WARSAW UNIVERSITY OF TECHNOLOGY DEVELOPMENT PROGRAMME



Friday 12th March, 2010: 11:00 -13:00

Covert Cryptography and Steganography





J M Blackledge

Stokes Professor Dublin Institute of Technology http://eleceng.dit.ie/blackledge

Distinguished Professor Warsaw University of Technology









Lectures co-financed by the European Union in scope of the European Social Fund



What is the Problem ?





Attack ? F7&^%p£#29hGS

Attack What ?

Have a nice day







Covert Encryption Information Hiding





Related issues include:

- Camouflage
- Disinformation
- Authentication
- Self-authentication









Steganography (Greek in origin) means Covered or Concealed Writing

At what time should I confirm our activities? kindly acknowledge.

Attack now



Watermarking and Authentication









Camouflage and Disinformation









Why Should Encrypted Information be Transmitted Covertly ?







Type VII U-boot







Kurzsignalheft







Focus of the Seminar



• AudioCode: Steganography for Digital Signals

home for	s & drama documentary commercials radio children & animation music composition studio	📣 AudioCode	
Tamb@rine	Tamborine lives and breathes sound Tamborine is 100% dedicated to audo post production for television pergamming and first from richarms programming to sourmentate; from doma to animation; from movies to commencial. Tamborine is the source of	Input: Path for input file Output: Path for output file	Browse Browse
14 Livosia Street, London, WIF BAG teb 020 7434 1	812 fax: 020 7434 1813 Emi@tambotite.co.uk matk@tambotite.co.uk	Mark Compress Audio code program status bar	Authenticate

• StegoCrypt. Steganography for Digital Images



S	stego	Crypt	DIT DIT North Antion Contraction
Input: Path for input file	Browse	Output: Path for output file	Browse
Host Image: Path for output file	Browse	PIN: Please put PIN	1
EE		D	Output -> PDF



Principal Publications





Audio Data Verification and Authentication using Frequency Modulation based Watermarking, J M Blackledge and O Farooq, International Society for Advanced Science and Technology, Transactions on Electronics and Signal Processing, No. 2, Vol. 3, 51-63, 2008; <u>http://eleceng.dit.ie/papers/111.pdf</u>

A Covert Encryption Method for Applications in Electronic Data Interchange, J M Blackledge and D Dubovitskiy, International Society for Advanced Science and Technology, Transactions on Electronics and Signal Processing, No. 1, Vol. 4, 107-128, 2009; http://eleceng.dit.ie/papers/140.pdf

Printed Document Authentication using Texture Coding,

J M Blackledge and K W Mahmoud,

International Society for Advanced Science and Technology, Transactions on Electronics and Signal Processing, No. 1, Vol. 4, 81-98, 2009; <u>http://eleceng.dit.ie/papers/135.pdf</u>



Contents of Presentation I



Part I:

- Principles of Steganography
- Signal Processing Model for Information Hiding
- Linear Frequency Modulation
- Chirp Coding
- Self-authentication
- Demonstration of Self-Authentication for Audio data
- Summary
- Interval (10 Minutes)



Contents of Presentation II



Part II:

- Hiding Information in Digital Images
- Fresnel Diffusion
- Stochastic Diffusion
- Demonstration of StegoCrypt
- Hardcopy Authentication
- Summary
- Research Project Proposals
- Q & A



Plaintext \rightarrow



Transmission



Hiding Data in Images









Stegotext image
= Covertext image
+ Plaintext image









Information Hiding: A Signal Processing Model

- f Information input (*Plaintext* or *Ciphertext*)
- *s* Output signal (*Stegotext*)
- *n* Noise (*Covertext* cover signal including a cipher)
- \hat{P} Linear transformation operator (signal processor)

$$s(t) = \hat{P}f(t) + n(t), \quad \|\hat{P}f(t)\| << \|n(t)\|$$

Diffusion + Confusion *Hiding condition*





Information Retrieval 1: *Diffuser/Covertext* Retrieval

$$f = \hat{P}^{-1}(s - n)$$

- Requires knowledge of both *processor* and *covertext*
- Inverse operator must be computationally stable
- If the covertext is a cipher, then retrieval is dependent on knowledge of a private key





