Fisher Information Determines Capacity of ε-secure Steganography

Tomáš Filler and Jessica Fridrich

Dept. of Electrical and Computer Engineering
SUNY Binghamton, New York

11th Information Hiding, Darmstadt, Germany, 2009

BINGHAMTON UNIVERSITY
State University of New York

Perfect vs. Imperfect Steganography

Stegosystems can be divided into two classes:

Perfectly secure stegosyst.

KL diverg. cover & stego
$D_{KL}(P||Q) = 0$

Detector does NOT exist.

Secure capacity is linear.

Communication rate is POSITIVE.

Imperfect stegosyst.

$D_{KL}(P||Q) = \varepsilon$

Detector DOES exist.

Sec. capacity is SUBlinear.

Communication rate is ZERO.

- perfectly secure stegosystems exist for artificial cover sources
- all known stegosystems for digital media are imperfect

Secure Capacity of ε-secure Stegosystems

From Taylor expansion ... ($n\beta = \#$ of changes)

$$D_{KL}\left(P^{(n)} || Q_\beta^{(n)}\right) = \frac{1}{2} n\beta^2 \mathtt{I} + O(\beta^3) = \varepsilon$$

$$\frac{1}{2} n\beta^2 \mathtt{I} \approx \varepsilon \quad \Rightarrow \quad n\beta \approx \sqrt{\frac{1}{\mathtt{I}}}\sqrt{2\varepsilon n}$$

$\mathtt{I}$ **... Fisher information (rate) at $\beta = 0$.**

Capacity of ε-secure stegosystems scales as $r\sqrt{n}$.

Root rate: (more refined measure of capacity)

$$r \approx \frac{1}{\sqrt{\mathtt{I}}}$$

Is Root Rate Useful?

Root rate: (more refined measure of capacity)

$$r \approx \frac{1}{\sqrt{\mathrm{I}}}$$

Fisher information rate can be expressed in a closed-form.

Applications:

- BENCHMARKING - compare stegosystems by their root rates.
- STEGANOGRAPHY DESIGN - maximize root rate w.r.t. embedding operation for fixed cover source.

Presentation Outline

1. ASSUMPTIONS
   model of cover objects, model of embedding impact
2. STEGANOGRAPHIC CAPACITY & ROOT RATE
   formal connection between detectability and root rate
3. APPLICATION
   comparison of LSB and ± 1 embedding in spatial domain
4. CONCLUSION

(1)

