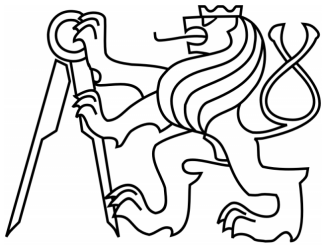
Modern Steganalysis Can Detect YASS

Jan Kodovský, Tomáš Pevný, Jessica Fridrich

January 18, 2010 / SPIE



YASS – Curriculum Vitae

- Birth Location: University of California, Santa Barbara
- Birth Date: More than 2 years ago [Solanki-2007], [Sarkar-2008]
- Deviation from the paradigm of minimizing embedding impact
- Steganalysis of 2007 failed to detect YASS reliably
- Two challenges for steganalysts:
  - Embedding in key-dependent domain
  - Embedding masked by JPEG compression

What Is This Talk About

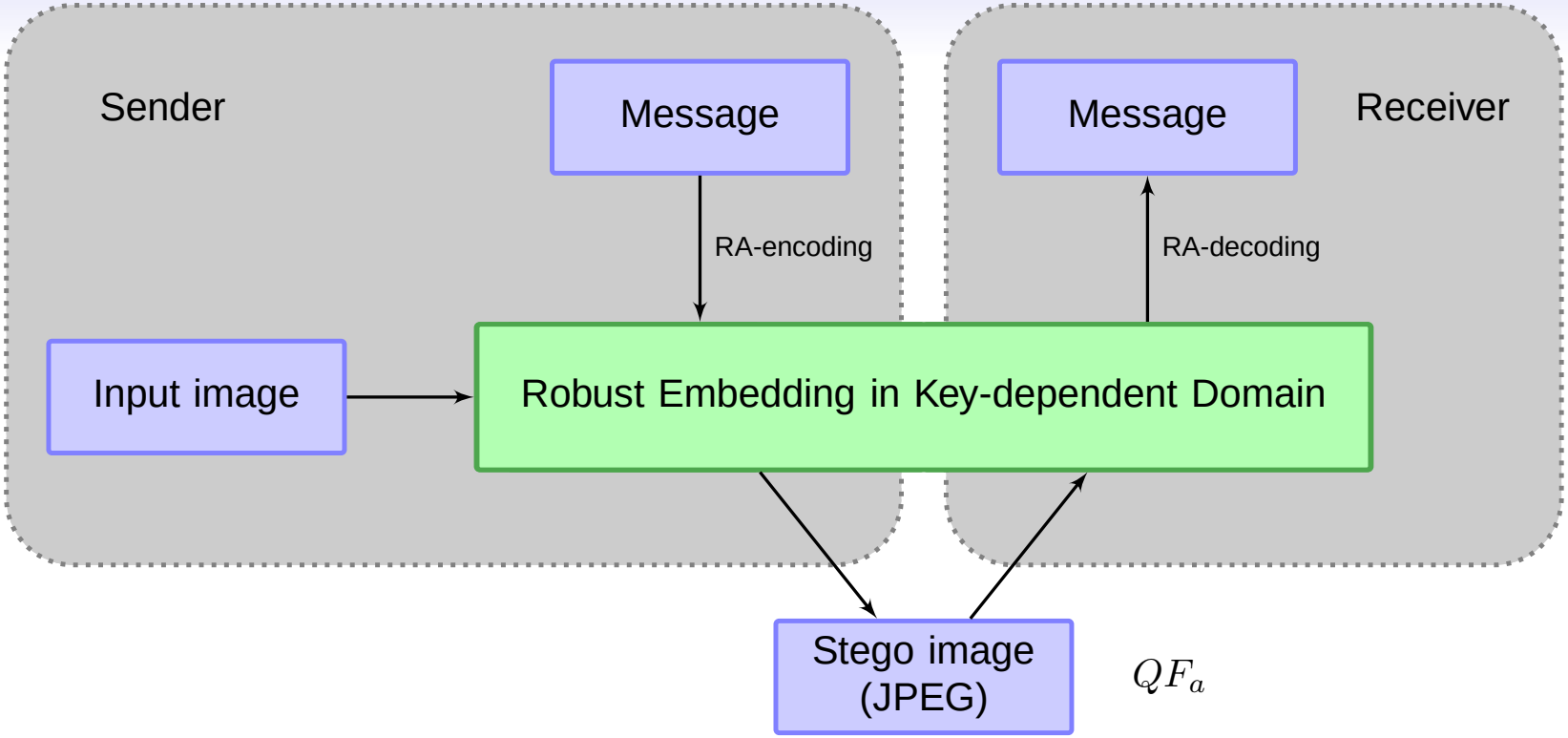
- YASS is indeed detectable (even for small images and payloads)
- Tool: state-of-the-art blind steganalysis
  - Several different general-purpose feature-sets
  - No utilization of implementation flaws of YASS
- Extended versions of YASS involved in tests as well
- Performance comparison to other methods

