Steganography and Forensic Steganalysis in JPEG using Benford's Law

Theo Tryfonas, Panagiotis Andriotis

IPICS 2014

Steganalysis with Benford's Law

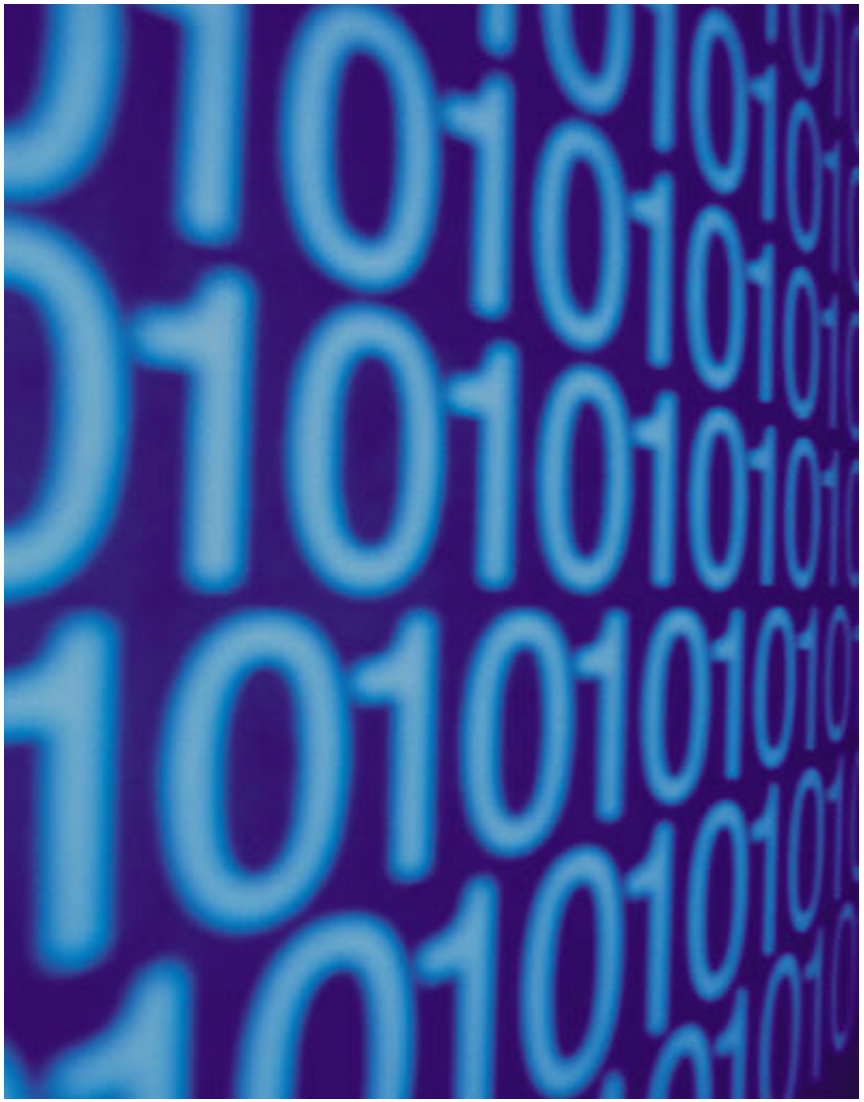
From various papers with contributions from:

Panagiotis Andriotis, Alex Zaharis, Dini Martini, Theo Tryfonas, George Oikonomou et al.

