A Novel Audio Steganalysis Based on High-Order Statistics of a Distortion Measure with Hausdorff Distance

Yali(Tracy) Liu, Rennie Archibald, Biswanath Mukherjee, Dipak Ghosal
University of California, Davis

Ken Chiang, Cherita Corbett
Sandia National Laboratories, Livermore

Outline

- Background
- Related Work
- Methodology
- Experimental Results
- Conclusion

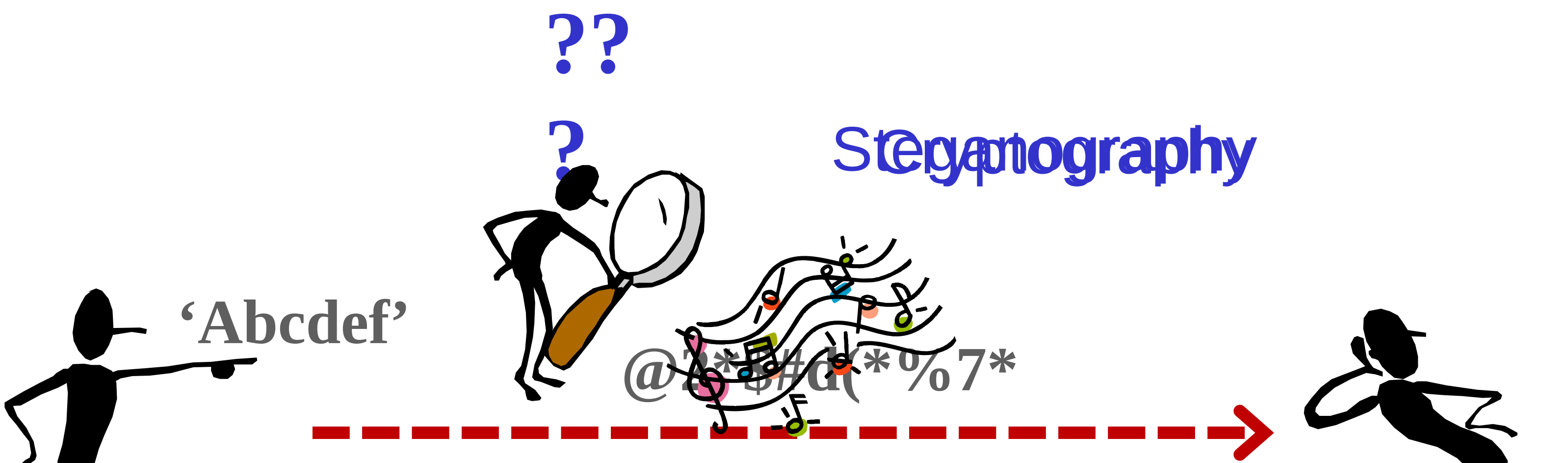
Introduction

- Steganography is having an enormous impact in our lives
  - Novel data hiding technology
  - Uncontrolled and unlimited information exchange
- News
  - In Feb. 2000, *USA Today* reported that terrorists are using steganography to hide their communication from law enforcements.
  - In Feb. 2001, *Wired News* reported that messages are being hidden in images posted to Internet auction sites like eBay or Amazon.
- Audio Steganography
  - Voice over Internet Protocol (VoIP) services
  - High hidden capacity of audio media
    - Inherent redundancy and transient and unpredictable characteristics

Steganography

The art of hiding information in ordinary-looking objects

- Goal
  - Conceal the existence of secret communication.



• **Storage carrier**: Images, audio, video, and natural languages
• **Non-storage carriers**: Networking protocols, timing information of the communication channel

Steganography

- Requirement
  - Media containing some hidden information (stego-objects) should be indistinguishable from media without any hidden information (cover-objects).