YASS = embedding paradigm

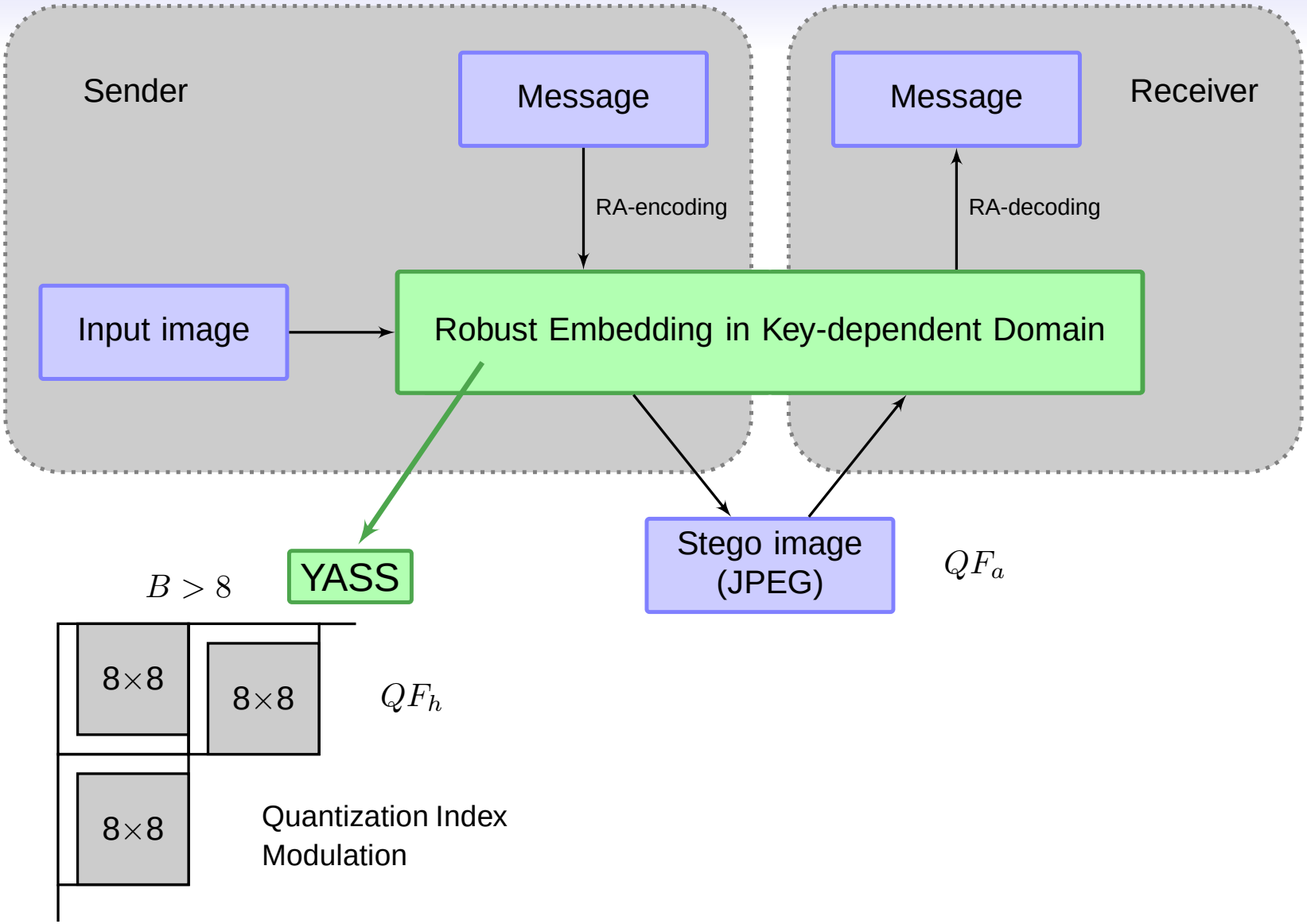
Selected Existing Attacks

- [Solanki-2007], [Sarkar-2008] – first blind attacks
  - YASS outperforms Outguess, Steghide and F5
- [Li-2008] – accurate targeted attack
  - YASS is not randomized enough
- [Huang-2008] – important insight
  - YASS effectively disables calibration
  - MB1 outperforms YASS
- [Kodovský-2009] – calibration revisited
  - Improved way of calibration – Cartesian product
  - Steganography minimizing emb. impact (MME3 and nsF5) is more secure than YASS