@TheoTryfonas

@PanosAndriotis

University of BRISTOL

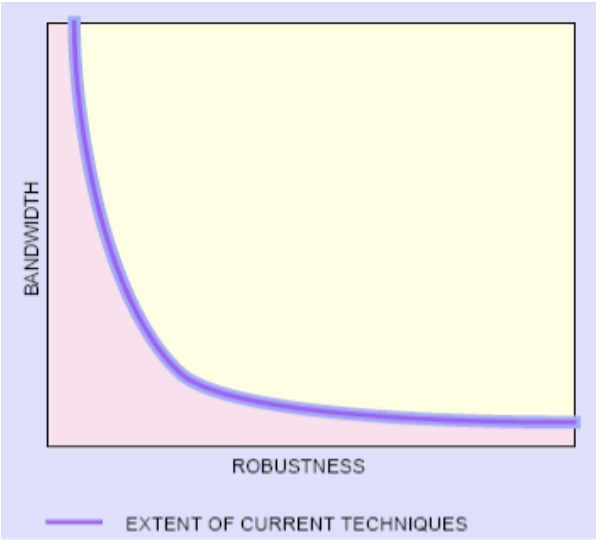


Outline

- Briefly on steganography and steganalysis
- Briefly on Benford's Law and applications
- Briefly on the JPEG format
- Applying Benford's Law to detect JPEG steganography
  - Raw byte values (most details, fairly straightforward)
  - DCT coefficients (briefly, requires some image processing background)
- Further work

Data Hiding

- Data insertion into existing data with the intention of:
  - fingerprinting
  - digital watermarking
  - covert communication



The robustness of the host signal reduces with the bandwidth (volume) of embedded data.

More information at: http://www.garykessler.net/library/fsc_stego.html

Types of Data Hiding

- Media Management Layer
  - Use of areas that the OS is unaware of (Unallocated space, Host Protected Area, Partition Gap, MBR-area*)
- File System Layer
  - Exploitation of file system structures vulnerabilities (Slack Space, NTFS Alternate Data Streams, Reserved inodes - EXT2/3)
- Application Layer
  - Steganography

*Master Boot Record area

Steganography is not Cryptography

Steganography

- Alice and Bob want to hide the fact they are exchanging data through a medium.

Cryptography

- Alice and Bob exchange messages using a special communication format. They do not hide their activity, they just protect their privacy.

Embedding secret messages in images

- "Fuse":
  - Embedding the secret information within the file exploiting its file structure.
  - Could be used with multiple file types.
- "Least Significant Bit (LSB) Encoding":
  - Hiding 1 bit of data in every pixel of 8-bit images.
  - Hiding 3 bits of data in every pixel of 24-bit images
    - Very sensitive in change of format and encoding of the images (e.g. save from .GIF to .JPEG).

Example of LSB encoding manipulation

- Hiding the letter G in the following bit stream:

10010101 00001101 11001001
10010110 00001111 11001011
10011111 00010000

- G → 01000111

10010100 00001101 11001000
10010110 00001110 11001011
10011111 00010001

Embedding secret messages in images (cont'd)

- Takes advantage of the limitations of the human vision system (HVS).
- Anything that can be coded into a bit stream can be embedded in an image.
- 8-bit:
  - Small.
  - Only 256 colours available.
- 24-bit:
  - Better for steganography
  - Large number of possible colours (>16M) exceeds HVS capabilities for differentiation.
- Compression:
  - "lossy", the secret message may lose integrity because the compression algorithm reduces the image fidelity (JPEG).
  - "lossless", retains image properties at the expense of image size - good for steganography (GIF, BMP).

Steganography

is the art of concealing a 'signal' within another 'signal' (informal definition).

Terminology:

Comes from the Greek words ***στεγανός*** and ***γράφειν*** (concealed writing)

Payload: data to be covertly communicated

Carrier: signal into which the payload is hidden

Channel: type of input, e.g. JPEG images

Stego: the resulting signal

Suspect, candidate: set of files considered likely to contain a payload

- *Historic Facts - Examples:*
- Using wax tablets
- On messengers' bodies
- Invisible Ink
- Using different typefaces (normal or italic) or spacing
- Microdots
- Hide messages behind postage stamps
- Etc.