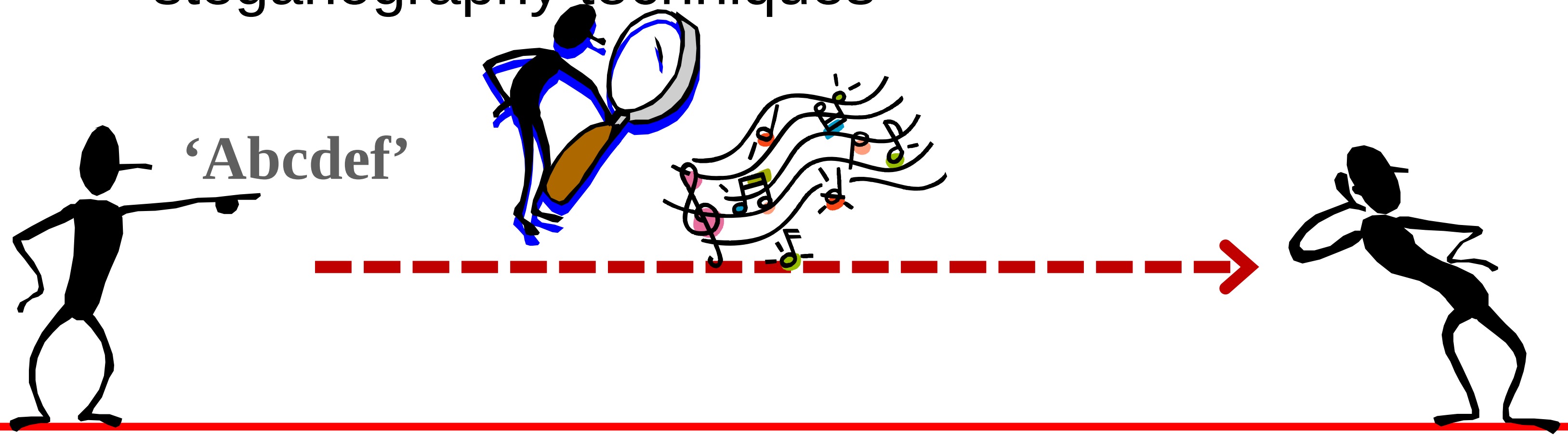
- Example: LSB Steganography
  - Replacing least-significant-bits (LSBs) of digital data with message bits.

1100100 0	Embedding ← Extracting	(0010110**1** (1010011**0** (1101001**0**	0001110**<u>1</u>** 1100010**<u>1</u>** 1010110**<u>0</u>**	1101110**0**) 0000110**0**) 01100011)

Steganalysis

The countermeasure to steganography

- Identify suspected information streams
- Detect of the presence of hidden content
- Message recovery
- Effective way for performance evaluation of steganography techniques



Steganalysis vs. Steganography

- Steganography
  - Stego-objects should be indistinguishable from cover-objects

Steganalysis vs. Steganography

- Take advantage of statistical (or perceptual) distinction of stego-signals from cover-signals
  - Cover-objects
    - Natural images or audios tend to be continuous and smooth
  - Hidden data
    - Independent to the cover media
  - Stego-objects
    - The continuity will be changed because it incurs random variation
- Design an efficient classifier that can distinguish between cover-objects and stego-objects

Existing Solutions

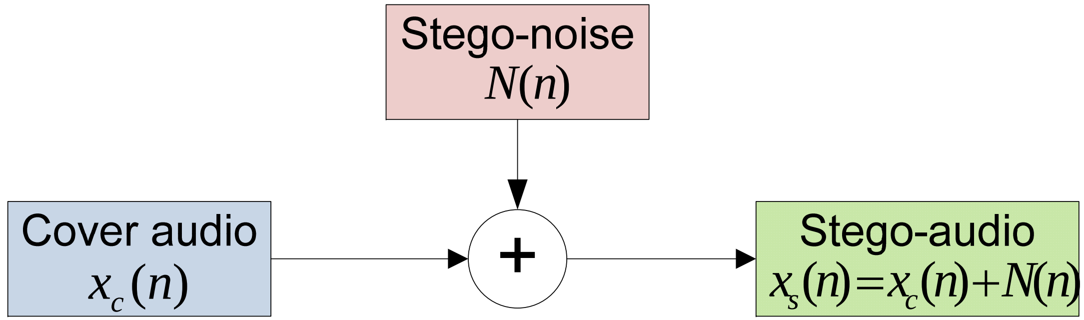
- High-Order Statistics and Steganalysis
  - A general steganalysis algorithm based on image high-order statistics [1]
  - The moments of the histogram characteristic function [2]
  - Statistical moments of the characteristic functions of the wavelet sub-bands [3]
- Distortion Measures and Steganalysis
  - The de-noised versions of the image signals can be used to represent close approximations of the cover-images [4]
  - The distortion of the cover-image to its de-noised version is different than the distortion between a stego-image and its de-noised version
  - Image quality metrics and distortion audio metrics are tested for their sensitivity to the presence of steganographic content [5]