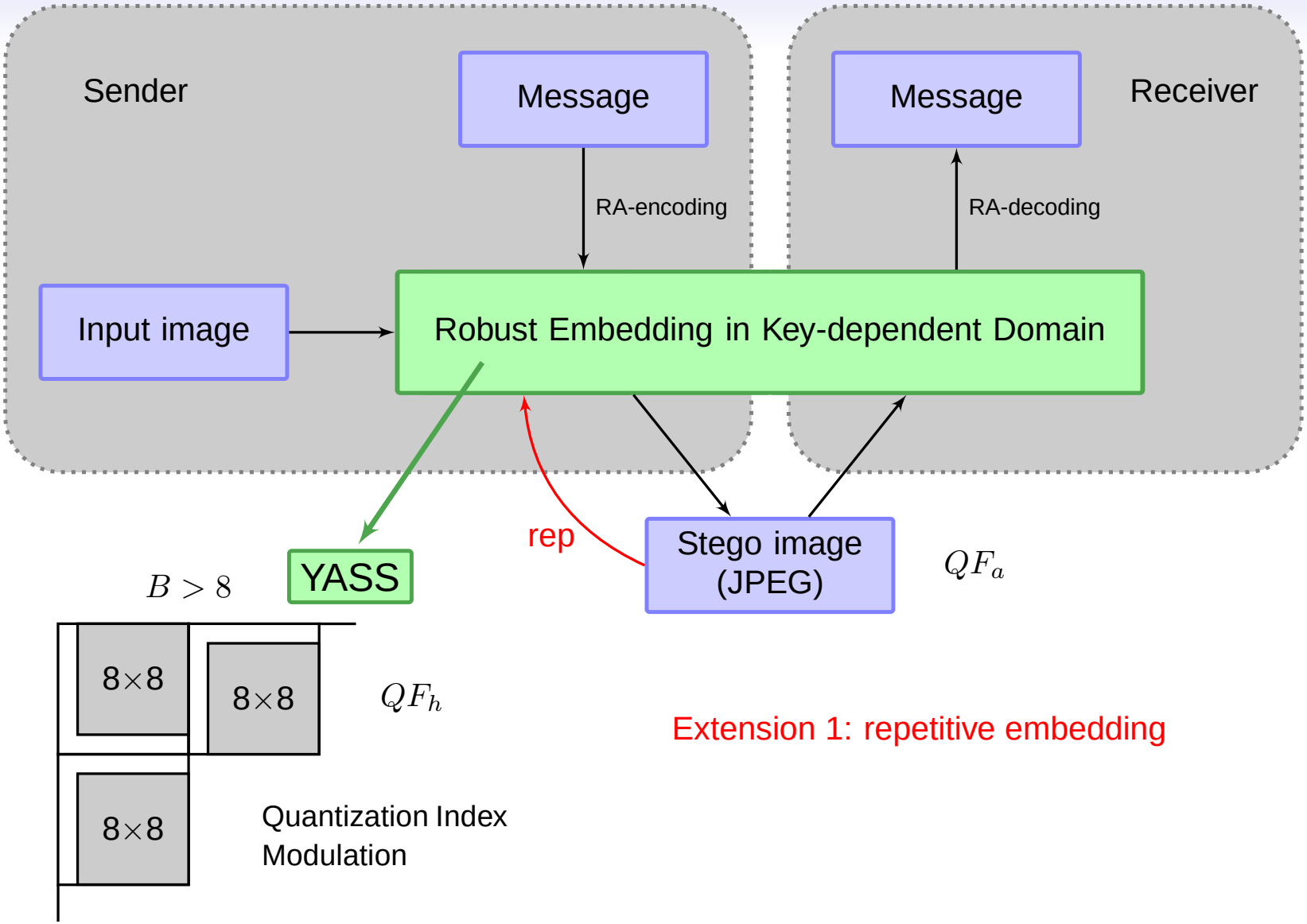
Mechanism of YASS



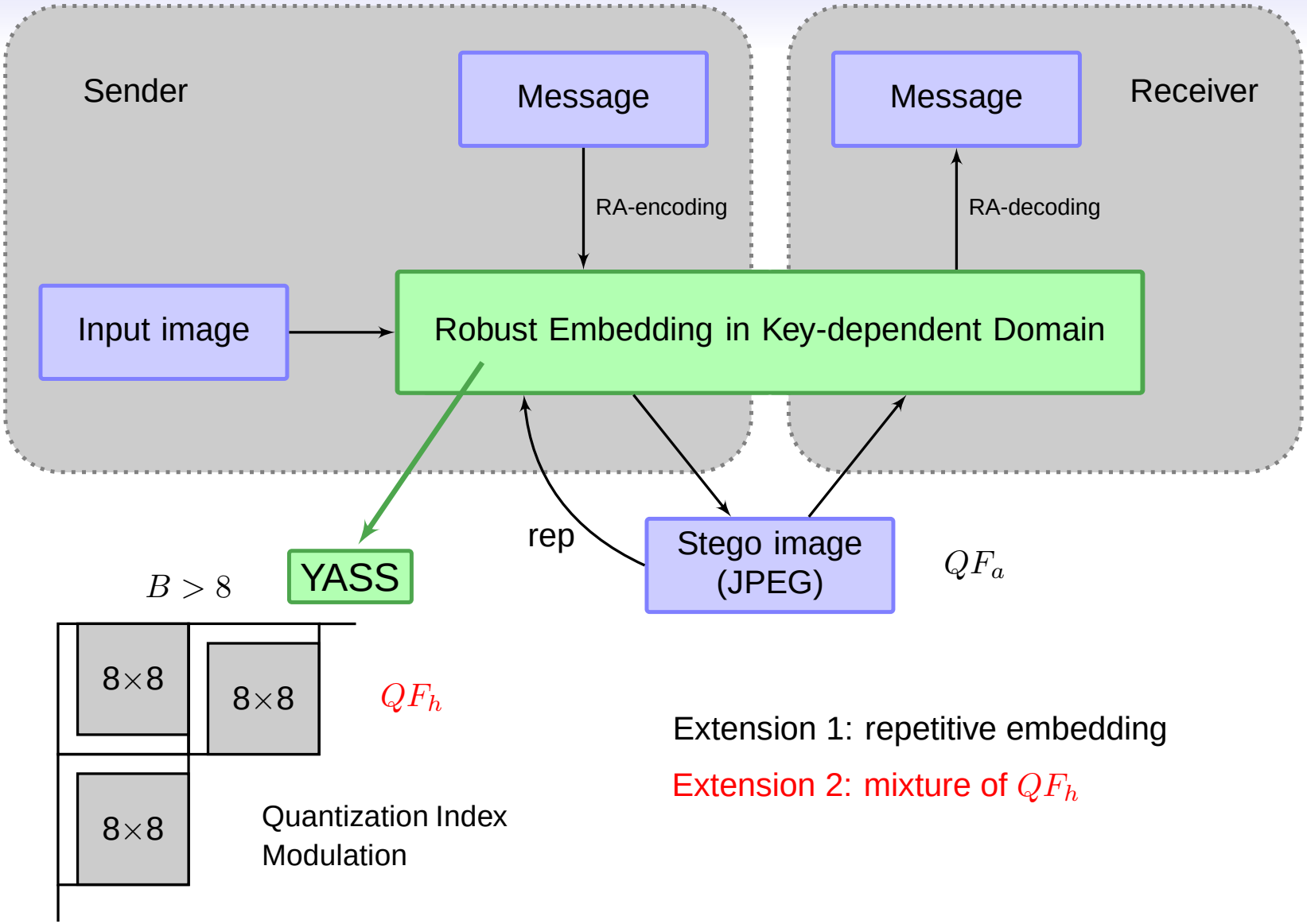
Mechanism of YASS



Mechanism of YASS



Mechanism of YASS



Extension 1: repetitive embedding

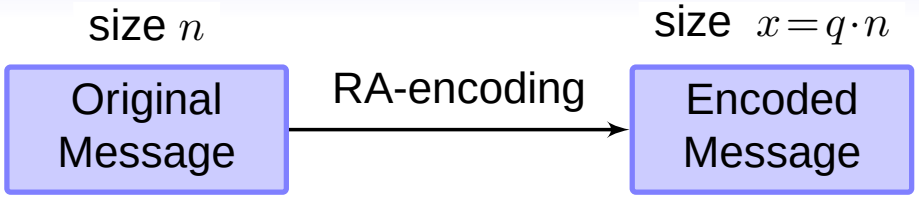
Extension 2: mixture of QF_h

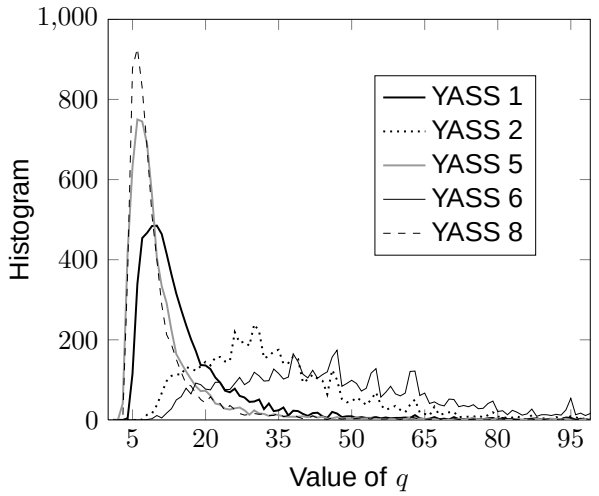
Different Setup Used in Our Tests

Setting	Extension	B	rep	QF_h	DBs	Bpac
YASS 1	2	9	0	65,70,75	3,7	0.110
YASS 2	none	9	0	75	-	0.051
YASS 3	1	9	1	75	-	0.187
YASS 4	2	9	0	65,70,75	2,5	0.118
YASS 5	2	9	0	50,55,60,65,70	3,7,12,17	0.159
YASS 6	none	10	0	75	-	0.031
YASS 7	2	10	0	65,70,75	3,7	0.078
