A Hierarchical Approach for Increasing the Stealthiness of Steganographic Methods

Mercan Topkara, Umut Topkara

Mikhail J. Atallah

Center for Education and Research in Information Assurance

Department of Computer Sciences

Cuneyt Taskiran, Eugene T. Lin

Edward J. Delp

Video and Image Processing Laboratory

School of Electrical and Computer Engineering

Key Ideas

- **A universal protocol to improve stealthiness**
  - not a pre-process or a post-process, guides embedding on the fly
- **Stego object's statistical features controlled at both fine and coarse levels**
  - tree-structured hierarchical representation

Outline of the Talk

- **Previous Work**
- **General Framework**
- **Hierarchical Representation**
- **The Protocol**
- **Setting an Upper Bound on Detectability**
- **Experiments**
- **Conclusion**

PURDUE
UNIVERSITY

Previous Work

- **Pfitzmann and Westfeld (IHW 1999)**
  - X^2 test over first order statistics
- **Westfeld, F5 (IHW 2001)**
  - decrements DCT coefficients' absolute values
  - spreads embedding using permutative straddling
  - uses matrix encoding to restrict necessary changes for a given code
- **Provos, Outguess (USENIX 2001)**
  - tags pixels with detectability values (-1,0,1)
  - embeds using the best group of bit selection
  - performs error correction using the redundant bits

Previous Work

- **Franz (IHW 2002)**
  - models embedding as a Markov source to find the best distribution and adjusts message accordingly
  - finds best groups of selection and noisy parts for alteration
- **Fridrich et al. (SPIE 2002, SPIE 2003)**
  - Steganalysis on embedding to JPEG: Searching for a macroscopic quantity that predictably changes with the length of the embedded secret message
  - Steganalysis on least significant bit embedding: RS-analysis
- **Lyu and Farid (IHW 2002)**
  - use wavelet decomposition to get the features of images
  - then these features are used for training a classifier (Fisher Linear Discriminant or Support Vector Machines) for stego or clean images
- **Sallee (IWDW 2003)**
  - uses a statistical model of the cover media to estimate

$$\hat{P}_{X_\beta | X_\alpha}(X_\beta \mid X_\alpha = x_\alpha)$$

Preliminaries

- A partitioning of the document is possible
- A quantifiable measure of *detectability* is defined
  - likelihood that a given region has embedded information
  - the degree that the statistics of the region deviate from normal behavior