Motivation

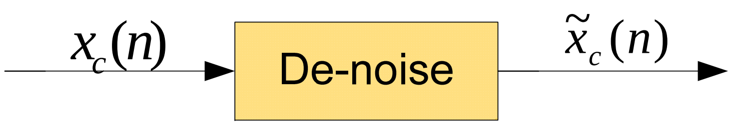
- From image to audio
  - The statistical regularities inherent to the spatial composition of images
- Quality to embedding
  - The primary motivation
    - Evaluate the perceptual and objective quality performance of images or audio
  - The limited distinguishing capability

Define a distortion measure that is designed specifically to detect modifications to audio content

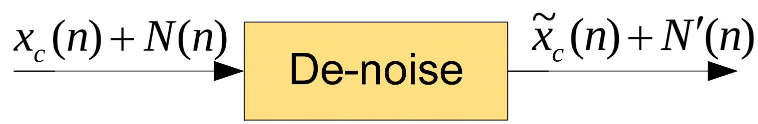
Steganography Model Descriptions



(a)



(b)

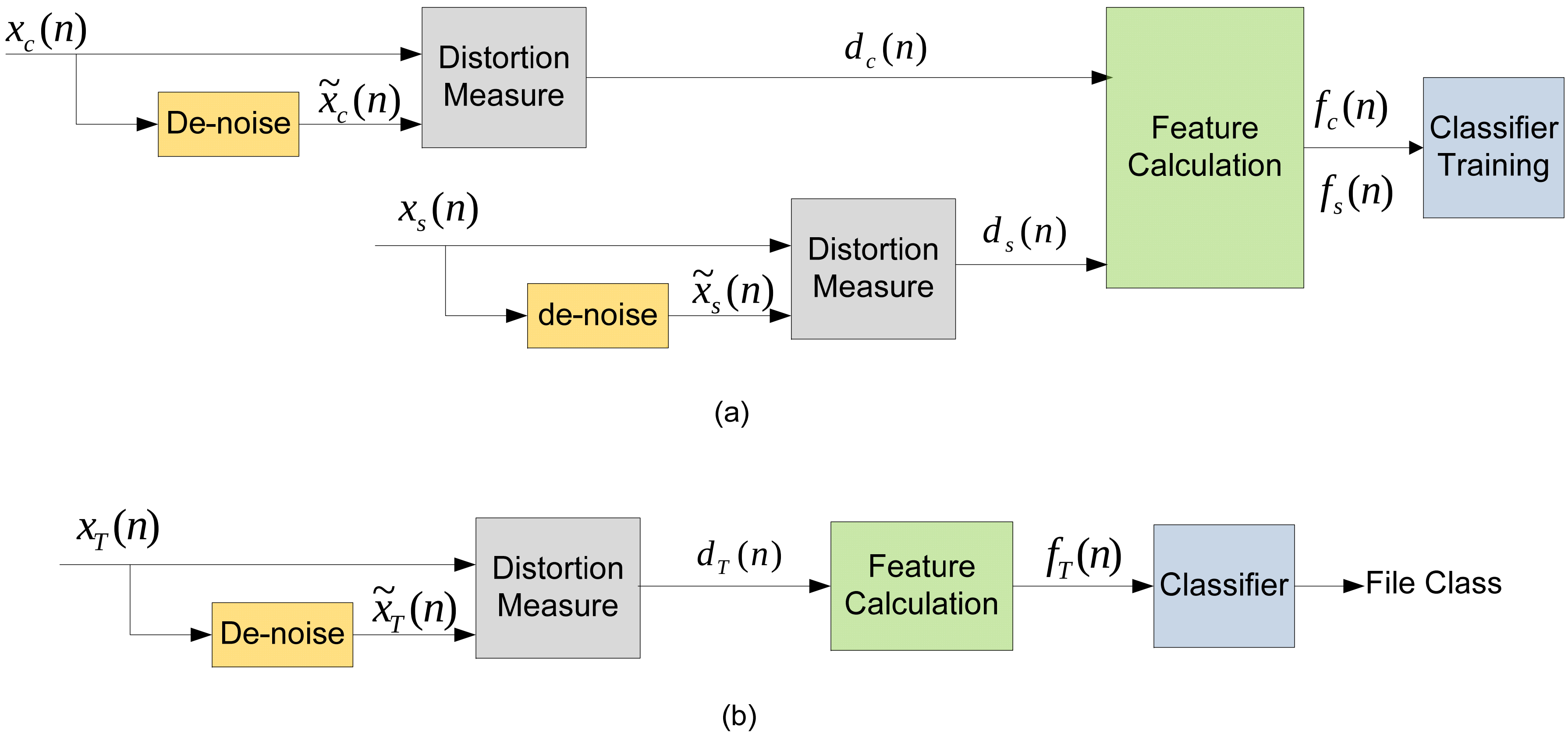


(c)

(a) additive noise steganography model,

(b) de-noising a cover audio object, and (c) de-noising a stego-audio object.

Schematic Description of Steganalysis