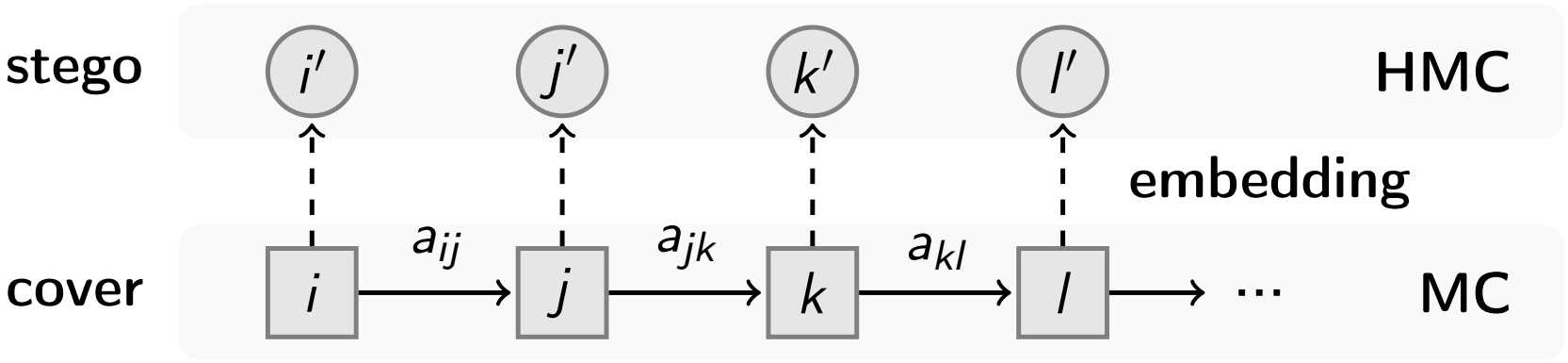
ASSUMPTIONS

Assumptions Summary

1. COVER SOURCE - Markov Chain
   first order Markov Chain with tran. prob. mat. $\mathbb{A} = (a_{ij})$

2. EMBEDDING ALGORITHM - MI embedding
   independent substitution of states (next slide)

3. STEGOSYSTEM - ε-secure (imperfect)
   stegosystem is ε-SECURE



Mutually Independent Embedding Operation

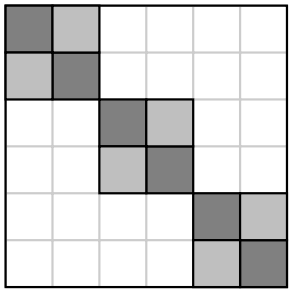
Impact of embedding is modeled as probabilistic mapping acting on each cover element independently - MI embedding.

$$Pr(Y_l = j | X_l = i) = b_{ij}(\beta)$$

X_l **... l-th cover element**
Y_l **... l-th stego element**
β **... change rate (rel. payload)**

Matrix $\mathbb{B} = (b_{ij})$ is "transition probability matrix" for $\beta \geq 0$.

LSB embedding:

$\mathbb{B} =$ 

$= 1 - \beta$ $= \beta$

Mutually Independent Embedding Operation

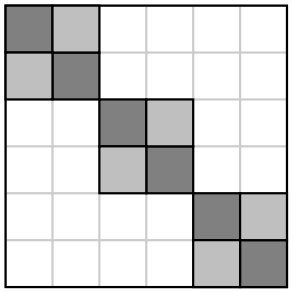
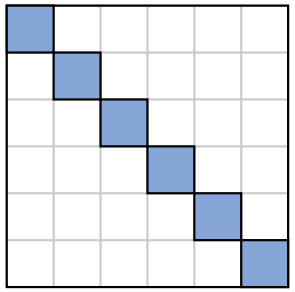
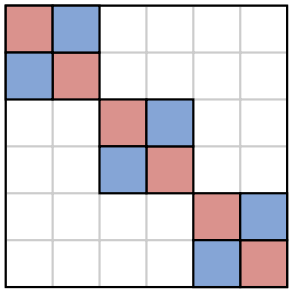
Impact of embedding is modeled as probabilistic mapping acting on each cover element independently - MI embedding.

$$Pr(Y_l = j | X_l = i) = b_{ij}(\beta)$$

X_l **... l-th cover element**
Y_l **... l-th stego element**
β **... change rate (rel. payload)**

Matrix $\mathbb{B} = (b_{ij})$ is "transition probability matrix" for $\beta \geq 0$.

LSB embedding:

$\mathbb{B} =$  $= \mathbb{I} + \beta\mathbb{C} =$  $+ \beta \cdot$ 

$= 1 - \beta$ $= \beta$ $= 1$ $= -1$

Mutually Independent Embedding Operation

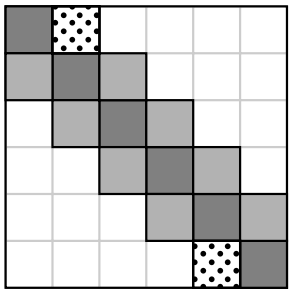
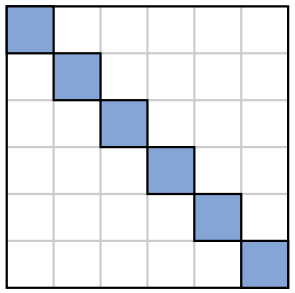
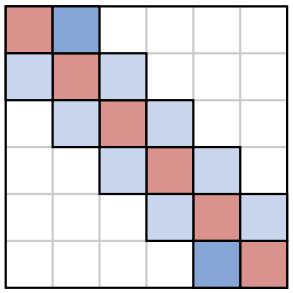
Impact of embedding is modeled as probabilistic mapping acting on each cover element independently - MI embedding.

$$Pr(Y_l = j | X_l = i) = b_{ij}(\beta)$$

X_l **... l-th cover element**
Y_l **... l-th stego element**
β **... change rate (rel. payload)**

Matrix $\mathbb{B} = (b_{ij})$ is "transition probability matrix" for $\beta \geq 0$.