General Framework

analyze

Protocol

embed

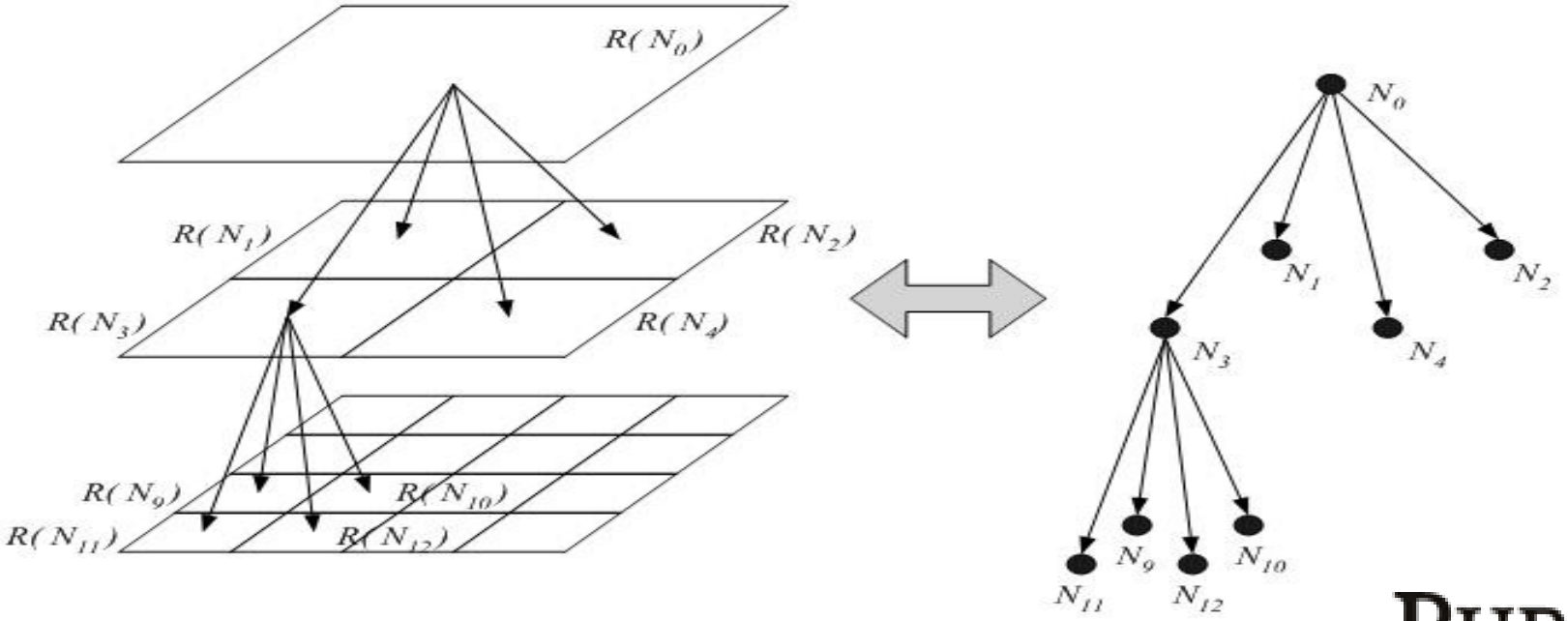
guide
&
restrict

Information Hiding System

PURDUE
UNIVERSITY

Hierarchical Representation

- **Partitioning done in a way that**
  - statistics of children directly effect statistics of parent node
  - updates in leaf level reflected to the root in O(h) steps
- **For images, representation may be a quad-tree**



PURDUE UNIVERSITY

Advantages of the Hierarchical Representation

- **Provides a structured view of the statistical properties of stego object at different resolutions**
- **Economically keeps track of the effect of embedding on the fly**
- **Possible to set an upper bound on the detectability of arbitrary regions in certain situations**
- **Provides a tool for finding the suitable regions to embed information**

The Protocol

Initialization Phase

// **Fills in the statistical information kept in the tree nodes**

for each $\mathbf{N_i}$ **in the tree in a bottom-up manner**

do $S(N_i) \longleftarrow 1$ // **Initially all nodes are marked as suitable**

if $\mathbf{N_i}$ **is a leaf node**

then **perform statistical analysis on** $\mathbf{R(N_i)}$ **to obtain** $\mathbf{T(N_i)}$

else $T(N_i) \longleftarrow \sum_{v \in children(N_i)} T(v)$

$d(R(N_i)) \longleftarrow D(T(N_i))$ // **D() returns detectability given** $\mathbf{T(N_i)}$

// **Marks the nodes that are not suitable for embedding**

for each $\mathbf{N_i}$ **in the tree in a top-down manner**

do if $d(G(b, R(N_i))) > \tau$

$S(N_i) \longleftarrow 0$ // **check current status of detectability**

then

for each $\mathbf{N_j}$ **in the sub-tree with root** $\mathbf{N_i}$

do $S(N_i) \longleftarrow 0$

The Protocol

Embedding & Dynamic Update Phase

for each $\mathbf{M_j}$ **in M**

// skip unsuitable regions

do repeat **obtain N* from embedding algorithm**

until **S(N*) = 1**

$R'(N^*) \longleftarrow G(M_j, R(N^*))$ **//embedding algorithm runs one step**

perform analysis on R`(N*) to obtain T(N*)

$N_p \longleftarrow$ parent(N*)

// reflect the statistics to upper levels

while $\mathbf{N_p}$ **is not root**

$$T(N_p) \longleftarrow \sum_{v \in \text{children}(N_p)} T(v)$$

$$d(R(N_i)) \longleftarrow D(T(N_i))$$

// check the current status of detectability

if $d(G(b, R(N_p))) > \tau$

then $S(N_p) \longleftarrow 0$

for each $\mathbf{N_j}$ **in the sub-tree with root** $\mathbf{N_p}$

do $S(N_j) \longleftarrow 0$

$N_p \longleftarrow$ parent(N_p)