YASS 8	1	10	1	75	-	0.138
YASS 9	both	9	2	65,70,75	3,7	0.237
YASS 10	1	10	2	75	-	0.159
YASS 11	1	11	1	75	-	0.114
YASS 12	2	11	0	65,70,75	3,7	0.077

$QF_a = 75$, Input image format: RAW (uncompressed)

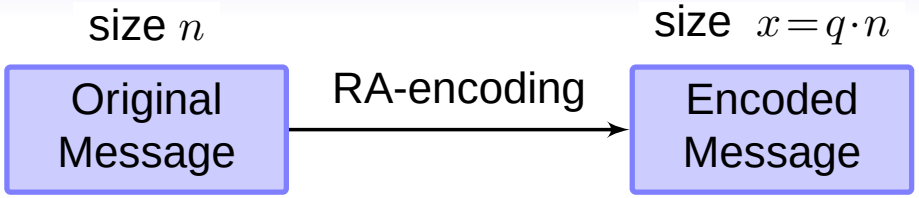
Determining The Payload

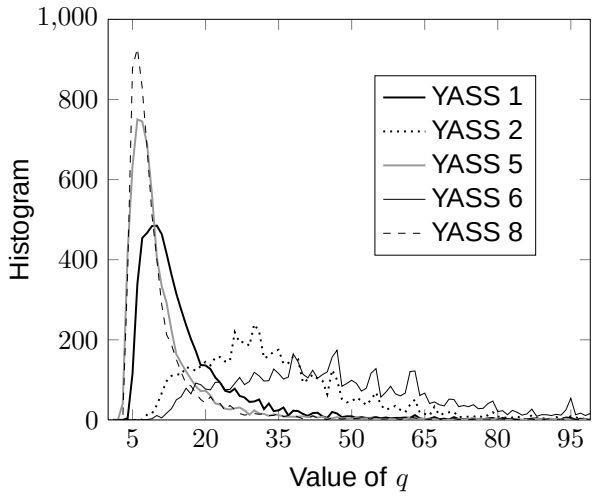




- Difficulties:
  - YASS embeds only full payload
  - YASS outputs x instead of n
- Existing approaches to this issue:
  - Do not report payload or report RA-encoded payload x
  - Report x/q for some value of q (fixed/random)
  - Report lower and upper bounds x/q_1 and x/q_2
  - Determine q for every image directly
    - Use estimate of q [Sarkar-2008]
    - Use repetitive embedding to determine the value of q

Determining The Payload





- Difficulties:
  - YASS embeds only full payload
  - YASS outputs x instead of n
- Existing approaches to this issue:
  - Do not report payload or report RA-encoded payload x
  - Report x/q for some value of q (fixed/random)
  - Report lower and upper bounds x/q_1 and x/q_2
  - Determine q for every image directly
    - Use estimate of q [Sarkar-2008]
    - Use repetitive embedding to determine the value of q