± 1 embedding:

$\mathbb{B} =$  $= \mathbb{I} + \beta \mathbb{C} =$  $+ \beta \cdot$ 

$= 1 - \beta$ $= \frac{\beta}{2}$ $= \beta$ $= 1$ $= 0.5$ $= -1$

Mutually Independent Embedding Operation

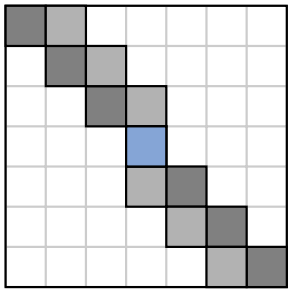
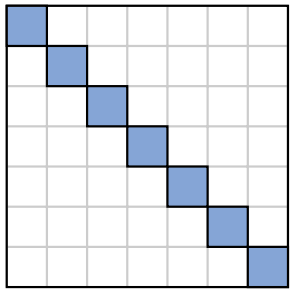
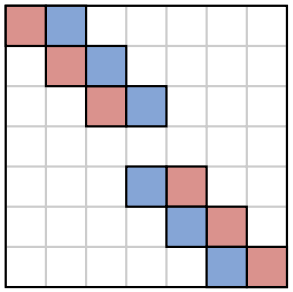
Impact of embedding is modeled as probabilistic mapping acting on each cover element independently - MI embedding.

$$Pr(Y_l = j | X_l = i) = b_{ij}(\beta)$$

X_l **... l-th cover element**
Y_l **... l-th stego element**
β **... change rate (rel. payload)**

Matrix $\mathbb{B} = (b_{ij})$ is "transition probability matrix" for $\beta \geq 0$.

F5 embedding:

$\mathbb{B} =$  $= \mathbb{I} + \beta\mathbb{C} =$  $+ \beta \cdot$ 

$= 1 - \beta$ $= \beta$ $= 1$ $= -1$

Mutually Independent Embedding Operation

Impact of embedding is modeled as probabilistic mapping acting on each cover element independently - MI embedding.

$$Pr(Y_l = j | X_l = i) = b_{ij}(\beta)$$

X_l **... l-th cover element**
Y_l **... l-th stego element**
β **... change rate (rel. payload)**

Matrix $\mathbb{B} = (b_{ij})$ is "transition probability matrix" for $\beta \geq 0$.

Assumption:

$$\mathbb{B}(\beta) = \mathbb{I} + \beta \mathbb{C}$$

$\mathbb{C}$ describes the inner workings of the embedding algorithm.

(2)

STEGANOGRAPHIC CAPACITY & $\sqrt{RATE}$

Square Root Law

1. If $\frac{n\beta_n}{\sqrt{n}} \to 0$ then the stegosyst. are asymptotically secure.
2. If $\frac{n\beta_n}{\sqrt{n}} \to +\infty$ then arbitrarily accurate stego detectors exist.

For fixed level of security,

$$\frac{n\beta_n}{\sqrt{n}} < C \quad \Rightarrow \quad \beta_n \to 0 \text{ as } n \to \infty.$$

[Filler, Ker, Fridrich, "The Square Root Law of Steganographic Capacity for Markov Covers", Proc. SPIE, 2009]

The Best Possible Steganalyzer

Hypothesis test:

$$\begin{array}{lcll} \mathrm{H}_0 & : & \beta = 0 & \text{decide cover} \\ \mathrm{H}_1 & : & \beta > 0 \text{ (known)} & \text{decide stego.} \end{array}$$

Detection statistics (log-likelihood ratio):

$$L_\beta^{(n)}(X) = \frac{1}{\sqrt{n}} \ln \frac{Q_\beta^{(n)}(X)}{P^{(n)}(X)}$$