Information Retrieval 2: *Diffuser Only* Retrieval

$$f = \hat{P}^{-1}(s - n) = \hat{P}^{-1}s - \hat{P}^{-1}n = \hat{P}^{-1}s$$

- Requires knowledge of processor only
- Any *covertext* can be used provided $\hat{P}^{-1}n = 0$
- Require a diffuser such that:
 - the inverse operator is computationally stable
 - simple to implement



Chirp based Diffusion



- A diffuser that provides these properties is a linear frequency modulated (FM) *chirp*.
- In complex form, a linear FM chirp is given by

$$\exp(i\alpha t^2)$$

- Operator is based on *convolution*
- Inverse operator is based on *correlation*



Linear Frequency Modulation



Let

$$\hat{P}f(t) = p(t) \otimes f(t) \equiv \int_{-\infty}^{\infty} p(t-\tau)f(\tau)d\tau$$

where

$$p(t) = \exp(i\alpha t^2), \quad |t| \le \frac{T}{2}$$

Then

$$\hat{f}(t) = \exp(-i\alpha t^2) \odot \exp(i\alpha t^2) \otimes f(t), \quad |t| \le \frac{T}{2}$$

where

$$p(t) \odot f(t) \equiv \int_{-\infty}^{\infty} p(t+\tau)f(\tau)d\tau$$



Evaluation of the Correlation Integral



$$\exp(-i\alpha t^2) \odot \exp(i\alpha t^2) \equiv \int_{-T/2}^{T/2} \exp[-i\alpha(\tau+t)^2] \exp(i\alpha\tau^2) d\tau$$

$$= \exp(-i\alpha t^2) \int_{-T/2}^{T/2} \exp(-2i\alpha t\tau) d\tau = T \exp(-i\alpha t^2) \operatorname{sinc}(\alpha Tt)$$

where

$$\operatorname{sinc}(x) \equiv \frac{\sin x}{x}$$



Application of the Condition *T* >> 1



$\cos(\alpha t^2)\operatorname{sinc}(\alpha Tt) \simeq \operatorname{sinc}(\alpha Tt)$ $\sin(\alpha t^2)\operatorname{sinc}(\alpha Tt) \simeq 0$

 $\hat{f}(t) = T \exp(-i\alpha t^2)\operatorname{sinc}(\alpha T t)$ $\simeq T\operatorname{sinc}(\alpha T t) \otimes f(t)$



Spectral Response



• In Fourier space (ignoring scaling constant)

$$\hat{F}(\omega) = \begin{cases} F(\omega), & |\omega| \le \alpha T; \\ 0, & |\omega| > \alpha T. \end{cases}$$

- Retrieved information is a *band-limited* version of the input signal
- **Band-width** is determined by lpha T



Retrieval with Covertext



$$s(t) = \exp(i\alpha t^2) \otimes f(t) + n(t)$$

 $\hat{f}(t) \simeq T \operatorname{sinc}(\alpha T t) \otimes f(t) + \exp(-i\alpha t^2) \odot n(t)$

Provided the covertext does not have any features that match with n(t), then

$$\|T\operatorname{sinc}(\alpha Tt) \otimes f(t)\| >> \|\exp(-i\alpha t^2) \odot n(t)\|$$





Graphical Example







Why use Chirps ? $s(t) = \exp(i\alpha t^2) \otimes f(t) + n(t)$ $\hat{f}(t) \simeq T\operatorname{sinc}(\alpha Tt) \otimes f(t)$





Microwave Imaging











Binary code

...0110010...

