



living.knowledge WWU Münster

Empirical Results

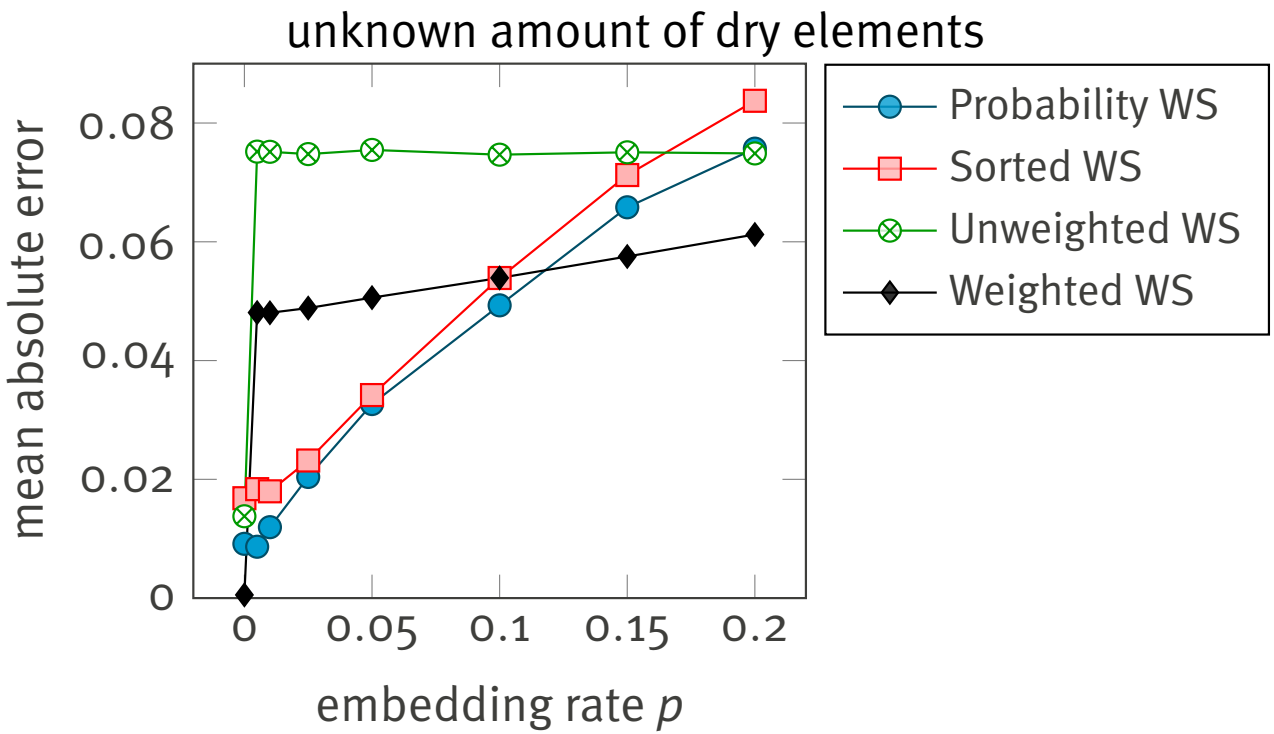
- 10,000 BOSSBase images, cropped to 128 × 128 pixels Bas et al. (2011)
- amount of dry elements unknown
- estimate by using the point where both methods yield a similar result



living.knowledge WWU Münster

Empirical Results

- 10,000 BOSSBase images, cropped to 128×128 pixels Bas et al. (2011)
- amount of dry elements unknown
- estimate by using the point where both methods yield a similar result