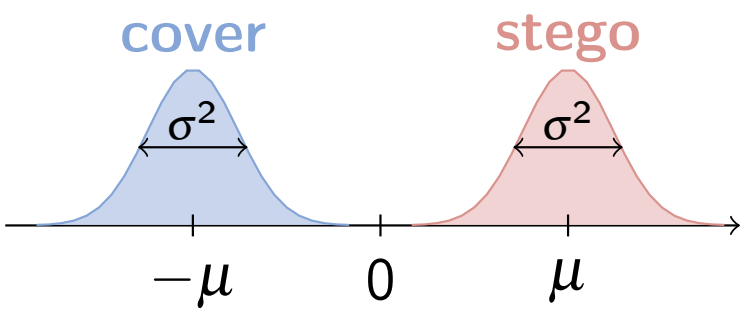
stego distribution with n elements

cover distribution with n elements

change rate

The Root Rate

For small β, $L_\beta^{(n)}(X)$ is Gaussian (Local Asymptotic Normality).



$$\mu = \sqrt{n}\beta^2 \mathtt{I}/2$$
$$\sigma^2 = \beta^2 \mathtt{I}$$

Fisher information rate $\mathtt{I} = \lim_{n\to\infty} \frac{1}{n}\frac{d^2}{d\beta^2} D_{KL}\left(P^{(n)} || Q_\beta^{(n)}\right)\Big|_{\beta=0}$

Deflection coefficient: $d^2 = (-\mu-\mu)^2/\sigma^2 = n\beta^2\mathtt{I} < \varepsilon$

$$n\beta < r\sqrt{\varepsilon n} \qquad r = \sqrt{\frac{1}{\mathtt{I}}} \text{ ROOT RATE}$$

Closed Form for Fisher Information Rate

Theorem (Fisher Information Rate)

Let $\mathbb{A} = (a_{ij})$ define the MC cover model and $\mathbb{B} = \mathbb{I} + \beta\mathbb{C}$ represent MI embedding. Then, the Fisher information rate $\mathtt{I}$ exists and can be written as

$$\mathtt{I} = \mathbf{c}^T \mathbb{F} \mathbf{c}.$$

- $\mathbf{c} = (c_{11}, \ldots, c_{NN})^T$ *is column vector obtained directly from* $\mathbb{C}$.
- *Matrix* $\mathbb{F} \in \mathscr{R}^{N^2 \times N^2}$, $\mathbb{F} = f(\mathbb{A})$ *and does not depend on* $\mathbb{B}$.

Maximizing root rate $\Rightarrow$ minimizing $\mathtt{I}$ w.r.t. $\mathbb{C}$.

Closed form for $\mathtt{I}$ enables us to

- **compare stegosystems**
- **maximize capacity w.r.t. embedding operation $\mathbb{C}$.**

(3)