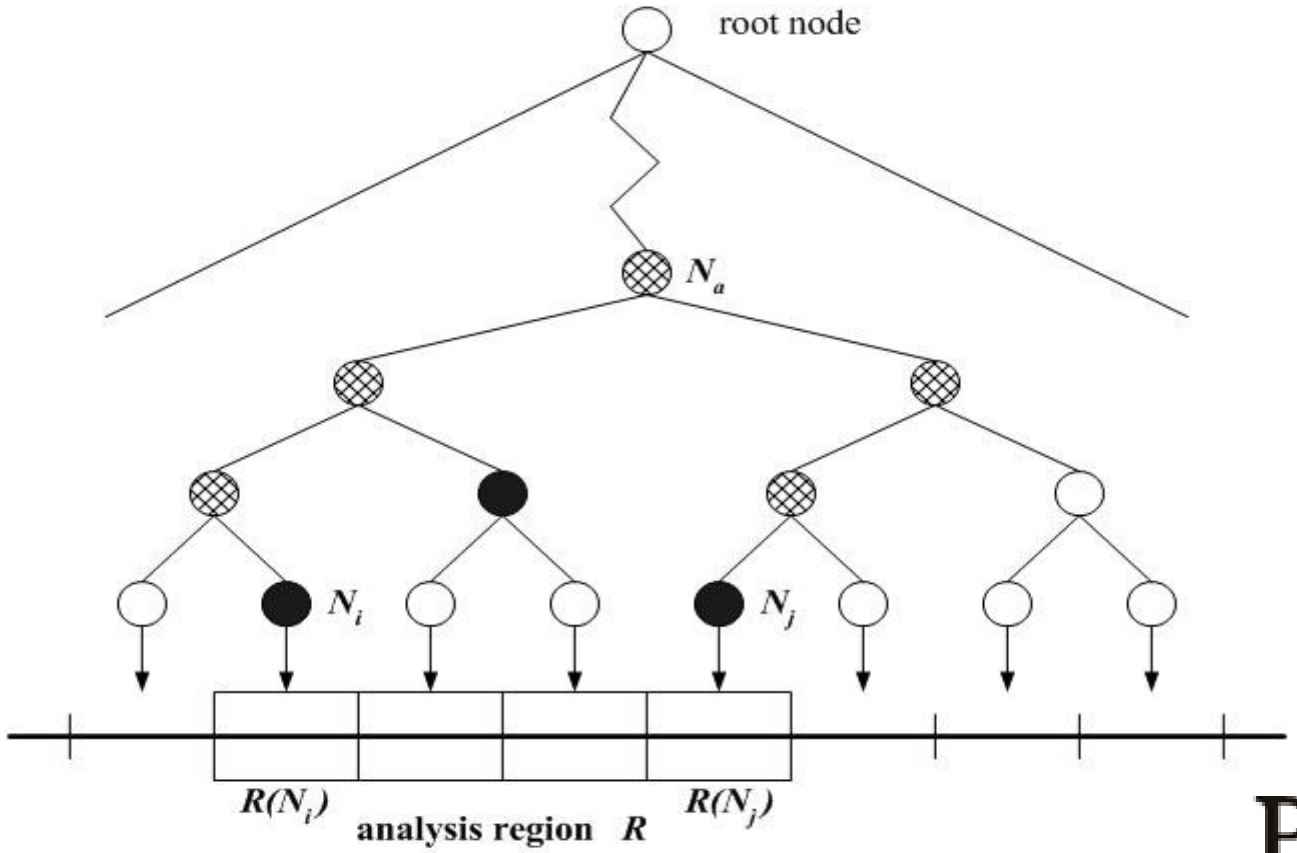
Extraction of Hidden Message

- **Fixed Threshold**
  - should be secret
  - embedding should not exceed the threshold
  - vulnerable to “try-all-thresholds” method
  - still forces the attacker to work on “hard-to-detect” regions
- **Markers (better approach)**
  - embed side information indicating skipped regions
  - sacrifices some bandwidth
  - for minimizing the markers tree-structure should be used to keep track of boundaries of avoided regions

Setting an Upper Bound on Detectability (1D)

If $d(R(N_i)) = \sum d(R(\text{children}(N_i)))$ then,

$d(R) = O(\tau \log_2 n)$



PURDUE UNIVERSITY

Setting an Upper Bound on Detectability (2D)