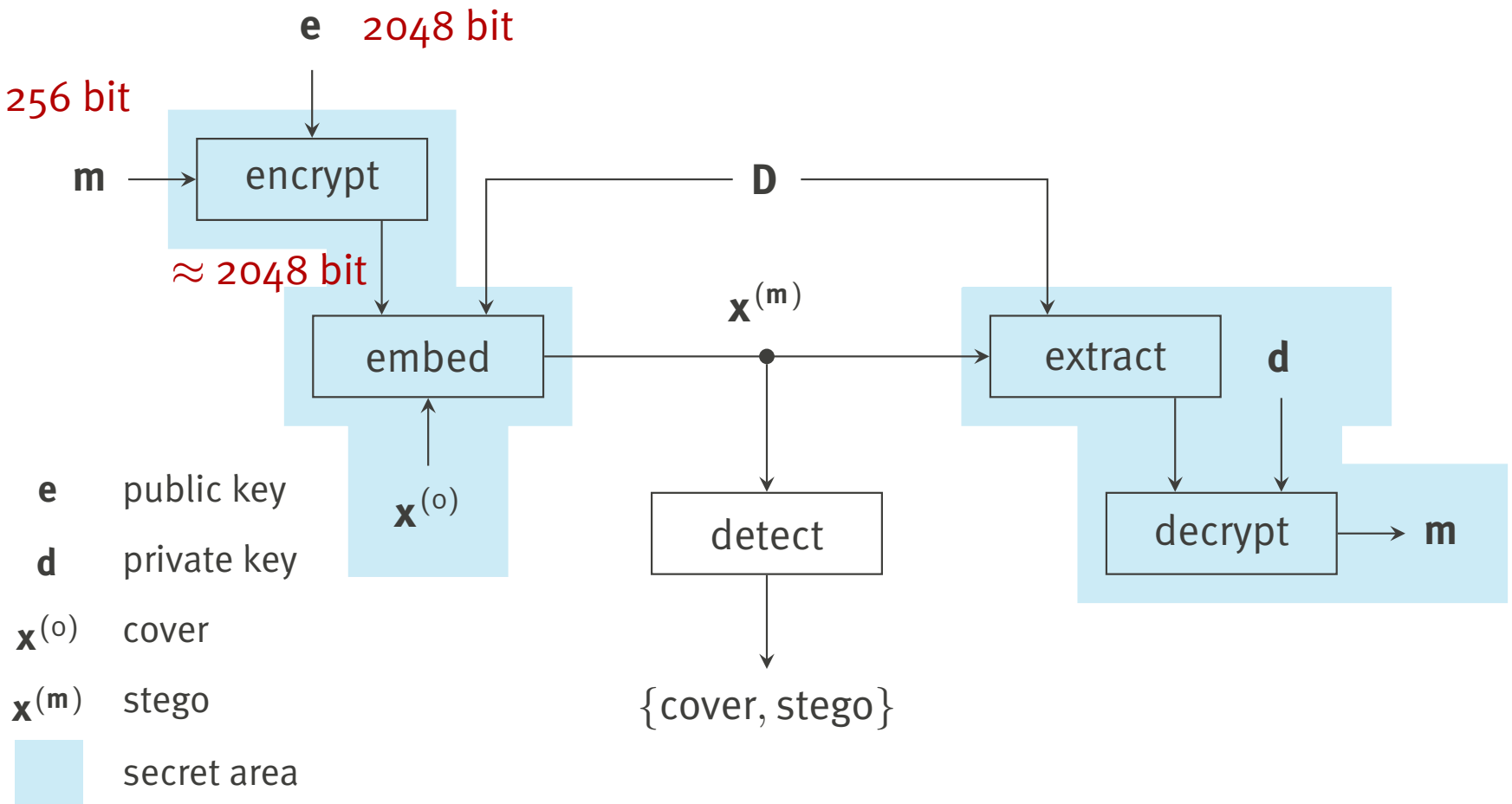


living.knowledge WWU Münster

Outline

living.knowledge
WWU Münster

Relevance



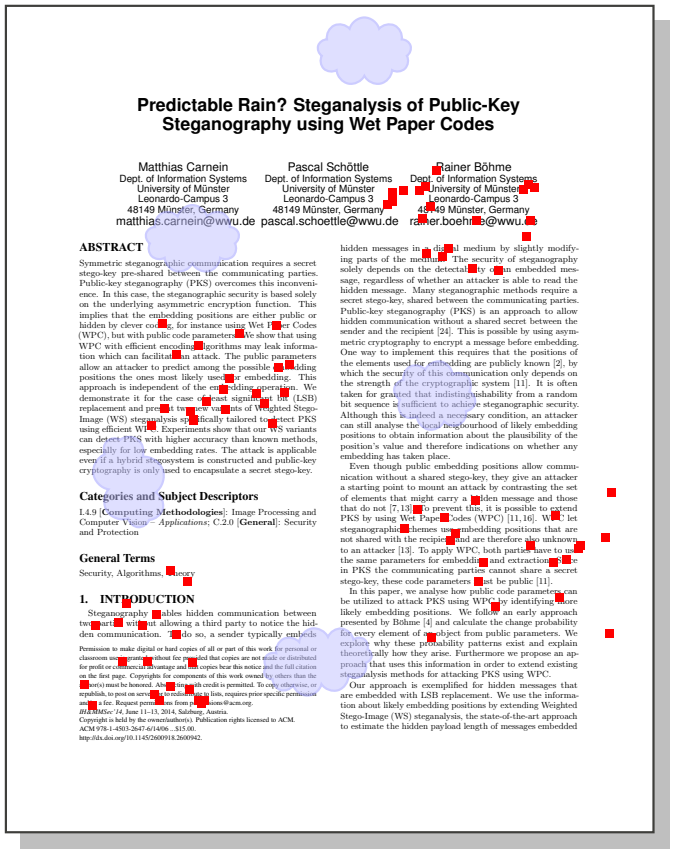
living.knowledge WWU Münster

Conclusion

- WPC enable secret embedding positions in PKS
- public matrix allows calculation of probability patterns
- probability patterns independent of used embedding method
- extend existing attacks to make them applicable
  - shown for LSB-Replacement
  - two extensions of WS steganalysis
- also relevant in a key encapsulation scheme
- Predictable Rain?

living.knowledge WWU Münster

Predictable Rain ?



Predictable Rain? Steganalysis of Public-Key Steganography using Wet Paper Codes

Matthias Carnein
Dept. of Information Systems
University of Münster
Leonardo-Campus 3
48149 Münster, Germany
matthias.carnein@wwu.de

Pascal Schöttle
Dept. of Information Systems
University of Münster
Leonardo-Campus 3
48149 Münster, Germany
pascal.schoettle@wwu.de

Rainer Böhme
Dept. of Information Systems
University of Münster
Leonardo-Campus 3
48149 Münster, Germany
rainer.boehme@wwu.de

ABSTRACT

Symmetric steganographic communication requires a secret stego-key pre-shared between the communicating parties. Public-key steganography (PKS) overcomes this inconvenience. In this case, the steganographic security is based solely on the underlying asymmetric encryption function. This implies that the embedding positions are either public or hidden by clever coding, for instance using Wet Paper Codes (WPC), but with public code parameters. We show that using WPC with efficient encoding algorithms may leak information which can facilitate an attack. The public parameters allow an attacker to predict among the possible embedding positions the ones most likely used for embedding. This approach is independent of the embedding operation. We demonstrate it for the case of least significant bit (LSB) replacement and present two practicable variants of Weighted