Fast and Efficient Steganalysis Methods for Spread Spectrum Steganography

KAI-SHENG SONG

Deparment of Statistics

Florida State University

Tallahassee, FL 32306

Steganography

- Steganography: “covered writing” in Greek
- Art and Science of Covert Communication: Convey message under cover
- Cryptography: Scramble the message to prevent eavesdroppers
- Digital Watermarking: A special case of information hiding
- Key Difference: Active adversary would attempt to remove invalidate or forge watermarks

Ancient Steganography

- Ancient Steganography: Histaiaeus (Greek) shaved the head of his messenger, wrote the message on his scalp, and then waited for the hair to regrow; the messenger could travel freely and upon arrival at his destination, he shaped his head and pointed his head at the recipient.
- Invisible Ink: Pliny the Elder described how the milk of the thithymallus plant dried to transparency when applied to paper but darkened to brown when heated.

Ancient Steganography

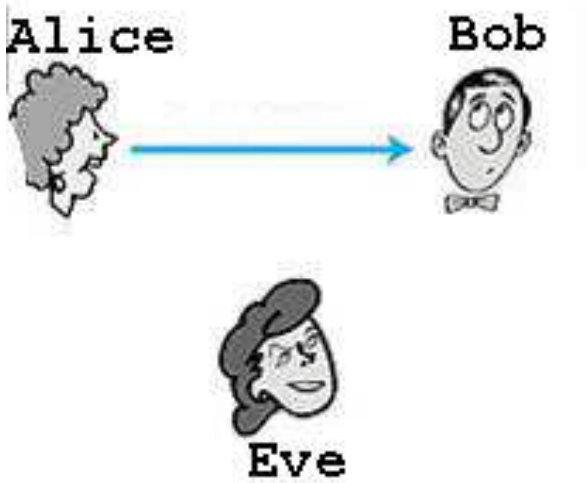
- Ancient Chinese: they wrote notes on small pieces of silk, wadded them into little balls, and coated them in wax, and then to be swallowed by a messenger and retrieved at his/her gastrointestinal convenience.
- Chinese Moon Cake: Hiding message in a cake in a simila way as fortunate cookies

Renaissance Steganography

- Johannes Trithemius (1404-1472) invented a steganographic cipher: each letter was represented as a word taken from a succession of columns. The resulting series of words would be a legitimate prayer.
- Giovanni Battista Porta (1535-1615): how to conceal a message within a hard-boiled egg ?
- Writing the message on the shell with a special ink made from an ounce of alum and a pint of vinegar.
- The solution penetrates the porous shell leaving no visible trace
- The message is stained on the surface of the hardened egg albumen and can be read when the shell is removed.

Modern Steganography

- The first informal definition of a steganographic scheme was formulated by Simmons (1983) as the prisoners' problem.
- Alice and Bob: inmates who wish to communicate in order to hatch an escape plan
- Eve: the warden examines all communication between them



LSB Embeding

- Least Significant Bit (LSB): Message bits are encoded as the LSBs of, say, pixel values
- To embed a 0 at a pixel with binary value 00110011, the last bit being the LSB is replaced by the message bit: 00110010
- Applicable to Any Numerical Samples: audio samples, quantized DCT coefficients
- OutGuess, F5 (decrementing the absolute value of DCT coefficient by 1)
- Weakness: Histogram attack

Spread Spectrum Steganography

Additive Embedding Scheme

$$Y_i = X_i + \gamma W_i \quad i = 1, \cdots, N$$

$X = \{X_i\}_{i=1}^{N}$ —- a sequence of the original data from cover Works:

- Luminance values of a picture,
- Sound intensity in an audio frame
- Coefficients of transformations such as DCT, DFT, DWT

- $W = \{W_i\}_{i=1}^{N}$: a pseudo-random sequence generated from a pseudo-random number generator (PRNG) initialized by a secret stego key.
- γ: Embedding strength parameter (gain factor)
- $Y = \{Y_i\}_{i=1}^{N}$: a sequence of possibly altered data
- Additive embedding may not be appropriate, *e.g.*, when the host signals X_i vary widely

Stochastic Modulation

Multiplicative Embedding Scheme:

$$Y_i = X_i(1 + \gamma W_i) \quad i = 1, \cdots, N$$

- Multiplicative embedding: robust against scale differences in X.
- Provide a way of perceptual masking of the hidden message in an image (Weber's Law)
- Consider multiplicative embedding in the transformed domains such as DCT, DFT, and DWT of an image