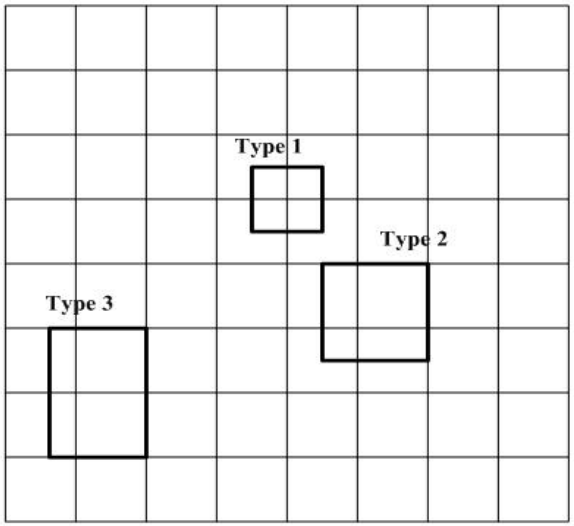
If $d(R(N_i)) = \sum d(R(\text{children}(N_i)))$ then,

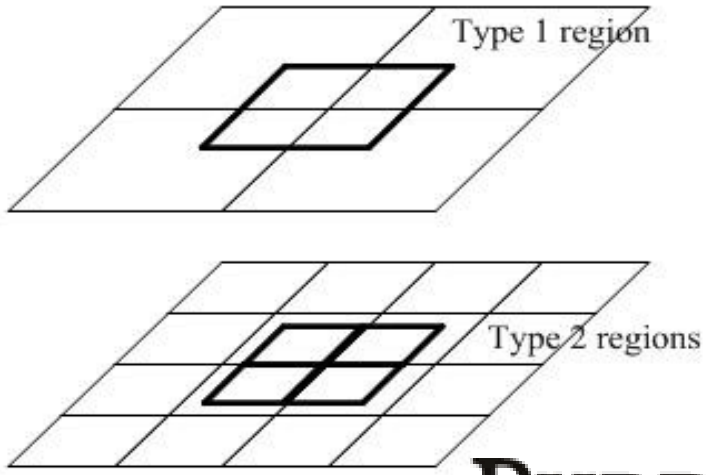
$d(R) = O(\tau\sqrt{n})$

$$d(R_1(l)) \le 4d(R_2(l-1))$$

$$d(R_2(h)) \le \tau + d(R_2(h-1)) + 2d(R_3(h-1))$$

$$d(R_3(h)) \le 2\tau + 2d(R_3(h-1))$$





PURDUE UNIVERSITY

Experimental Setup

- **A simple least significant bit (LSB) embedding algorithm for color TIFF images developed by VIPER**
- **141 TIFF images of size 512x512 obtained from the Watermark Evaluation Testbed**
- **The first chapter of the *Tale of Two Cities* as secret message (1010 words, 6KB)**
  - Embedding algorithm randomizes the message before embedding
- **Actual message length is 18%, embedding length is 100% for plain embedding and 42% on the average for embedding with the protocol**

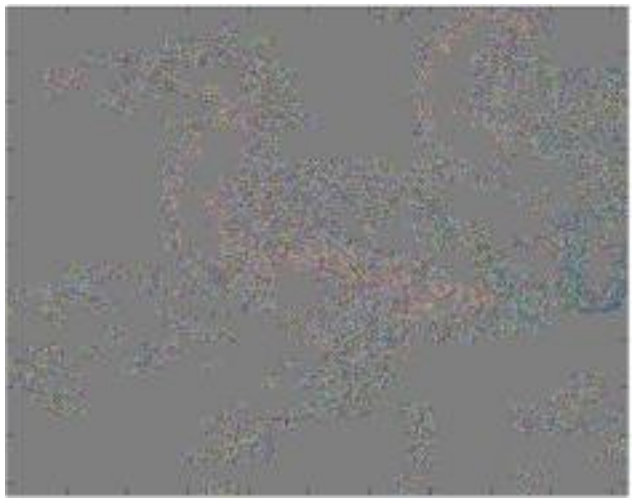
Experimental Setup

$$d(R(N_i)) = \begin{cases} -Var(R(N_i)) & \textbf{for leaf nodes} \\ \sum d(R(children(N_i)) & \textbf{for internal nodes} \end{cases}$$

$$S(N_i) = \begin{cases} 1 \text{ if } d(R(N_i)) \quad \tau_h(N_i) \text{ and } d(R(parent(N_i))) \quad \tau_h(N_i) \\ 0 \end{cases}$$