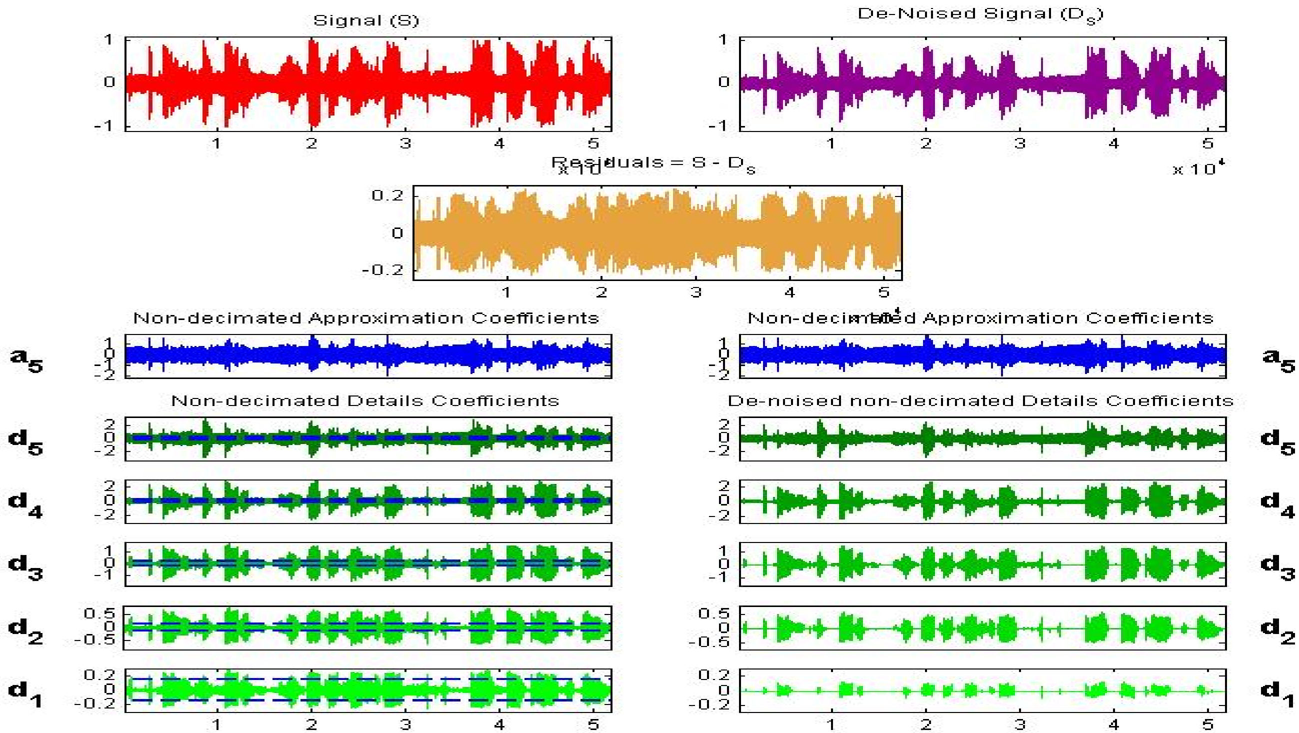


(a) training (b) testing

Wavelet Denosing

- Goal: recover the characteristics of the original cover-audio object while removing as much noise as possible.
  - Adaptive represent of signal discontinuities

$$X_s[n] = X_c[n] + N[n]$$



Distortion Measure

- Measure space - signatures instead of audio files
  - Wavelet coefficients at different levels of resolution
    - Multi-resolution decomposition
- Local distortion measure - pre-defined small segments
  - Partial modification of the cover-objects
  - Time-frequency localization from wavelet
- Measure metrics - Hausdroff Distance
  - Successful applications in object matching
  - Sensitive to modification
    - Max-min distance