$$chirp(t) = a\cos(\alpha t^2), \quad \forall t \in [0,T)$$

$$s(t) = \begin{cases} -\operatorname{chirp}(t), & t \in [0, T); \\ +\operatorname{chirp}(t), & t \in [T, 2T); \\ +\operatorname{chirp}(t), & t \in [2T, 3T); \\ -\operatorname{chirp}(t), & t \in [3T, 4T); \\ -\operatorname{chirp}(t), & t \in [4T, 5T); \\ +\operatorname{chirp}(t), & t \in [5T, 6T); \\ -\operatorname{chirp}(t), & t \in [6T, 7T). \end{cases}$$







$$s(t) \odot \operatorname{chirp}(t) = \begin{cases} -a, & t \in [0, T); \\ +a, & t \in [T, 2T); \\ +a, & t \in [2T, 3T); \\ -a, & t \in [3T, 4T); \\ -a, & t \in [4T, 5T); \\ +a, & t \in [5T, 6T); \\ -a, & t \in [6T, 7T). \end{cases}$$

Decoding

Chirp function must be an *identical replica* of that used to chirp code the binary stream



Applications



- Covert information exchange using digital signals
 - plaintext
 - ciphertext
- Covert key exchange
- Authentication of digital signals
 - Copyright protection
 - Digital Rights Management
- Self-authentication of digital signals
 - Speech
 - Audio



Self-authentication of Audio Data: The Problem



f(t) - audio signal

w(t) - watermark obtained from the audio signal s(t) - watermarked signal

Find transforms \hat{T} and \hat{L} where

$$w(t) = \hat{T}f(t)$$
 and $s(t) = f(t) + \hat{L}w(t)$

such that

$$\begin{aligned} \|\hat{L}w(t)\| << \|f(t)\| \\ \hat{T}s(t) = w(t) \quad \text{and} \quad \hat{L}^{-1}s(t) = w(t) \end{aligned}$$
 Signal Coding (?) Chirp Coding (OK





Signal Coding using the Wavelet Transformation



$$F_L(t) = \frac{1}{\sqrt{L}} \int f(\tau) W\left(\frac{t-\tau}{L}\right) d\tau$$
$$E_L = \frac{100}{E} \int |F_L(t)|^2 dt, \quad E = \sum_L E_L$$

Binary[Round(E_L)] is concatenated to produce a binary string which is then

Chirp Coded







http://eleceng.dit.ie/arg/downloads/Audio_Self_Authentication.zip

📣 AudioCode

_ 🗆 🗡

Input:	Path for input file		Browse
Output:	Path for output file		Browse
Mark	Compress	,	Authenticate





Perceptual Evaluation of Audio Quality: BS.1387







lothouse

Commercial Applications





Technology to License

Self-authentication of Audio Data for Copyright Protection



http://www.tamborine.co.uk

C ENTERPRISE IRELAND







On the Search for Extraterrestrial Intelligence



- Chirp coding provides a solution for communicating over 'channels' with very noisy environments
- Interstellar space becomes very noisy when radio waves propagate over many light years
- Suggests correlating SETI data with different chirp codes and searching for an output with minimum Information Entropy









- Covert encryption uses Steganography to hide encrypted information in a Covertext
- Chirp coding provides an effective method of hiding bit streams in digital signals which has many applications including
 - key exchange
 - authentication and copyright protection
- Chirp coding is unique in that it provides a method of self-authenticating a digital signal


In the Following Lecture...



- We shall investigate a method to hide encrypted information in digital images using the process of stochastic diffusion
- Consider an approach for *e-fraud* prevention of *e-documents*
- Investigate a method for authenticating hardcopy documents based on *texture coding*
- Provide a demonstration of the product





Questions + Interval (10 Minutes)



Contents of Presentation II



Part II:

- Hiding Information in Digital Images
- Fresnel Diffusion
- Stochastic Diffusion
- Demonstration of StegoCrypt
- Hardcopy Authentication
- Summary
- Research Project Proposals
- Q & A



Hiding Information in Digital Images





Etude no.12 "Revolution"

Chopin







Sheet Music from Snotes.com @ Copyright 2008 Red Balloon Technology Ltd

Information



Host Image



Retrieval



Basic Model



stegotext = ciphertext + covertext

 $ciphertext = cipher \otimes \otimes plaintext$

$\otimes \otimes$ denotes the 2D convolution integral

- Ciphertext generated by process of Diffusion
- Stegotext generated by process of Confusion