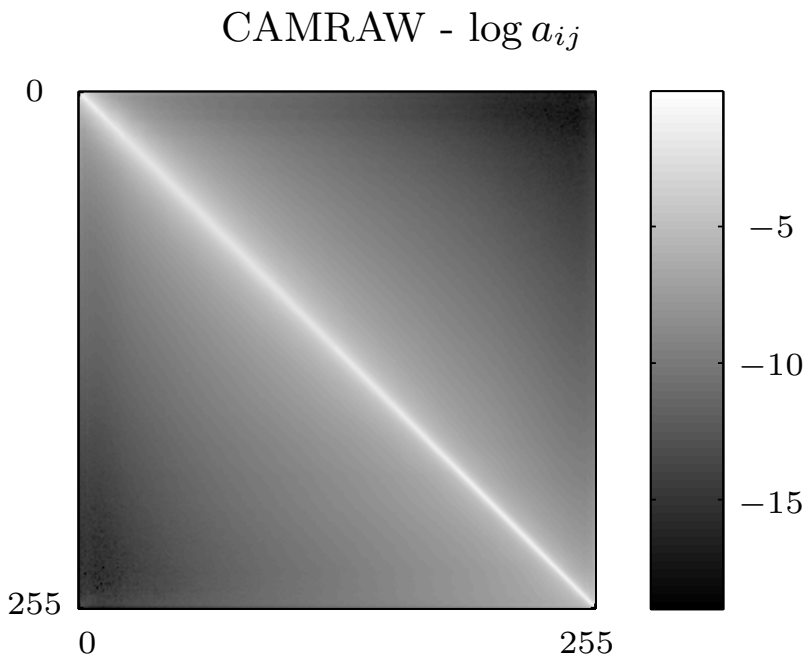
APPLICATIONS

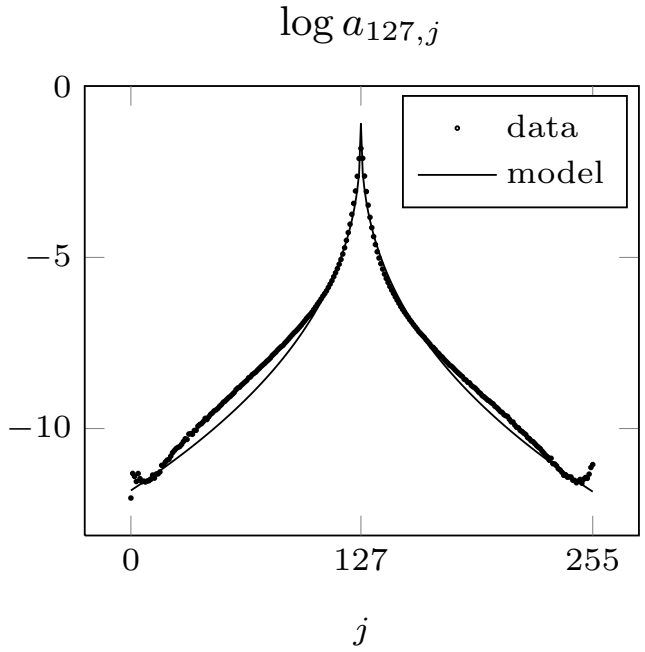
Markov Model of Spatial Domain Images

Image databases: CAMRAW, NRCS, NRCS-JPEG70.

Spatial domain 8-bit grayscale images: $\mathbb{A} = (a_{ij}) \in \mathscr{R}^{256\times 256}$

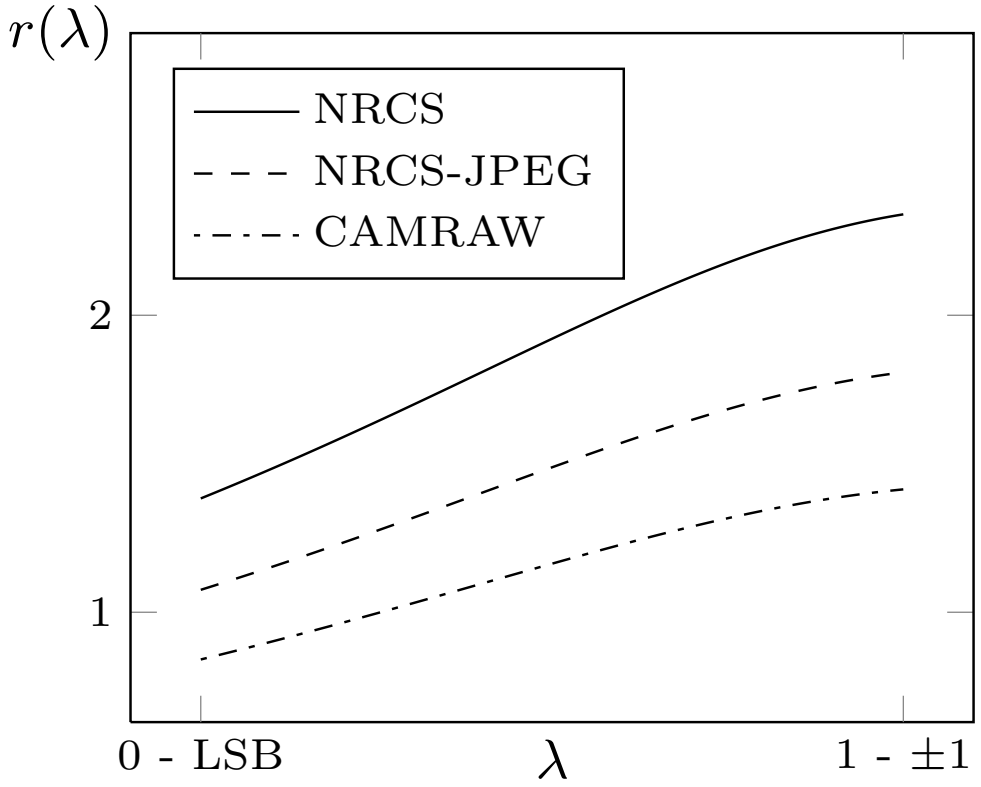
$$a_{ij} = \frac{1}{Z_i} e^{-(|i-j|/\tau)^\gamma}$$