Distortion Measure with Hausdorff distance

- Suppose the length of each segment of the audio file is *M*.
- Wavelet decomposition at level *p*
- For *m*th segment, the wavelet coefficients of the audio file and its de-noised version are $C_m^p = \{C_m^1, C_m^2, \ldots, C_m^q\}$ and $\tilde{C}_m^p = \{\tilde{C}_m^1, \tilde{C}_m^2, \ldots, \tilde{C}_m^q\}$ where $q = M/2^p$
- The distortion measure with Hausdorff distance is

$$H_m^p = max\{h(C_m^p, \tilde{C}_m^p), h(\tilde{C}_m^p, C_m^p)\}$$

where

$$h(C_m^p, \tilde{C}_m^p) = max_{i=1,2,\ldots,q}\{min_{j=1,2,\ldots,q}||c_m^j - \tilde{c}_m^j||\}$$

h ranks each point of one set based on its nearest point of the other set and uses the most mismatched point

Feature Calculation

- Trade off between good local distortion estimation and computation cost
  - The segment size *M* should not be very large
    - In our experiment, we set *M* as 1024 audio samples.
  - The dimension of the feature vector determines the training and detection complexity
- Features are extracted after distortion measure
  - High-order statistics based on the moments
  - The feature vector $V^p = \langle v_1^p, v_2^p, \ldots, v_K^p \rangle$ can be extracted with

$$v_i^p = \frac{\sum_{j=1}^{n} (f_j^p)^i \cdot d_j^p}{\sum_{j=1}^{n} d_j^p}, i = 1, 2, \ldots, K$$

    - where d_j^p is the amplitude of j^{th} frequency component f_j^p to the distortion distances D^p and *K* is the total number of moments.

Algorithm Summary

- **Step 1.** For a given audio file $x(n)$, apply wavelet de-noising to get its de-noised version $\tilde{x}(n)$.
- **Step 2.** Partition the signal $x(n)$ and $\tilde{x}(n)$ with pre-defined segment length *M*. Calculate the wavelet coefficients c_m^p and $\tilde{C}_m^p$ at different levels *p* for segment *m*.
- **Step 3.** For each wavelet decomposition level *p*, calculate the distortion measure H_m^p for all the segments.
- **Step 4.** Set up the feature vector V^p by calculating the moments of D^p for each wavelet decomposition level .
- **Step 5.** Set up the high-dimensional feature.