Embedding Rule

To embed M bits of information, consider partition $\{\mathbf{S}_i\}_{i=1}^{M}$

$$\mathbf{S}_i \cap \mathbf{S}_j = \emptyset \qquad \text{for} \qquad i \neq j$$

$$\cup_{k=1}^{M} \mathbf{S}_k = \{1, \cdots, N\}$$

Embedding Rule

$$Y_i = X_i(1 + \gamma_k W_i), \quad i = 1, \cdots, n_k,$$

$\gamma_k := \gamma b_k$

$n_k := card(\mathbf{S}_k)$

$b_k = +1$ for bit 1 and $b_k = -1$ for bit 0

Spreading sequence modulated by the message bit b_k

Efficient Scores Method:

- Total Efficient Score Vector:

$$\mathbf{U}(\lambda_k) := \frac{\partial \mathcal{L}(\lambda_k; \mathbf{Y}_k)}{\partial \lambda_k},$$

- Fisher Information Matrix:

$$I(\lambda_k) := E\left\{ -\frac{\partial^2 \mathcal{L}(\lambda_k; \mathbf{Y}_k)}{\partial \lambda_k^2} \right\}.$$

Blind Detection of Hidden Message

- $\{X_i\}_{i=1}^N$: not available for detecting the existence of hidden message
- Parameter vector δ_k: typically unknown at the receiver end

Hypothesis Testing

$$H_0 : \gamma_k = 0 \quad \text{versus} \quad H_A : \gamma_k \neq 0$$

- Treat unknown δ_k as a nuisance parameter vector
- Parameter Space under H_0: $\Lambda_0 = \{\lambda_k : \gamma_k = 0\}$
- Parameter Space under H_A: $\Lambda_A = \{\lambda_k : \gamma_k \neq 0\}$

Efficient Scores Statistic:

$$\mathcal{S}(0,\hat{\delta}_k) = B_k^T I^{\gamma_k\gamma_k}(0,\hat{\delta}_k) B_k$$

$$B_k := \mathbf{U}_{\gamma_k}(0,\hat{\delta}_k) - \mathbf{R}(0,\hat{\delta}_k)\mathbf{U}_{\delta_k}(0,\hat{\delta}_k)$$

$$\mathbf{R}(0,\hat{\delta}_k) := I_{\gamma_k\delta_k}(0,\hat{\delta}_k) I^{-1}_{\delta_k\delta_k}(0,\hat{\delta}_k)$$

$\mathcal{S}(0, \hat{\delta}_k)$: chi-squared $df := dim(\Lambda_A) - dim(\Lambda_0)$

Under H_0: central chi-squared

If $\mathcal{S}(0, \hat{\delta}_k) > \chi^2_{df,1-\alpha}$, declare hidden message present

Once we reject H_0, the information bit b_k is determined by

$$\hat{b}_k = \text{sign}(B_k)$$

Applications

Generalized Gaussian (GG) Model

- Model Coefficients $\{X_i\}_{i\in S_k}$ of Transforms (DWT, DCT, etc) Statistically by GG distributions:

$$f_X(x;\delta_k) = \frac{\theta_k}{2\sigma_k\Gamma(\frac{1}{\theta_k})} e^{-|\frac{x}{\sigma_k}|^{\theta_k}}, \quad x \in R$$