otherwise

Examples



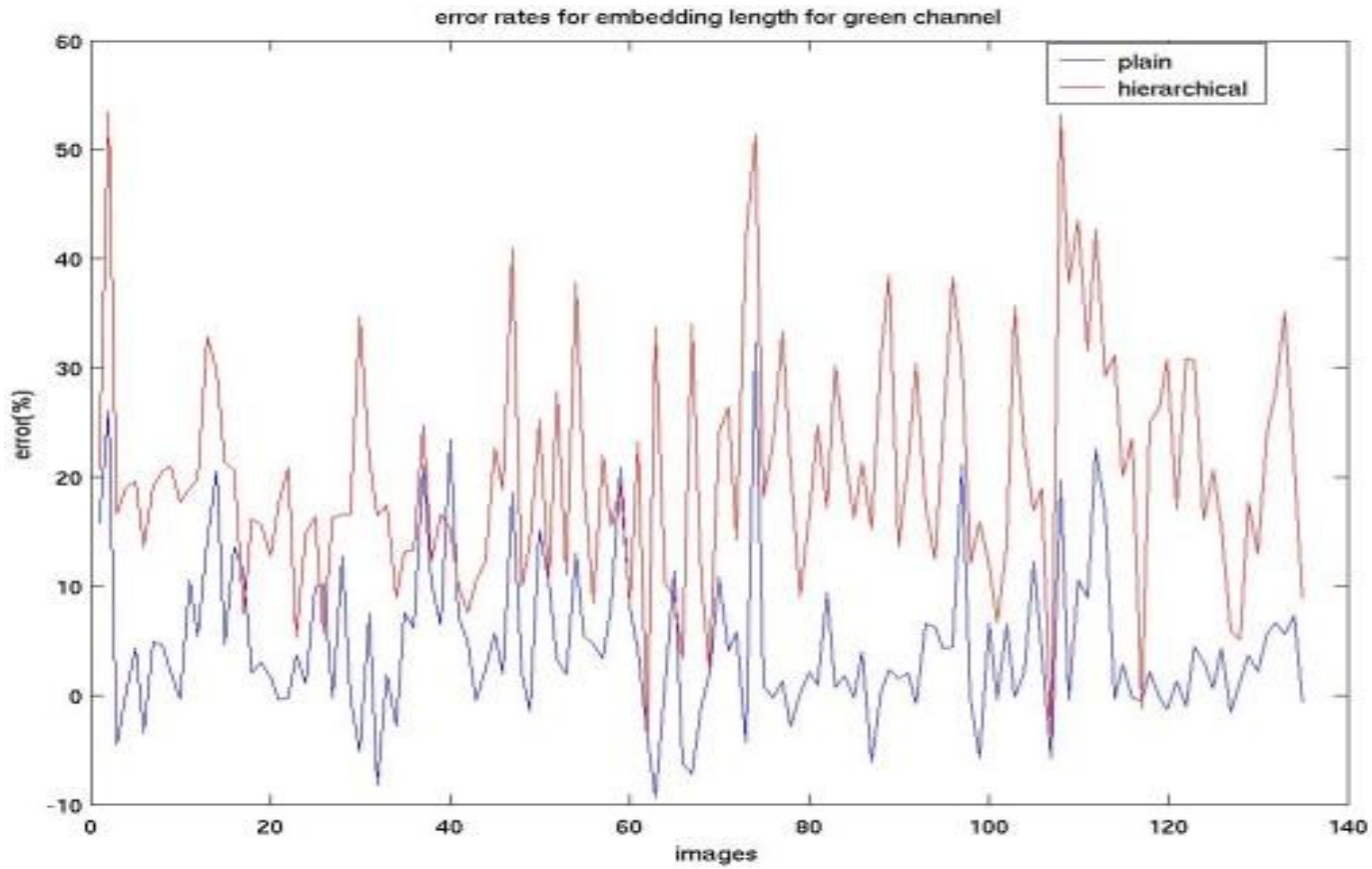
Results

- **Farid and Lyu proposed a universal attack mechanism for information hiding**
- **Support Vector Machines (SVM) and Fisher Linear Discriminant (FLD) are trained with higher order statistics of images**

classification method	plain embedding	embedding with hierarchical protocol
SVM	49.65%	42.65%
FLD	76.92%	69.23%

Results

- **RS steganalysis over the red and green color plane**
- **More successful with grayscale images and for messages that are randomly scattered over the stego-image**
- **Still increase in error rate with our system**



PURDUE UNIVERSITY

Conclusions

- **A universal protocol for improving the stealthiness of information-hiding systems**
- **Provides a mechanism for controlling statistical anomalies at both fine and coarse scales of granularity**
- **Allows continuous control on detectability**
- **Quantified how bounds on the detectability of regions from the hierarchy translate into detectability bounds for arbitrary regions in certain cases**