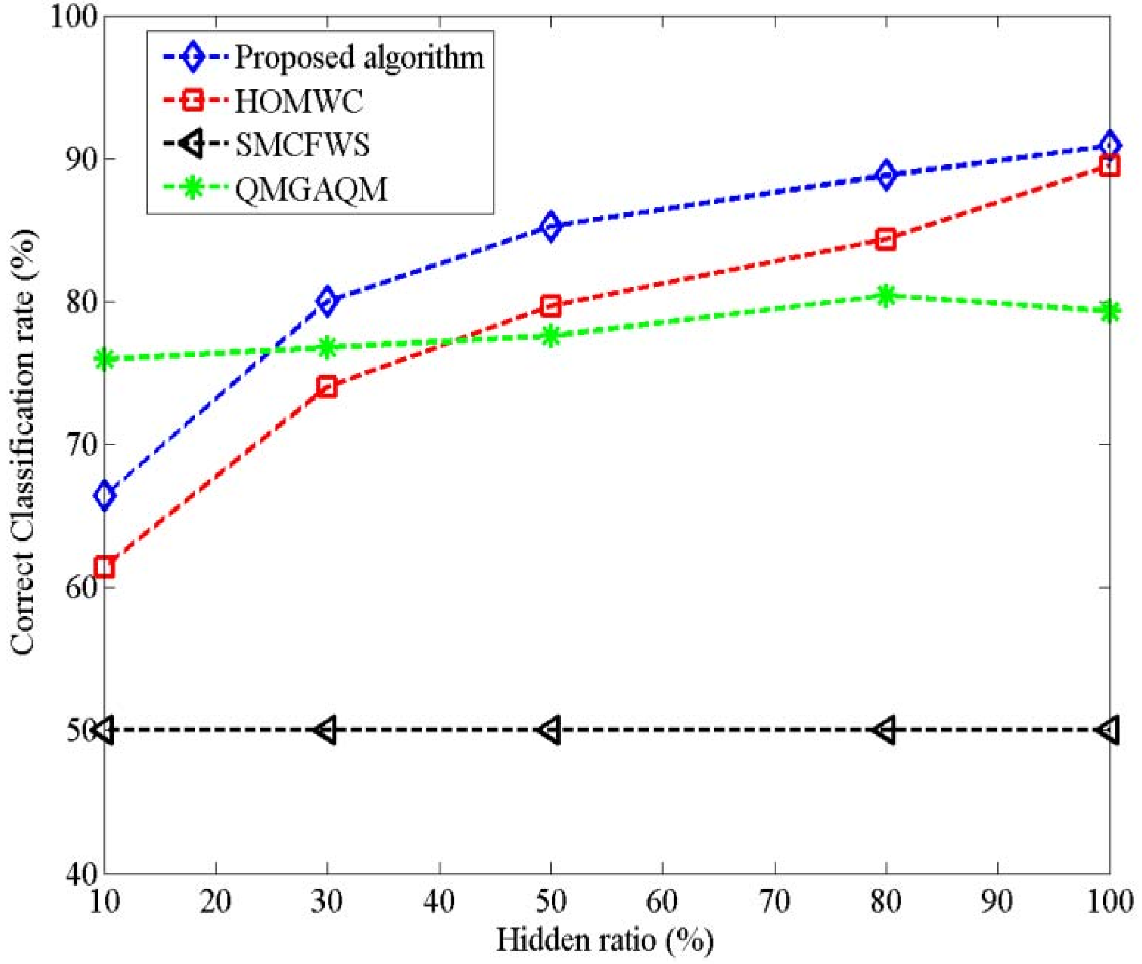
Experimental Setup

- The database
  - 994 audio files in total (from wav surfer)
- Hidden ratio
  - Percentage to the maximum hidden capacity for each audio file
- Feature vector
  - 25-dimensional feature (up to level 4 and 5 moments each)
- Steganography tools
  - Steghide
    - LSB Steganography
    - Robust against a number of different steganalysis tools.

Performance Comparison(1)

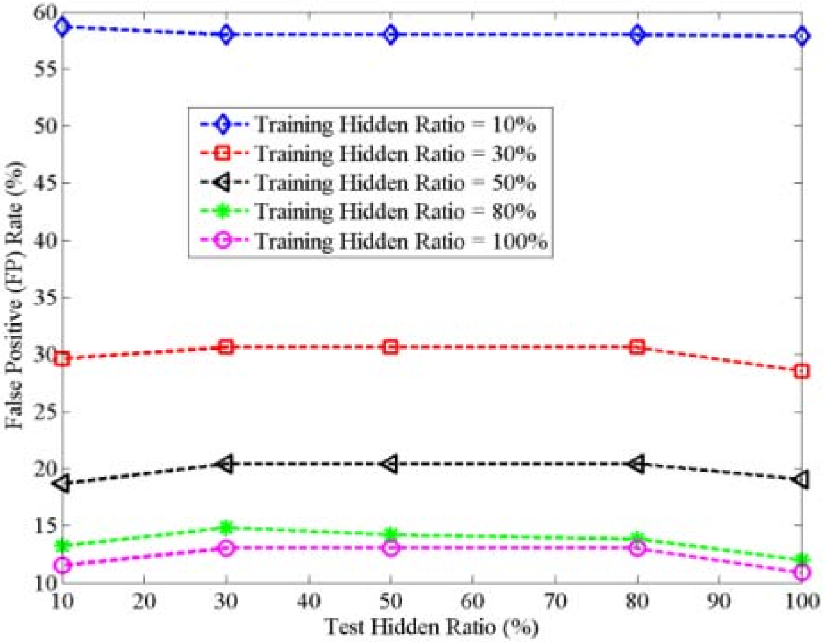
- Training set
  - 895 pair of audio files
- Test set
  - 99 pair of audio files (10% in total audio database)
- Hidden ratios
  - 10%, 30%, 50%, 80%, 100%
  - Information is known before testing, i.e the same in the training and testing
- Comparison algorithms
  - High-order moments of the wavelet coefficients (HOMWC)
  - Statistical moments of the characteristic functions of wavelet sub-bands (SMCFWS)
  - Quality measurement with general audio quality metrics (QMGAQM)