Convex Combination of LSB and ± 1 Embedding

Use ± 1 embedding in first λn pixels and LSB embedding in the rest.



Root rate $r(\lambda)$:

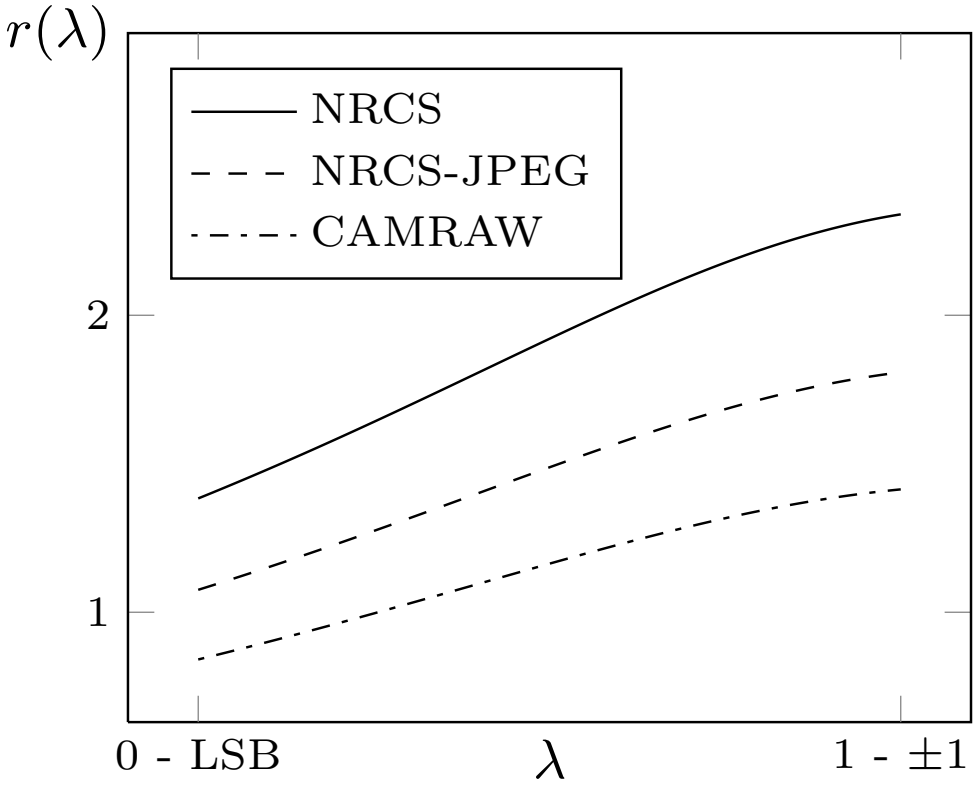
$$\mathbf{c}_\lambda = \lambda \mathbf{c}_{\pm 1} + (1-\lambda)\mathbf{c}_{\mathrm{LSB}}$$

$$r(\lambda) = \sqrt{\frac{1}{\mathbf{c}_\lambda^T \mathbb{F} \mathbf{c}_\lambda}}$$

Higher root rate = better method.

Convex Combination of LSB and ± 1 Embedding

Use ± 1 embedding in first λn pixels and LSB embedding in the rest.