Stego-Image (WS) steganalysis specifically tailored to detect PKS using efficient WPC. Experiments show that our WS variants can detect PKS with higher accuracy than known methods, especially for low embedding rates. The attack is applicable even if a hybrid stegosystem is constructed and public-key cryptography is only used to encapsulate a secret stego-key.

Categories and Subject Descriptors

I.4.9 [**Computing Methodologies**]: Image Processing and Computer Vision – *Applications*; C.2.0 [**General**]: Security and Protection

General Terms

Security, Algorithms, Theory

1. INTRODUCTION

Steganography enables hidden communication between two parties without allowing a third party to notice the hidden communication. To do so, a sender typically embeds hidden messages in a digital medium by slightly modifying parts of the medium. The security of steganography solely depends on the detectability of an embedded message, regardless of whether an attacker is able to read the hidden message. Many steganographic methods require a secret stego-key, shared between the communicating parties. Public-key steganography (PKS) is an approach to allow hidden communication without a shared secret between the sender and the recipient [24]. This is possible by using asymmetric cryptography to encrypt a message before embedding. One way to implement this requires that the positions of the elements used for embedding are publicly known [2], by which the security of this communication only depends on the strength of the cryptographic system [11]. It is often taken for granted that indistinguishability from a random bit sequence is sufficient to achieve steganographic security. Although this is indeed a necessary condition, an attacker can still analyse the local neighbourhood of likely embedding positions to obtain information about the plausibility of the position's value and therefore indications on whether any embedding has taken place.