$x, q \Rightarrow$ calculate n, take average over all images

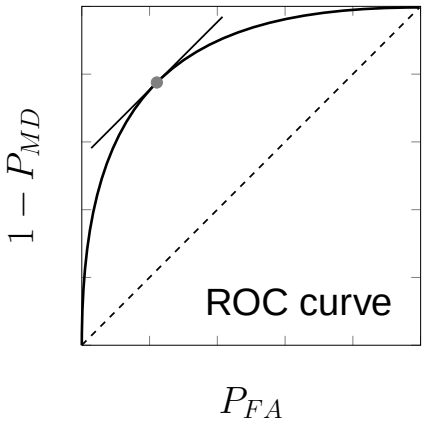
Steganalysis Feature Sets

- MP (486) – Markov Process
  - Sample transition probability matrices of 1^{st} order Markov chains of DCT coefficients (within and between DCT blocks)
  - Introduced in [Chen-2008]
- CC-PEV ($2 \times 274 = 548$) – Cartesian calibrated Pevný features
  - Basis: 274 features [Pevný-2007]
  - Introduced in [Kodovský-2009]
- SPAM (686) – Subtractive Pixel Adjacency Model
  - Differences between pixels modeled as 2^{nd} order Markov chains
  - Introduced in [Pevný-2009]
- CDF (1,234) – Cross-Domain Features
  - Merged CC-PEV and SPAM features

Steganalysis Methodology

- Testing database
  - 6,500 images acquired in the raw format
  - Converted to 8-bit grayscale, resized to 512 pixels
- Classification tool
  - Soft-margin SVM with Gaussian Kernel
  - Hyperparameters (C, γ) optimized over a fixed grid of values
  - Five-fold cross-validation
- Measure of security
  - Minimal probability of misclassification P_E
  - Equal prior probabilities of cover and stego

$$P_E = \min \tfrac{1}{2} \left(P_{FA} + P_{MD} \right)$$



Experimental Results

		MP	CC-PEV	SPAM	← feature set
Algorithm	bpac	(486)	(548)	(686)	← dimension
YASS 1	0.110	0.110	0.123	0.140	
YASS 2	0.051	0.155	0.164	0.152	
YASS 3	0.187	0.117	0.086	0.111	
YASS 4	0.118	0.098	0.112	0.130	
YASS 5	0.159	0.054	0.069	0.094	
YASS 6	0.031	0.270	0.260	0.145	
YASS 7	0.078	0.237	0.222	0.133	
YASS 8	0.138	0.232	0.180	0.121	
YASS 9	0.237	0.068	0.046	0.093	
YASS 10	0.159	0.202	0.141	0.119	
YASS 11	0.114	0.186	0.159	0.178	
YASS 12	0.077	0.179	0.194	0.179	