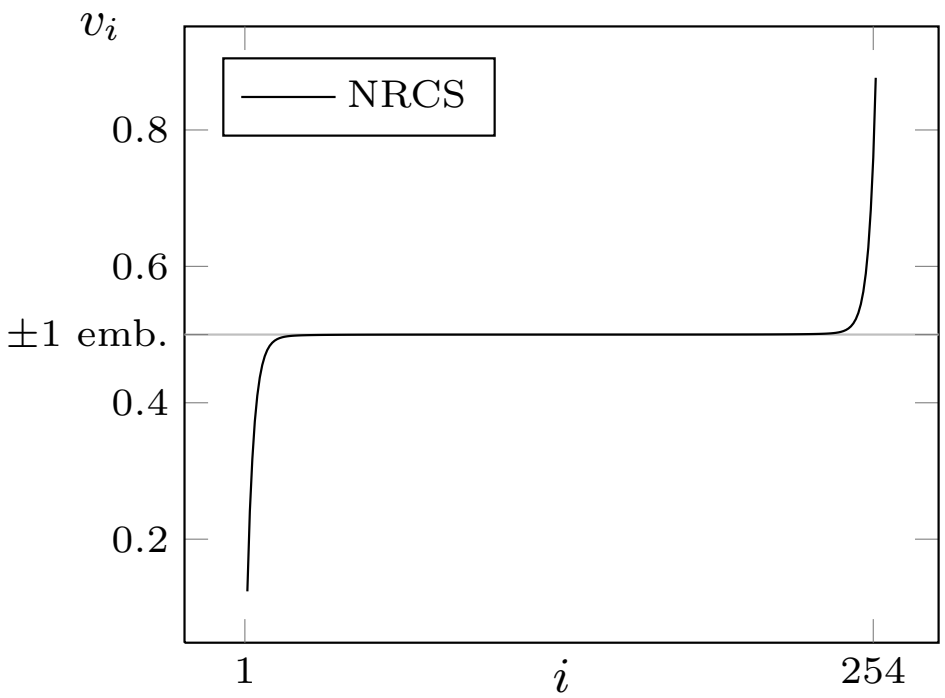


Pevný, Bass, Fridrich: Steganalysis by Subtractive Pixel Adjacency Matrix, submitted to ACM MM&SEC 2009, Princeton, NJ

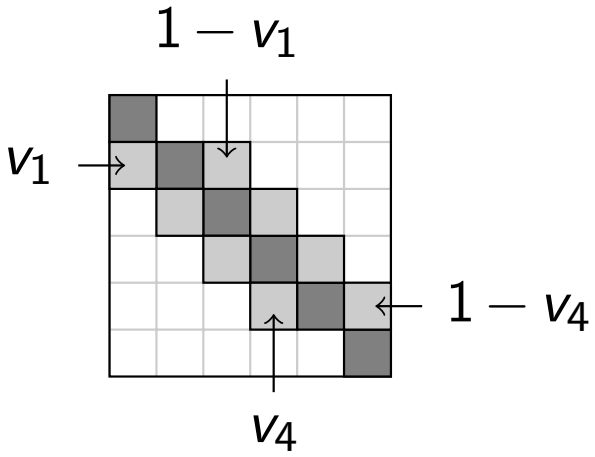
Higher root rate = better method.

± 1 is Asymptotically Optimal

What embedding method minimizes Fisher information rate and modifies cover by at most 1?



Class of embedding operations:



± 1 embedding is asymptotically optimal as number of grayscales $N \to \infty$.

Conclusion and Future Directions

Steganographic capacity of ε-SECURE stegosystems $\approx r\sqrt{n}$.

We coin a new term for constant r - the Root Rate.

Root rate

- is determined by Fisher information rate.
- can be expressed in a closed-form amenable to optimization (under mentioned assumptions).
- was used to compare spatial domain stegosystems.

Future work: use this framework for JPEG images.

Conclusion and Future Directions

Steganographic capacity of ε-SECURE stegosystems $\approx r\sqrt{n}$.

We coin a new term for constant r - the Root Rate.

Root rate

- is determined by Fisher information rate.
- can be expressed in a closed-form amenable to optimization (under mentioned assumptions).
- was used to compare spatial domain stegosystems.

Future work: use this framework for JPEG images.

Thank you!

tomas.filler@binghamton.edu