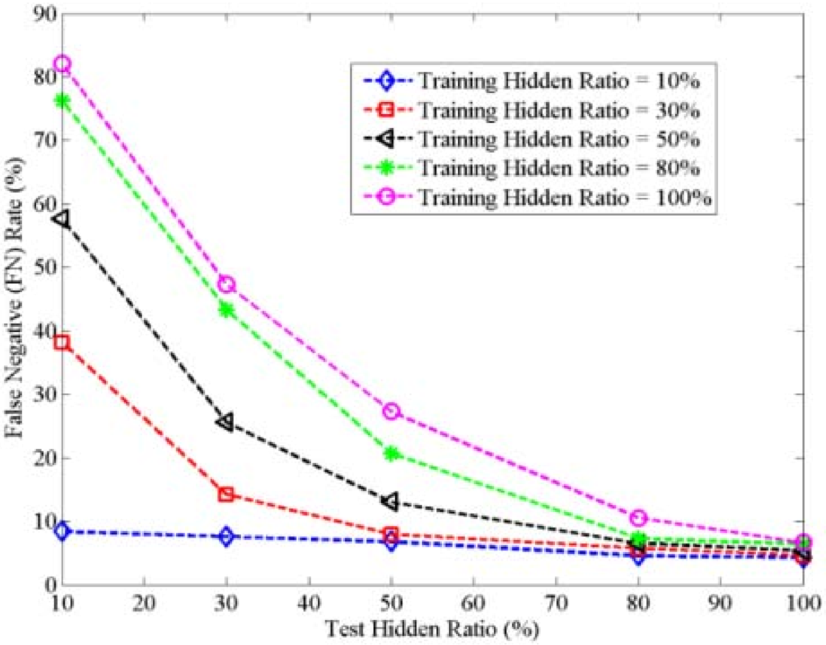
Performance Comparison (2)



Performance with Different Hidden Ratios

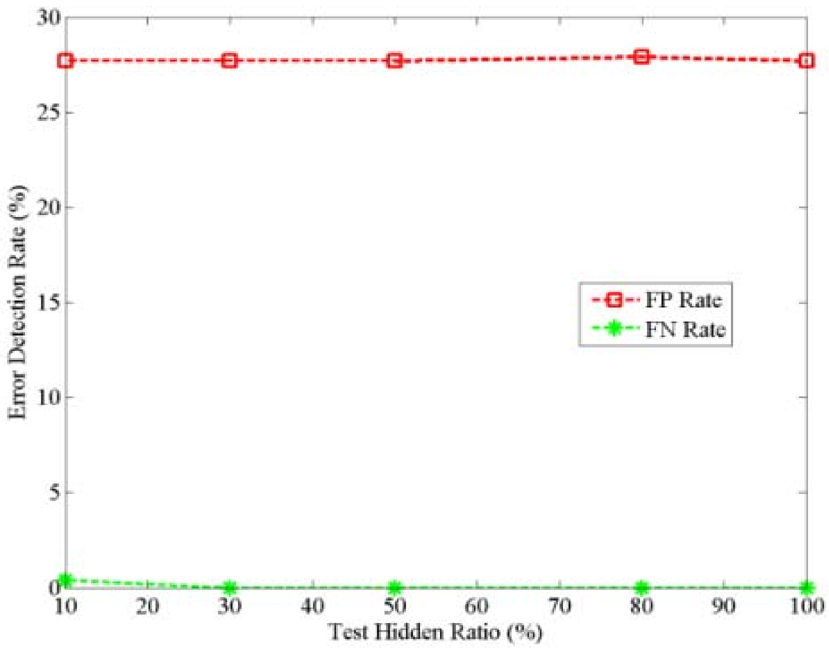
- In a real system, the hidden ratio information will be unknown before the test.