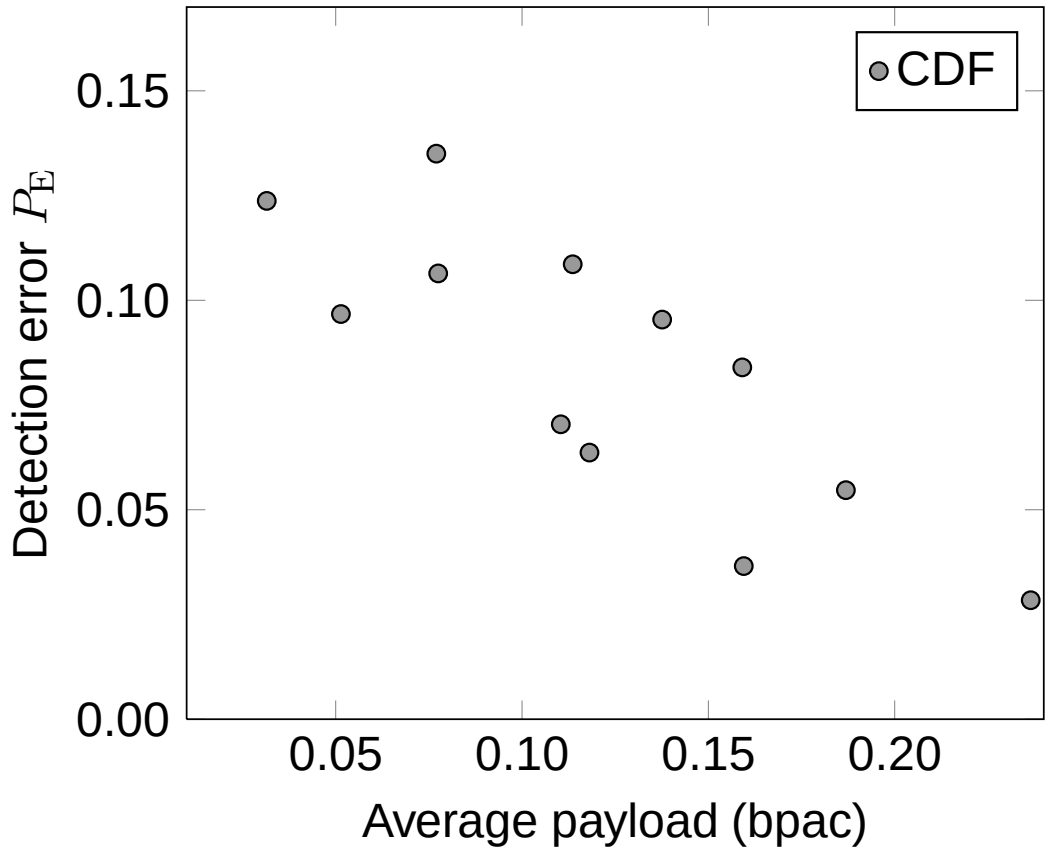
Experimental Results

		MP	CC-PEV	SPAM	CDF
Algorithm	bpac	(486)	(548)	(686)	(1,234)
YASS 1	0.110	0.110	0.123	0.140	0.070
YASS 2	0.051	0.155	0.164	0.152	0.097
YASS 3	0.187	0.117	0.086	0.111	0.055
YASS 4	0.118	0.098	0.112	0.130	0.064
YASS 5	0.159	0.054	0.069	0.094	0.037
YASS 6	0.031	0.270	0.260	0.145	0.124
YASS 7	0.078	0.237	0.222	0.133	0.106
YASS 8	0.138	0.232	0.180	0.121	0.095
YASS 9	0.237	0.068	0.046	0.093	0.028
YASS 10	0.159	0.202	0.141	0.119	0.084
YASS 11	0.114	0.186	0.159	0.178	0.109
YASS 12	0.077	0.179	0.194	0.179	0.135