$\sigma_k > 0$ —- scale parameter

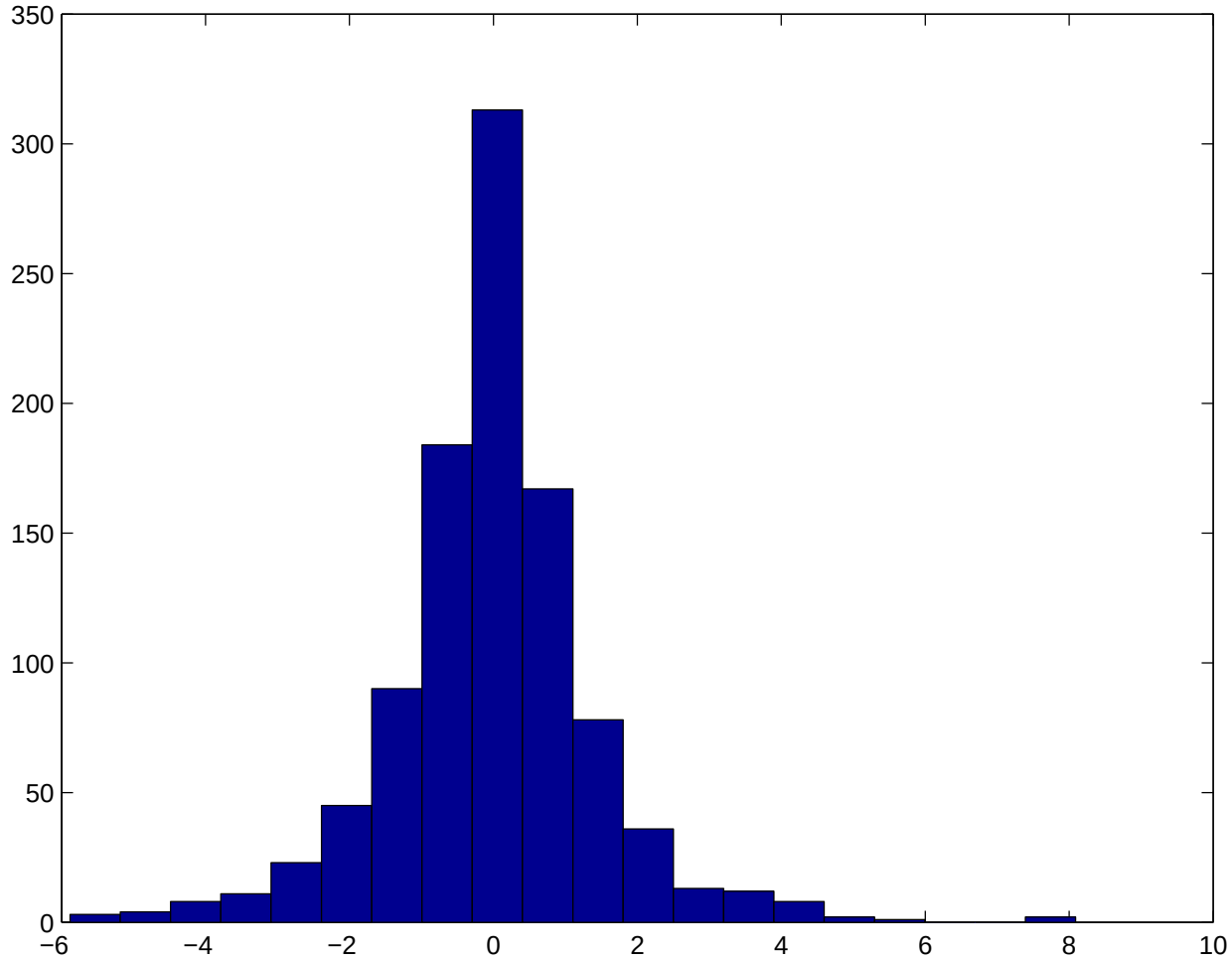
$\theta_k > 0$ —- shape parameter