Fresnel Diffusion



Consider a watermarking model given by

$$I_3(x,y) = rp(x,y) \otimes \otimes I_1(x,y) + I_2(x,y)$$

with 'Fresnel' Point Spread Function (PSF)

$$p(x,y) = \frac{1}{2}(1 + \cos[\alpha(x^2 + y^2)]$$

and where

 $||p(x,y) \otimes \otimes I_1(x,y)||_{\infty} = 1$ and $||I_2(x,y)||_{\infty} = 1.$



Watermark Retrieval



$$I_1(x,y) = \frac{1}{r} p(x,y) \odot \odot [I_3(x,y) - I_2(x,y)]$$

where $\odot \odot$ denote two-dimensional correlation.

Implemented using a Fast Fourier Transform and application of the two-dimensional convolution and correlation theorems, i.e.

$$p \otimes \otimes f \Longleftrightarrow PF$$

and

$$p \odot \odot f \Longleftrightarrow P^*F$$

respectively, where \iff denotes transformation from 'image space' to 'Fourier space'.









Stochastic Diffusion









How Does it Work? 1: *Data Diffusion*



Let n(x, y) be a cipher with Fourier transform $N(k_x, k_y)$ and compute

$$m(x,y) = \mathcal{F}_2^{-1} \left[\frac{N(k_x, k_y)}{|N(k_x, k_y)|^2} \right], \quad |N(k_x, k_y)|^2 > 0$$

so that the diffused field is given by

 $I(x,y) = m(x,y) \otimes \otimes I_0(x,y).$





How Does it Work? 2: Condition for Regularisation



 $\forall k_x, k_y$

if $|N(k_x, k_y)|^2 = 0$

then $|N(k_x, k_y)|^2 = 1$



How Does it Work? 3: **Data Retrieval**



$$n(x,y) \odot \odot I(x,y) \iff N^*(k_x,k_y)\widetilde{I}(k_x,k_y)$$

and

$$N^{*}(k_{x}, k_{y})\widetilde{I}(k_{x}, k_{y}) = N^{*}(k_{x}, k_{y})M(k_{x}, k_{y})\widetilde{I}_{0}(k_{x}, k_{y})$$
$$= N^{*}(k_{x}, k_{y})\frac{N(k_{x}, k_{y})}{|N(k_{x}, k_{y})|^{2}}\widetilde{I}_{0}(k_{x}, k_{y}) = \widetilde{I}_{0}(k_{x}, k_{y})$$

so that

$$I_0(x,y) = n(x,y) \odot \odot I(x,y).$$





How Does it Work? 4: Covertext Model

 $I_3(x,y) = rm(x,y) \otimes \otimes I_1(x,y) + I_2(x,y)$

 $||m(x,y) \otimes \otimes I_1(x,y)||_{\infty} = 1$ and $||I_2(x,y)||_{\infty} = 1$

- r is the *Diffusion-to-Confusion* watermarking ratio
- *m* is a *pre-conditioned* stochastic field
- *n* is a **key dependent cipher**



Further Example of Watermarking by Stochastic Diffusion



















Image Data Diffusion







Data Redundancy



- For binary plaintext images, stochastic diffusion (with a grey level stochastic field) yields a field that is data redundant.
- The data field can therefore be binarized to compress the encrypted information





Principal Algorithms 1



Algorithm I: Encryption and Watermarking Algorithm

Step 1: Read the binary plaintext image from a file and compute the size $I \times J$ of the image.

Step 2: Compute a cipher of size $I \times J$ using a private key and pre-condition the result.

Step 3: Convolve the binary planitext image with the preconditioned cipher and normalise the output.

Step 4: Binarize the output obtained in Step 3 using a threshold based on computing the mode of the Gaussian distributed ciphertext.

Step 5: Insert the binary output obtained in Step 4 into the lowest 1-bit layer of the host image and write the result to a file.



Principal Algorithms 2



Algorithm II: Decryption Algorithm

Step 1: Read the watermarked image from a file and extract the lowest 1-bit layer from the image.

Step 2: Regenerate the (non-preconditioned) cipher using the same key used in Algorithm I.