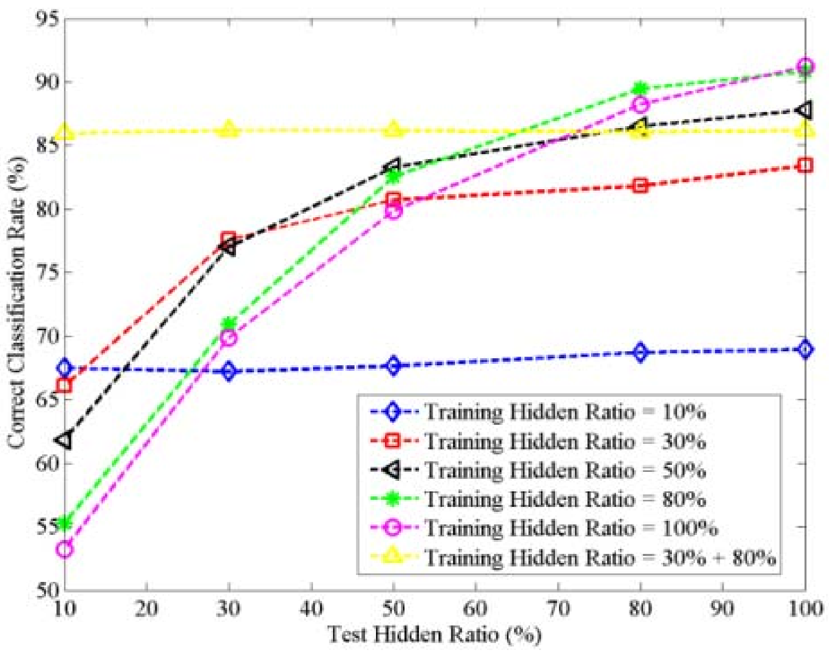




Performance with Multiple Training Hidden Ratio

- 30% + 80% training hidden ratio





Feature Effectiveness

- Correct classification rate for the training data (100% hidden ratio)

Moment Order	Level 0	Level 1	Level 2	Level 3	Level 4
1	55.30	57.42	55.02	51.79	51.06
2	62.94	65.56	55.91	53.01	51.78
3	65.90	77.23	56.19	53.57	52.40
4	66.35	79.80	58.25	55.80	52.57
5	67.18	82.75	59.04	56.25	52.68
5-D feature vector	68.97	85.77	69.87	60.71	54.30
25-D feature vector			89.45		

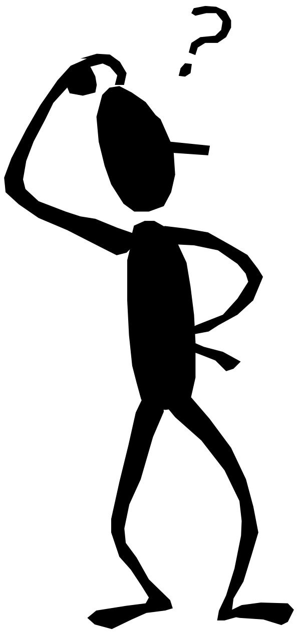
Conclusion

- We present an audio steganalysis method that is based on audio distortion measurement and high order statistics in the feature selection.

- Simulations with numerous audio sequences show that our algorithm provides higher detection rates than existing schemes that use standard audio quality metrics or statistical moments without considering audio quality.

References

- [1] Farid, H.: Detecting hidden messages using higher-order statistical models. In: IEEE International Conference on Image Processing, vol. 2, pp.905--908. New York (2002)
- [2] Harmse, J. J.: Steganalysis of Additive Noise Modelable Information Hiding. Master Thesis of Rensselaer Polytechnic Institute, Troy, New York. Thesis advisor: W. A. Pearlman (2003)
- [3] Xuan, G., Shi, Y. Q., Gao, J., Zou, D., Yang, C., Zhang, Z., Chai, P., Chen, C., Chen, W.: Steganalysis Based on multiple features formed by statistical moments of wavelet characteristic functions. In: 7th Information Hiding Workshop, pp. 262--277. Barcelona (2005)
- [4] Avcibas, I., Memon, N., Sankur, B.: Steganalysis using image quality metrics. IEEE Transaction on Image Processing, 12, 221--229 (2003)
- [5] Ozer, H., Avcibas, I., Sankur, B., Memon, N.: Steganalysis of audio based on audio quality metrics. In: Security and Watermarking of Multimedia Contents, pp.55--66. Santa Clara (2003)



Yali(Tracy) Liu
yliu@ucdavis.edu
http://networks.cs.ucdavis.edu

Questions and Comments