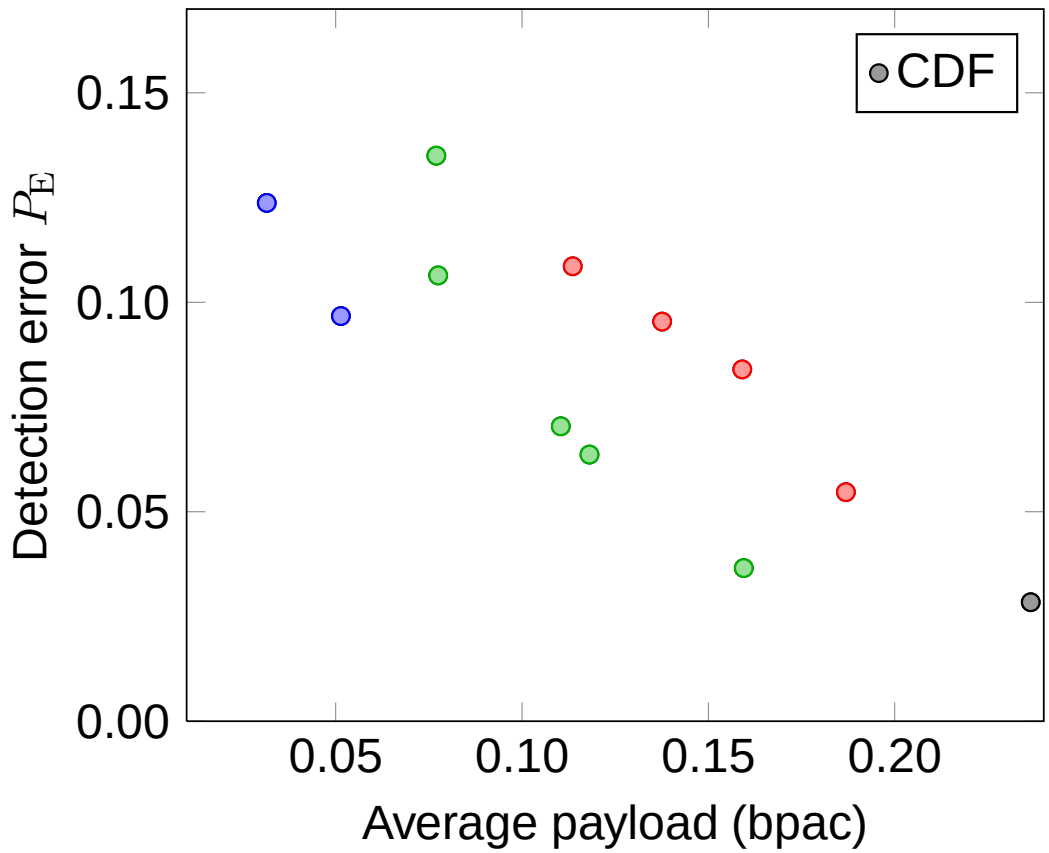
← feature set

← dimension

Experimental Results, cont'd



Experimental Results, cont'd



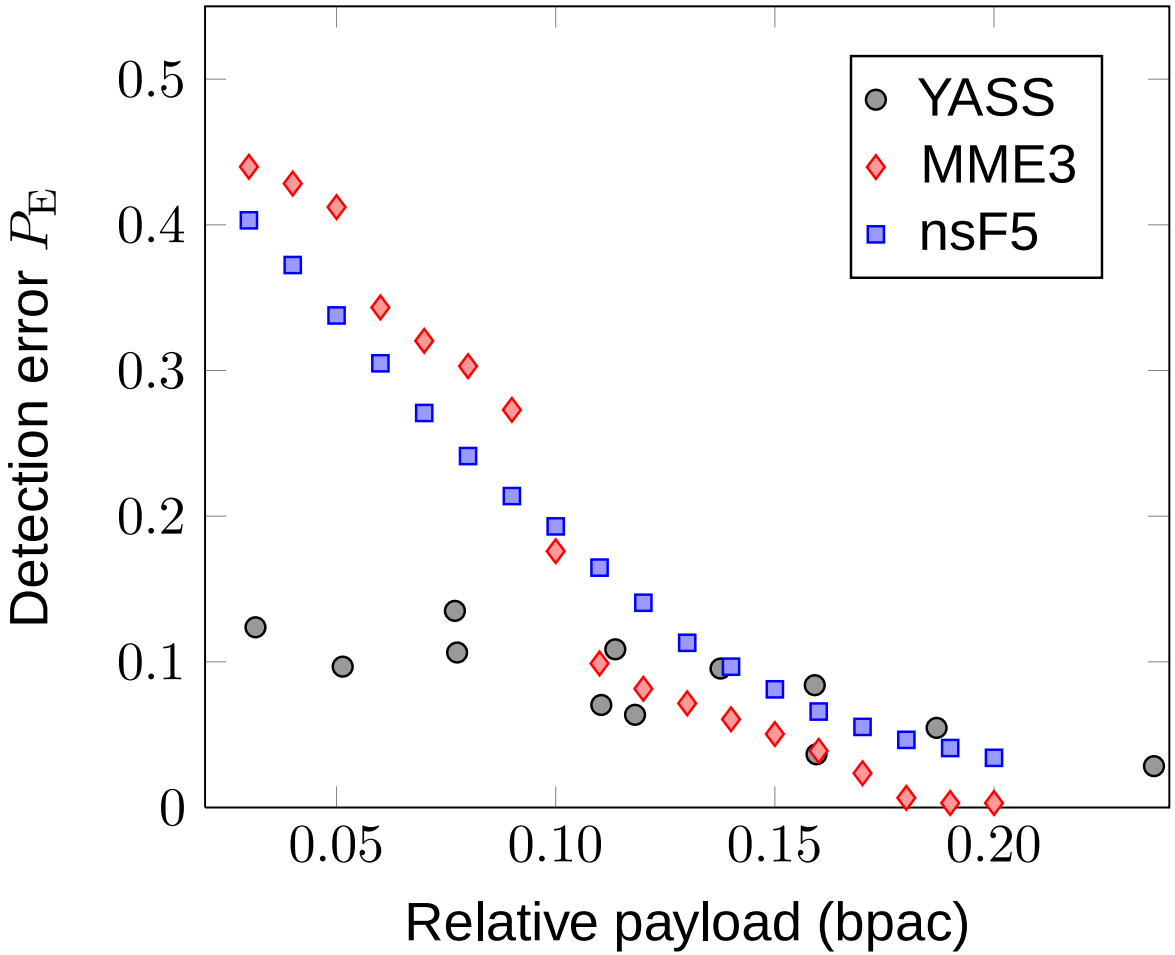
Original YASS

Extension 1 – repetition

Extension 2 – mixture of QF_h

Both Extensions

Comparison to Other Methods



Conclusions

- Modern steganalysis can detect YASS reliably

  $P_E < 15\%$ even for small payloads

- No implementation weakness employed

  $\Rightarrow$ detectability of further modifications

- Minimization of embedding impact seems like more secure steganographic strategy

[jan.kodovsky@binghamton.edu]