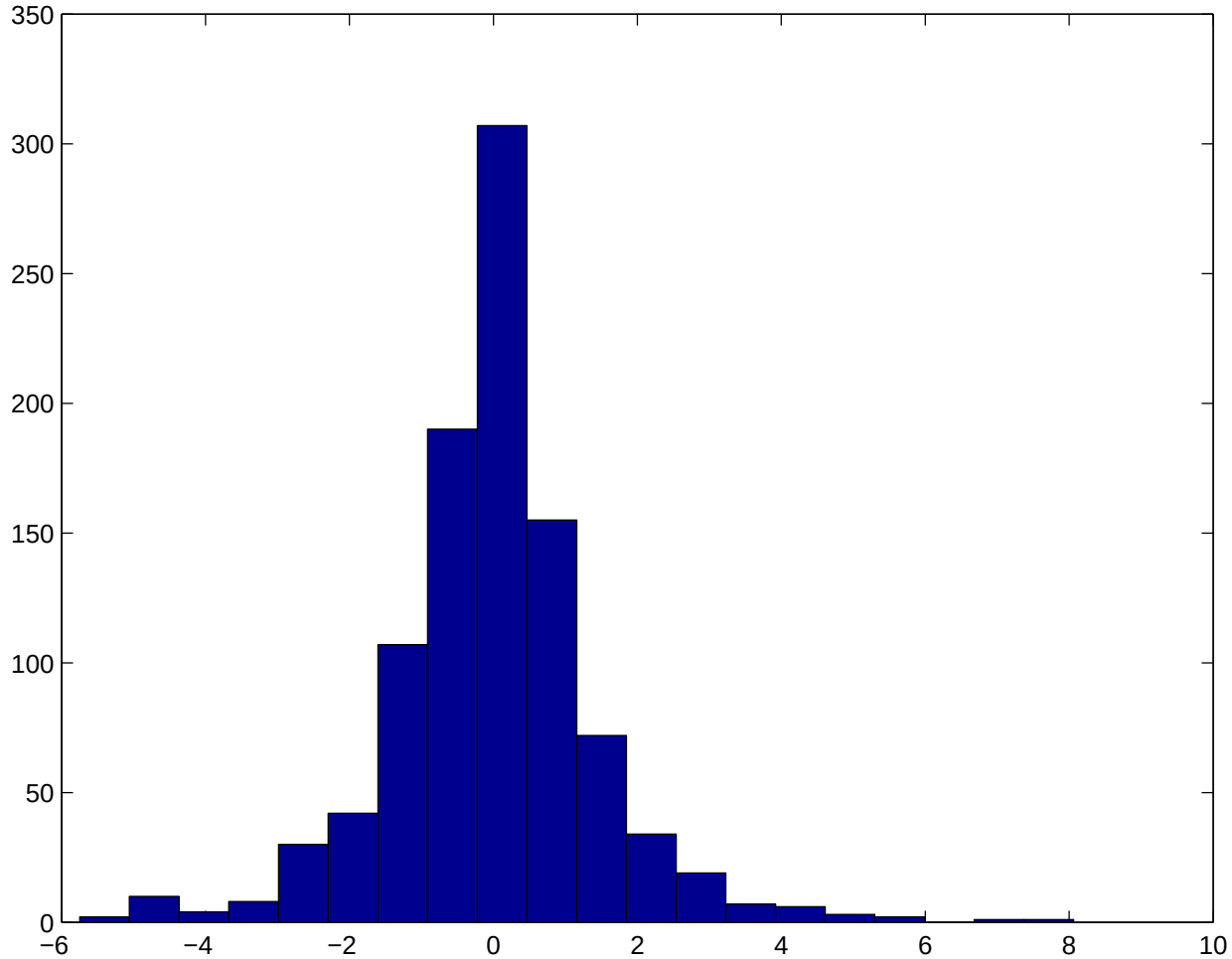
$\Gamma(p) = \int_0^\infty x^{p-1}e^{-x}dx$ —- Gamma function