Steganography in the news

USA TODAY
Home
News
Money
Sports
Life
Tech
Main Categories
Tech briefs
Web Guide
Tech Investor
Product reviews
More Tech
Columnists
Shareware Shelf
Talk Today
Weather

ARCHIVES
SEARCH FOR
:WSPAPER ARTICLES
CLICK HERE

NEW E-MAIL
GET NEWS
IN YOUR INBOX
ick here to get the
ily Briefing in your
inbox

Tech

▪ E-mail this story ▪ Subscribe to the newspaper ▪ Sign-up for e-mail news

02/05/2001 - Updated 05:22 PM ET

Terrorist instructions hidden online

By Jack Kelley, USA TODAY

WASHINGTON — Osama bin Laden and other Muslim extremists are posting encrypted, or scrambled, photographs and messages on popular Web sites and using them to plan terrorist activities against the United States and its allies, U.S. officials say. The officials say bin Laden and his associates are using the Internet to conduct what some are calling "e-*jihad*," or holy war. Bin Laden, a dissident Saudi businessman, has been indicted for the 1998 bombing of two U.S. embassies in East Africa and is believed to be responsible for last fall's bombing of the USS Cole in Yemen. Four alleged bin Laden associates went on trial Monday in federal court in New York for the embassy bombings. "To a greater and greater degree, terrorist groups, including Hezbollah, Hamas, and bin Laden's al Qaida group, are using computerized files, e-mail, and encryption to support their operations," CIA Director George Tenet wrote last March to the Senate Foreign Relations Committee. The testimony, at a closed-door hearing, was later made public.

Read more

Related story

- Terror groups hide behind Web encryption

Through weeks of interviews with U.S. law-enforcement officials and experts, USA TODAY has learned new details of how extremists hide maps and photographs of terrorist targets — and post instructions for terrorist activities — on sports chat rooms, pornographic bulletin boards and other popular Web sites. Citing security concerns, officials declined to name the sites. Experts say it's

USA TODAY
Home
News
Main Categories
Top News
Nation
States
Washington/Politics
World
Editorial/Opinion
Health & Science
Census
Offbeat
More News
Columnists
Lotteries
City Guides
Government Guide
Talk Today
Money
Sports
Life
Tech
Weather

Search Go
Site Web
By LYCOS

ARCHIVES
SEARCH FOR
NEWSPAPER ARTICLES
CLICK HERE

NEW E-MAIL
GET NEWS
IN YOUR INBOX
Click here to get the

World

▪ E-mail this story ▪ Subscribe to the newspaper ▪ Sign-up for e-mail news

07/10/2002 - Updated 04:23 PM ET

Militants wire Web with links to jihad

By Jack Kelley, USA TODAY

Jihadunspun.net supports a holy war against the West.

ISLAMABAD, Pakistan — One Web site urges Muslims to travel to Pakistan to "slaughter American soldiers." Another solicits donations to buy dynamite to "blow up Israeli Jews." A third shows new videotape of Osama bin Laden and promises film clips of American casualties in Afghanistan. As the United States and its allies hunt them in caves, mountains and jungles, al-Qaeda, Hamas and dozens of other militant Muslim groups are increasingly turning to the Internet to carry on their jihad, or holy war, against the West, U.S. law enforcement officials and experts say. It has become one of al-Qaeda's primary means of communication, they say. The groups use Web sites to plan attacks, recruit members and solicit donations with little or no chance of being caught by the FBI or other law enforcement agencies, officials say.

Read more below

Stories

- Agents pursue terrorists online
- Researchers: No secret bin Laden messages on sites
- This Jihad Web site brought to you by...Visa?
- Bin Laden's cybertrail proves elusive

This new cyber-battlefield is allowing al-Qaeda and other groups to stay "several

Steganography on the TV

Applications of steganography over computer networks

- “Covert Channels” - creation of a secret channel over a communication network
- “Containers” - secret messages in seemingly innocent communications e.g. picture attachments containing design documentation
- “Digital Watermarking” - for protection of intellectual property and the detection of illegal use of copyrighted materials

Digital Watermarking

Are media features added during their production or distribution.

Digital watermarks are steganographically embedded within the data.

Watermarks are not DRM mechanisms, but they can be used as part of a copyright enforcement system.

Watermarks can be used to:

- record the copyright owner or the distributor
- track the distribution chain and identify the purchaser of the media

Steganalysis

- Steganalysis is the process of detection and extraction of hidden messages from a carrier.
- It uses statistical and mathematical techniques to reduce as much as possible the range of suspicious files.
  - But sometimes all files may be suspected.
  - Embedded content may be encrypted.