Step 3: Correlate the cipher with the input obtained in Step 1 and normalise the result.

Step 4: Quantize and format the output from Step 3 and write to a file.







http://eleceng.dit.ie/arg/downloads/StegoCrypt.zip

Technology to License

Document Authentication for Electronic Data Interchange

Dublin Institute of Technology (DIT) is seeking companies to license a novel technology that provides a facility for authenticating documents (letters, certificates, speed sheets etc.) communicates via the Internet as attachments.

StegoCrypt			DIT.
Input: Path for input file	Browse	Output: Path for output file	Browse
Host Image: Path for output file	Browse	PIN: Please put PIN]
StegoCrypt program status bar	E	D	Output -> PDF

Encryption Mode	Decryption Mode	
Inputs:	Inputs:	
Plaintext image	Stegotext image	
Covertext image	Private key (PIN)	
Private Key (PIN)		
Output:	Output:	
Watermarked Covertext image	Decrypted watermark	
Operation:	Operation:	
Encrypt by clicking on	Decrypt by clicking on	
button E (for Encrypt)	button D (for Dycrypt)	









CITY AND GUILDS OF LONDON INSTITUTE Incorporated by Royal Charter

This Diploma certifies that the Council of the Institute has conferred upon

Professor Jonathan Michael Blackledge 1850 PhD DIC FBCS FIGE FIMeche FIOP FRSS FIMA FELLOWSHIP

of the City and Guilds of London Institute in recognition of outstanding professional and personal achievement

2007



Michael Howell Chairman of Council





CITY AND GUILDS OF LONDON INSTITUTE Incorporated by Royal Charter

This Diploma certifies that the Council of the Institute has conferred upon

Professor Jonathan Michael Blackledge BSC PHD DIC FBCS FIEF FIMECHE FIOP FRSS FIMA FELLOWSHIP









Authentication of e-Letters





SCHOOL OF ELECTRICAL ENGINEERING SYSTEMS. FACULTY OF ENGINEERING

Jonathan Blackledge Stokes Professor of DSP http://eleceng.dit.ie/blackledge



cc: Prof Eugene Coyle Dr Marek Rebow



Dublin Institute of Technology Kevin Street, Dublin 8, Ireland

Tel: +35 3 1 402 4707 Email: jonathan.blackledge@dit.ie

4 August, 2009

Dear Sir

Re: A Covert Encryption Method for Applications in Electronic Data Interchange

Please find enclosed the manuscript for the above paper which I am submitting to the ISAST Transactions on Electronics and Signal Processing.

Yours Faithfully

J M Blackledge Stokes Professor



of TECHNOLOGY Incompared Designation from Minute Asky Climb

Dublin Institute of Technology Kevin Street, Dublin & Ireland

Tet +35/3 1:402:4707. Email: jonathan black ledge@dit.ie

4 August 2009

Re: A Covert Encryption Method for Applications in Electronic Data Interchange

Please and enclosed the menuscript for the above paper which i am submitting to the ISAST Transactions on Electronics and Signal Processing.

Yours Faithfully



J M Blockledge Stokes Professor



Camouflage





SCHOOL OF ELECTRICAL ENGINEERING SYSTEMS, FACULTY OF ENGINEERING

Jonathan Blackledge Stokes Professor of DSP http://eleceng.dit.le/blackledge



cc: Prof Eugene Coyle Dr Marek Rebow DIT DUBLIN INSTITUTE of TECHNOLOGY Institutió Teicueldistre Braile Arba Citach

> Dublin Institute of Technology Kevin Street, Dublin 8, Ireland

Tel: +35 3 1 402 4707 Email: jonathan.blackledge@dit.ie

4 August, 2009

Dear Sir

Re: A Covert Encryption Method for Applications in Electronic Data Interchange

Please find enclosed the manuscript for the above paper which I am submitting to the ISAST Transactions on Electronics and Signal Processing.