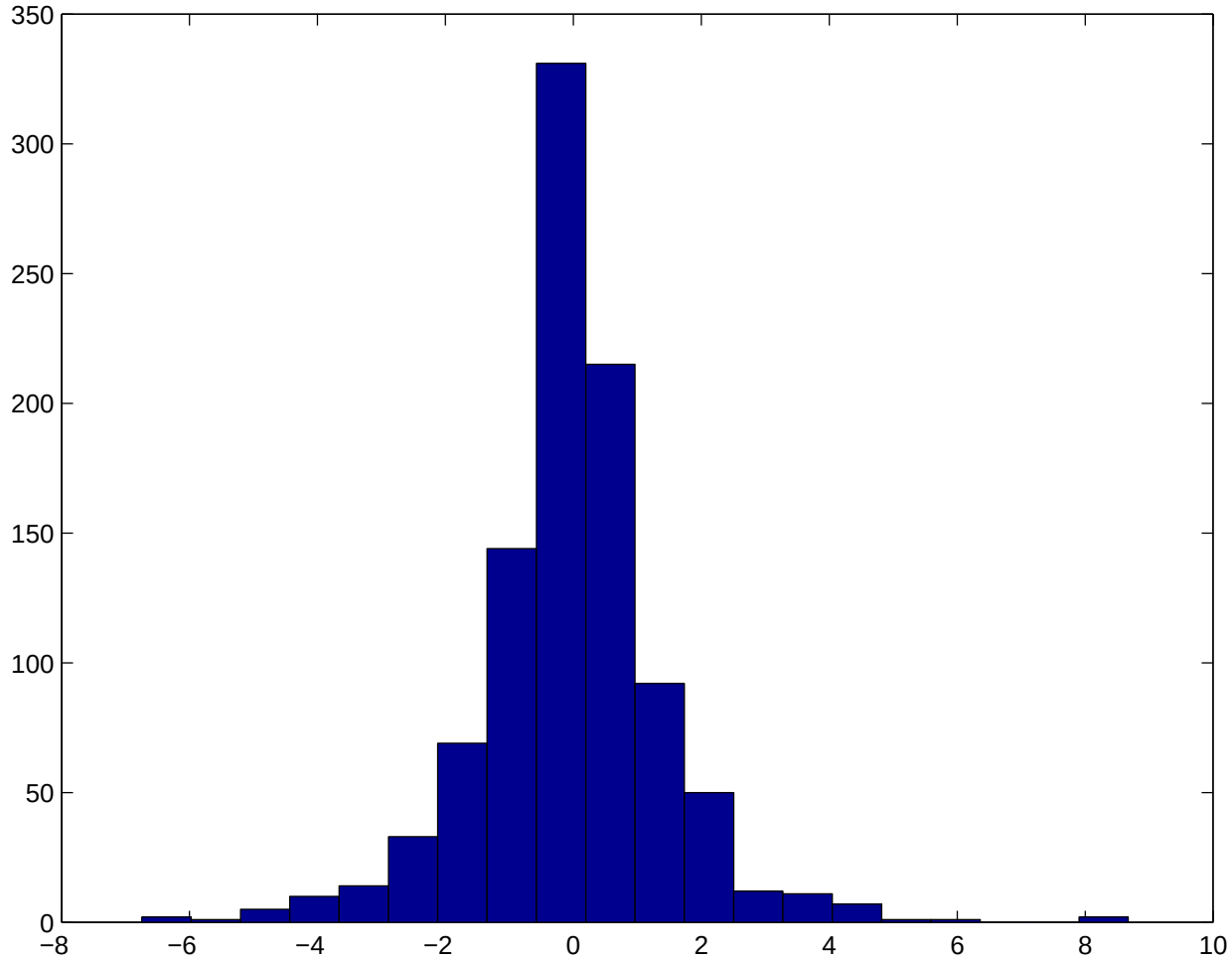
Applications

Generalized Gaussian (GG) Model

- Flexible parametric family: Laplace, Gaussian and uniform distributions.
- Parameter vector (σ_k, θ_k): a vector of nuisance parameters for testing the composite null hypothesis $H_0 : \gamma_k = 0$ versus the composite alternative $H_A : \gamma_k \neq 0$


350
300
250
200
150
100
50
0
−6
−4
−2
0
2
4
6
8
10


350
300
250
200
150
100
50
0
−6
−4
−2
0
2
4
6
8
10


350
300
250
200
150
100
50
0
−8
−6
−4
−2
0
2
4
6
8
10

Applications

Generalized Gaussian (GG) Model

- Components of the vector Y are independent but not identically distributed
- By writing down the log likelihood of the output data vector Y and calculating the vector of scores, we obtain, after extensive algebra, the score test statistic:

Efficient Scores Statistic

$$\mathcal{S}(0, \hat{\delta}_k) = \frac{[\sum_{i \in S_k} W_i \left(\hat{\theta}_k |Y_i / \hat{\sigma}_k|^{\hat{\theta}_k} - 1 \right)]^2}{\hat{\theta}_k \sum_{i \in S_k} W_i^2}$$

- $(\hat{\sigma}_k, \hat{\theta}_k)$: MLE of (σ_k, θ_k) subject to $\gamma_k = 0$
- Any root n consistent estimator of (σ_k, θ_k) works.

Efficient Scores Detection

- Obtain a threshold value $\chi^2_{1,1-\alpha}$ from the limiting central chi-squared distribution with one degree of freedom.
- Decision Rule:

$$\text{If} \mathcal{S}(0, \hat{\delta}_k) > \chi^2_{1,1-\alpha} \Longrightarrow \text{hidden message present}$$

- Alternatively compute P-value:

$$\text{If P-value } < \alpha \Longrightarrow \text{hidden message present}$$

Message Bit Recovery

- The message bit b_k is recovered by

$$\hat{b}_k = \text{sign}\left(\sum_{i\in\mathbf{S}_k} W_i \left(\hat{\theta}_k \left|\frac{Y_i}{\hat{\sigma}_k}\right|^{\hat{\theta}_k} - 1\right)\right)$$

Cover Image

Stego Image