Types of steganalysis

- Stego only attack – where available is only the stego-object (carrier).
- Known cover attack – initial cover object and corresponding stego object available to the analyst
- Known message attack – the secret message is available along with the stego object.
- Chosen stego attack – the algorithm (stego tool) and the stego-object are available.
- Chosen message attack – for given secret message we can create the corresponding stego object.
- Known stego attack – the algorithm (stego tool), the cover object and the stego-object are available.

*More info: SANS Institute

Steganography Tools

Text Steganographic Tools	Plain Text	Other	Source Code	License	Production
PGPn123		Yes		Shareware	Yes
Nicetext	Yes		Yes	Open Source	Yes
Snow	Yes		Yes	Open Source	Yes
Texto	Yes		Yes	Open Source	Yes
Sam's Big Play Maker	Yes		Yes	Open Source	Yes
Steganosaurus	Yes		Yes	Open Source	Yes
FFEncode	Yes			Open Source	Yes
Mimic	Yes			Open Source	Yes
wbStego	Yes	HTML, PDF	Yes	Open Source	Yes
Spam Mimic	Yes			Not Specified	Yes
Secret Space	Yes			Not Specified	Yes
WitnesSoft	Yes	Yes		No longer in production	
MergeStreams		Hides excel file in word		Freeware	Yes
Steganos	Yes	HTML		Commercial	Yes
Invisible Secrets		HTML		Commercial	Yes

Text

Audio Steganographic Tools	MP3	WAV	Others	Production	License
Info Stego	Yes			Yes	Shareware
ScramDisk		Yes		Yes	Shareware
MP3Stego	Yes			Yes	Open Source
StegoWav		Yes		Yes	Open Source
Hide4PGP	Yes		VOC	Yes	Open Source
Steghide		Yes	AU	Yes	Open Source
S-Tool		Yes		Yes	Open Source
Invisible Secrets		Yes		Yes	Commercial
Paranoid			Yes	Yes	Commercial
Steganos		Yes	VOC	Yes	Commercial

Sound

Image Steganographic Tools	BMP	JPEG	GIF	PNG	TGA	Other	Production	License
Crypto123	Yes	Yes					Yes	S
Hermetic Stego	Yes						Yes	S
IBM DLS	Yes	Yes	Yes	Yes			Yes	S
Invisible Secrets	Yes	Yes		Yes			Yes	S
Info Stego	Yes	Yes	Yes				Yes	S
Syscop		Yes					Yes	S
StegMark	Yes	Yes	Yes	Yes	Yes	TIF	Yes	S
Cloak	Yes							S
Contraband Hell	Yes						Yes	F
Contraband	Yes						Yes	F
Dound	Yes						Yes	F
Gif it Up			Yes				Yes	F
Camouflage				Yes	Yes		Yes	F
Hide and Seek	Yes		Yes				Yes	F
InThePicture	Yes						Yes	F
S-Tools	Yes						Yes	F
Jpegx		Yes					Yes	F
Steganos	Yes					DIB	Yes	F
BMP Secrets	Yes							
DCT-Steg		Yes						
Digital Picture Envelope	Yes							
EikonAmark		Yes						
Empty Pic			Yes					
Encrypt Pic	Yes							
EzStego			Yes					
BMP Embed	Yes							
BMPTable	Yes							
StegoTif					Yes	TIF		
Hide Unhide						TIF		
In Plain View	Yes							
Invisible Encryption			Yes					
JK-PGS						PPM		
Scytale						PCX		
appendX		Yes	Yes	Yes				
Total	20	10	9	5	3	6	17	