Even though public embedding positions allow communication without a shared stego-key, they give an attacker a starting point to mount an attack by contrasting the set of elements that might carry a hidden message and those that do not [7, 13]. To prevent this, it is possible to extend PKS by using Wet Paper Codes (WPC) [11, 16]. With it let steganographic schemes use embedding positions that are not shared with the recipient and are therefore also unknown to an attacker [13]. To apply WPC, both parties have to use the same parameters for embedding and extraction. Since in PKS the communicating parties cannot share a secret stego-key, these code parameters must be public [11].

In this paper, we analyse how public code parameters can be utilized to attack PKS using WPC by identifying more likely embedding positions. We follow an early approach presented by Böhme [4] and calculate the change probability for every element of an object from public parameters. We explain why these probability patterns exist and explain theoretically how they arise. Furthermore we propose an approach that uses this information in order to extend existing steganalysis methods for attacking PKS using WPC.

Our approach is exemplified for hidden messages that are embedded with LSB replacement. We use the information about likely embedding positions by extending Weighted Stego-Image (WS) steganalysis, the state-of-the-art approach to estimate the hidden payload length of messages embedded

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
IH&MMSec '14, June 11–13, 2014, Salzburg, Austria.
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-2647-6/14/06...$15.00.
http://dx.doi.org/10.1145/2600918.2600942.

Conclusion

- WPC enable secret embedding positions in PKS
- public matrix allows calculation of probability patterns
- probability patterns independent of used embedding method
- extend existing attacks to make them applicable
  - shown for LSB-Replacement
  - two extensions of WS steganalysis
- also relevant in a key encapsulation scheme
- Predictable Rain?
  - predict the most likely *hideouts in the storm*

living.knowledge WWU Münster

Questions?

living.knowledge WWU Münster

References I

Anderson, Ross J. (1996). 'Stretching the limits of steganography'. In: *Information Hiding (1st International Workshop)*. Ed. by Ross J. Anderson. Vol. 1174. Lecture Notes in Computer Science. Berlin Heidelberg: Springer-Verlag, pp. 39–48.

Bas, Patrick, Tomáš Filler and Tomáš Pevný (2011). 'Break our steganographic system — the ins and outs of organizing BOSS'. In: *Information Hiding (13th International Workshop)*. Ed. by Tomáš Filler, Tomáš Pevný, Scott Craver and Andrew Ker. Vol. 6958. Lecture Notes in Computer Science. Berlin Heidelberg: Springer-Verlag, pp. 59–70.

Fridrich, Jessica (2009). *Steganography in Digital Media. Principles, Algorithms, and Applications*. 1st. New York, NY, USA: Cambridge University Press.

Fridrich, Jessica and Miroslav Goljan (2004). 'On estimation of secret message length in LSB steganography in spatial domain'. In: *Security, Steganography, and Watermarking of Multimedia Contents VI*. Ed. by Edward J. Delp and Ping Wah Wong. Vol. 5306. Proceedings of SPIE. San Jose, CA: SPIE, pp. 23–34.

Fridrich, Jessica, Miroslav Goljan, Petr Losiněk and David Soukal (2005a). 'Writing on wet paper'. In: *IEEE Transactions on Signal Processing* 53.10, pp. 3923–3935.

living.knowledge WWU Münster

References II

Fridrich, Jessica, Miroslav Goljan, Petr Losiněk and David Soukal (2005b). 'Writing on wet paper'. In: *Security, Steganography and Watermarking of Multimedia Contents VII*. Ed. by Edward J. Delp and Ping Wah Wong. Vol. 5681. Proceedings of SPIE. San Jose, CA: SPIE, pp. 328–340.

Fridrich, Jessica, Miroslav Goljan and David Soukal (2004). 'Perturbed quantization steganography with wet paper codes'. In: *Proceedings of the ACM workshop on Multimedia and Security (MM&Sec)*. New York, NY: ACM Press, pp. 4–15.

Ker, Andrew D. (2007). 'A weighted stego image detector for sequential LSB replacement'. In: *Third International Symposium on Information Assurance and Security*. IEEE Computer Society, pp. 453–456.

Luby, Michael (2002). 'LT codes'. In: *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*. FOCS 2002. IEEE Computer Society, pp. 271–280.

living.knowledge WWU Münster