

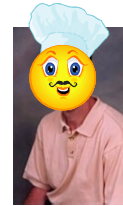
Steganalysis with Streamwise Feature Selection

Steven D. Baker
University of Virginia
sdb7e@cs.virginia.edu

Steganography: An Example



Original Image



"Hello, I am the amazing Mr. Moulin!"

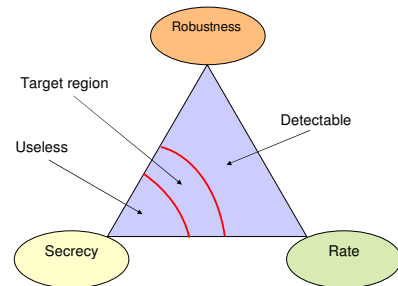


"Hello, I am the amazing Mr. Moulin..."

Motivation

- Catch bad people trying to communicate in secret
- Catch good people trying to communicate in secret?
- Research opportunities:
 - Improve detection
 - Disrupt secret communication without harming legitimate image sharing
 - Improve theoretical guarantees

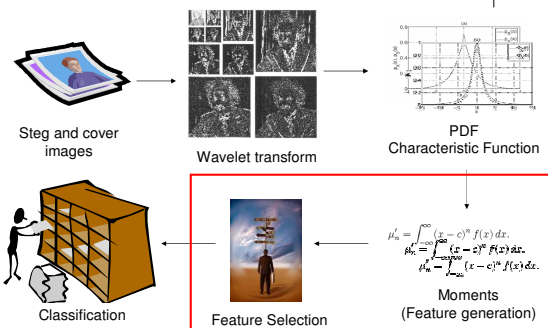
Triangle of Peril



Theoretical work in Steganography

- Complexity theory
 - Provably Secure Steganography [Hopper et al.]
- Information theory
 - An Information-Theoretic Model for Steganography [Cachin]
 - Perfectly Secure Steganography [Wang and Moulin]
- Basic conclusion: perfect security means useless rate
- No available, practical algorithm allows smooth adjustment of rate, robustness, and secrecy*

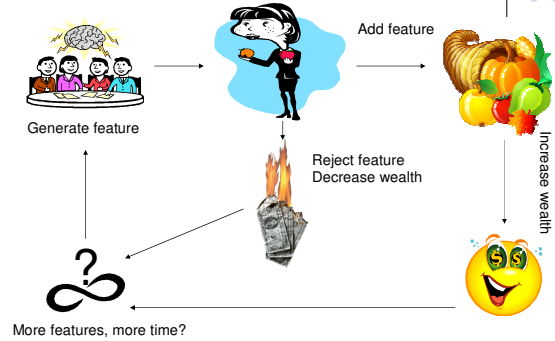
Method of Wang and Moulin



Let Intel do the work

- Can we combine existing features to form useful new features?
- Have a computer separate the useless features from the good ones
- Do this in a suboptimal but very fast way, so that you can evaluate loads of features, more than there are observations
- Streamwise Feature Selection [Zhou et al.]

Feature generation/selection (Alpha-investing)



Feature generation

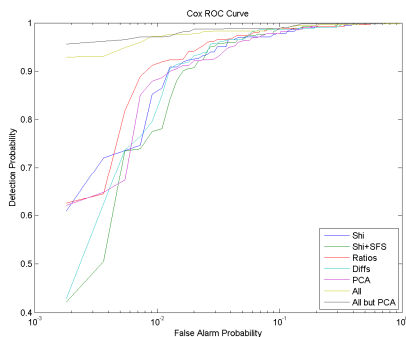
- Pair-wise ratios
- Pair-wise differences
- Principal Component Analysis
- Untested possibilities:
 - Log
 - Square root
 - Nonsmooth Nonnegative Matrix Factorization

Experimental data (Shi et al.)

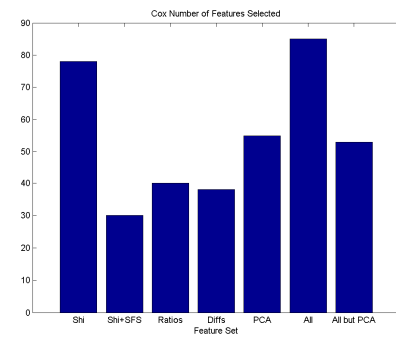
- CorelDraw images
 - ~1000 images
 - 78 Original features
 - > 6000 generated features
- Steganographic techniques
 - Cox et al. (SS)
 - Piva et al. (SS)
 - Huang et al. (SS)
 - Generic QIM
 - Generic LSB



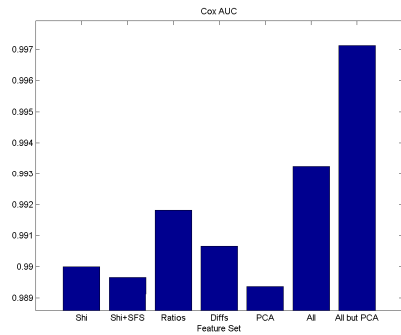
Results: ROC



Results: Model Complexity



Results: AUC



Questions?