S – Shareware License
F – Freeware License

Image

Image Steganographic Tools	JPEG	BMP	Others	Embedding Approach	Production
Blindside		Yes		SDS	Yes
Camera Shy	Yes			SDS	Yes
dc-Steganograph			PCX	TDS	
F5	Yes	Yes	GIF	TDS	Yes
Gif Shuffle			GIF	Change the order of the color map	Yes
Hide4PGP		Yes		SDS	Yes
JP Hide and Seek	Yes			SDS	Yes
Jsteg Jpeg	Yes			SDS	Yes
Mandelsteg			GIF	SDS	Yes
OutGuess	Yes		PNG	TDS	Yes
PGM Stealth			PGM		Yes
Steghide		Yes		SDS	Yes
wbStego		Yes		SDS	Yes
WnStorm			PCX		Yes

TDS - Transform Domain Steganography
SDS - Spatial Domain Steganography (LSB Replacement and LSB Matching)

Steganography Tools

File System Steganographic Tools	Location of Embedding	Source Code	License	Production
Disk Hide	Windows Registry	No		No
Drive Hider	Windows Registry	No		No
Easy File & Folder Protector	VXD driver, Windows Kernel	No	Shareware	Yes
Invisible Files 2000	Hard Disk	No	Shareware	Yes
Magic Folders	File System	No	Shareware	Yes
Dark Files	File system	No	Shareware	Yes
bProtected 2000	File system	No	Shareware	Yes
BuryBury	File system	No	Shareware	Yes
StegFS	File system	Yes	Open Source	Yes
Folder Guard Jr	File System	No	Freeware	Yes
Dmagic	File System	No	Freeware	Yes
BackYard	File System	No		No
Snowdisk	Disk space			No
Masker	Any file (Image, Text, Audio, Video)	No	Shareware	Yes
Anahtar	3.5-inch diskette	No		No
Hide Folders		No	Shareware	Yes
Hidden		No		No
Paranoid		No		No
Diskhide		No		No

Disc and file system

Miscellaneous

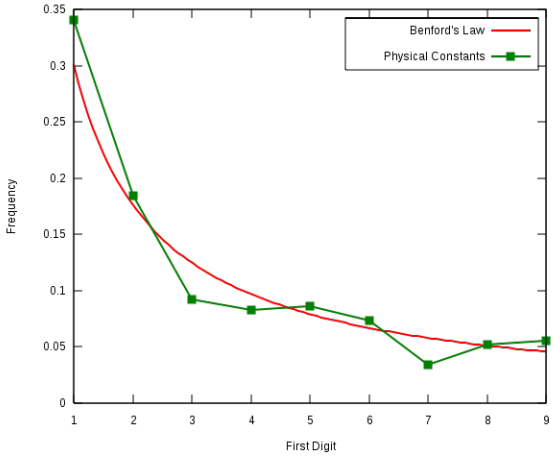
Miscellaneous Steganographic Tools	Cover Media	Source Code	License
GZSteg	.gz files	Yes	
InfoStego	Image, audio, video		Shareware
KPK File	Word, BMP		Shareware
S-Mail	.exe and .dll files		
Hiderman	Many different media		Shareware
StegMark	Image, audio, video		
Steghide	JPEG, BMP, WAV, AU	Yes	
S-Tools	BMP, GIF, WAV	Not sure	
Hydan	Program Binaries	Yes	Open Source
Covert.tcp	TCP/IP Packets	Yes	Open Source