Yours Faithfully

J M Blackledge Stokes Professor







Jonathan Blackledge Stokes Professor of DSP http://eleceng.dit.ie/blackledge



Dublin Institute of Technology Kevin Street, Dublin 8, Ireland

Tel: +35 3 1 402 4707 Email: jonathan.blackledge@dit.ie

cc: Prof Eugene Coyle Dr Marek Rebow

4 August, 2009

Dear Sir

Re: A Covert Encryption Method for Applications in Electronic Data Interchange

Please find enclosed the manuscript for the above paper which I am submitting to the ISAST Transactions on Electronics and Signal Processing.

Yours Faithfully

J M Blackledge Stokes Professor

MS Word

(Format→Background→Fill Effect...) (Format→Background→Printed Watermark...)



Other Applications



• Disinformation:

Watermark one letter (consisting of disinformation to be intercepted) with another (secret information)

• Plausible Deniability

Watermark one letter (consisting of information of value to an attacker) with another (consisting of secrete information) and encrypt the result

- Cribb Camouflage
- Covert Key Exchange



Hardcopy Authentication using Stochastic Diffusion



• The covertext model

$$I_3(x,y) = rm(x,y) \otimes \otimes I_1(x,y) + I_2(x,y)$$

can not be applied to hardcopy applications due to the de-registration and distortion of pixels that occurs with covertext removal

• However, we can use a *diffusion only* approach

 $I(x,y) = m(x,y) \otimes \otimes I_0(x,y)$ Texture Coding







 $I_{\text{print}} = p_{\text{print}} \otimes \otimes m \otimes \otimes I_0$

$$I_{\rm scan} = p_{\rm scan} \otimes \otimes I_{\rm print}$$

Because convoution is *commutative*

 $I_{\text{scan}} = p_{\text{scan}} \otimes \otimes p_{\text{print}} \otimes \otimes m \otimes \otimes I_0$ $= m \otimes \otimes p_{\text{scan/print}} \otimes \otimes I_0$



Conditions Required for Hidden Data Retrieval



- I_{scan} must be re-sampled to the size of the original e-image I_0 before correlating with n
- Fidelity of the reconstruction critically depends on:
 - orientation
 - cropping
- Method is robust to *hardcopy soiling*



Applications of Texture Coding 1: Identity Cards







with flat-bed scanner at 300dpi

Printed at 600dpi; scanned with mobile phone camera



Applications of Texture Coding 2: Signature Authentication











Applications of Texture Coding 3: Passport Authentication











Printed at 400dpi;

Scanned with flat-bed scanner at 300dpi



Applications of Texture Coding 4: *Currency Authentication*









Binary texture code printed using UV ink at 150 dpi

Scanned with camera at at 300dpi under UV lamp



Applications of Texture Coding 5: Statistical Authentication











Texture code generated of basic statistics associated with a scan of a high value bank bond and printed on the back of the bond at 300dip; flat-bed scanned at 150dpi



Attack and Robustness Analysis



Printed Document Authentication using Texture Coding,

J M Blackledge and K W Mahmoud, International Society for Advanced Science and Technology, Transactions on Electronics and Signal Processing, No. 1, Vol. 4, 81-98, 2009; <u>http://eleceng.dit.ie/papers/135.pdf</u>





Summary



Fundamental steganographic model

$$I_3(x,y) = rm(x,y) \otimes \otimes I_1(x,y) + I_2(x,y)$$

Diffusion+ConfusionCiphertext+Covertext

Retrieval of I₁ requires knowledge of the
Covertext and the Key used to compute m



Summary (Continued)



$$I_3(x,y) = rm(x,y) \otimes \otimes I_1(x,y) + I_2(x,y)$$

- Self-Authentication: $I_1 = I_2$
- Stegocrypt: Based on binarisation of ciphertext
- Binary ciphertext embedded in covertext using 1-bit layer replacement method





- Diffusion + Confusion model suitable for electronic-to-electronic (e-to-e) applications
- For hardcopy authentication, a *diffusion only* approach is used called *Texture Coding*
- Based on an application of the model

$$I(x,y) = m(x,y) \otimes \otimes I_0(x,y)$$



Research Project Proposal FP7 Security (SCRYPTOMED














Q&A