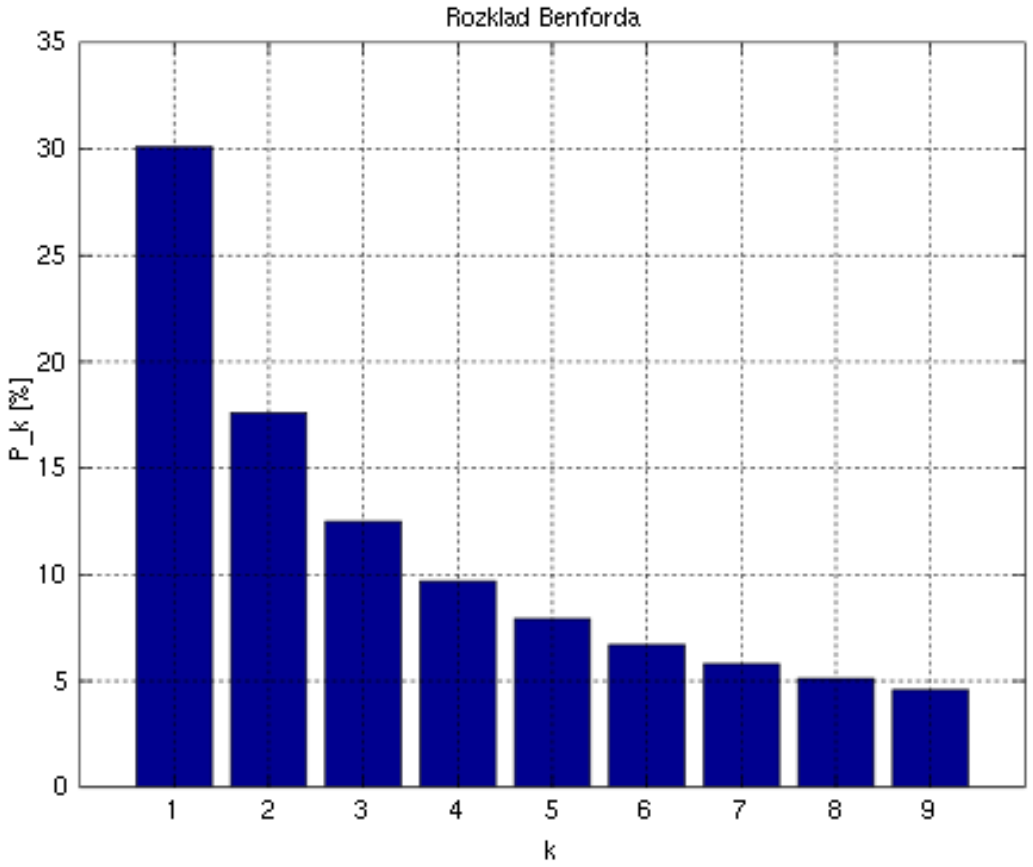
Steganalysis Tools

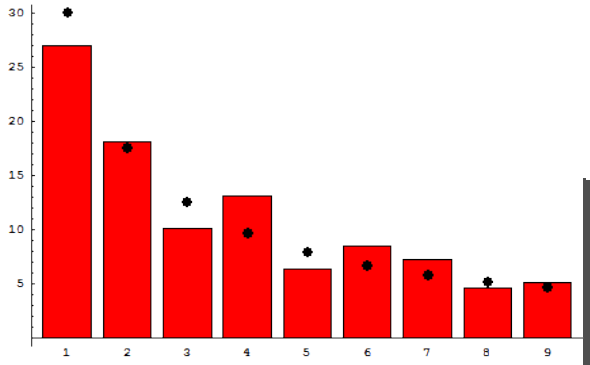
Hard Disk Steganographic Tools	Tools Analyzed	Detection Approach	Extraction Approach	Destruction Approach
2Mosaic	Removes stego content from any images			Break Apart
StirMark Benchmark	Removes stego content from any images			Resample
Phototile	Removes stego content from any images			Break Apart
Steganography Analyzer Real-Time Scanner	Analyzes Network Packets	Signature		
StegBreak	Jsteg-shell, JPhide, and Outguess 0.13b		Dictionary	
StegDetect	Jsteg, JPhide, Invisible Secrets, Outguess 01.3b, F5, appendX, Camouflage	Statistical		
StegSpy	Hiderman, JPHIde and Seek, Masker, JPegX, Invisible Secrets			
Stego-Suite	Detects Stego Image and Audio file		Dictionary	

Modern steganalytic methods using ML

- "Detection of Double-Compression in JPEG Images for Applications in Steganography (IEEE TIFS 2008)"
- Neural Networks (NN) and Support Vector Machines (SVM) utilized.
- Benford's Law was used among other features
- "Ensemble Classifiers for Steganalysis of Digital Media (IEEE TIFS 2012)"
- Accuracy: supervised classification with SVM
- Drawback: Long training steps, high complexity
- Ensemble classifiers here are implemented as random forests tested on nsF5, YASS, and MBS algorithms







The Benford's Law

(Images from Wikipedia)

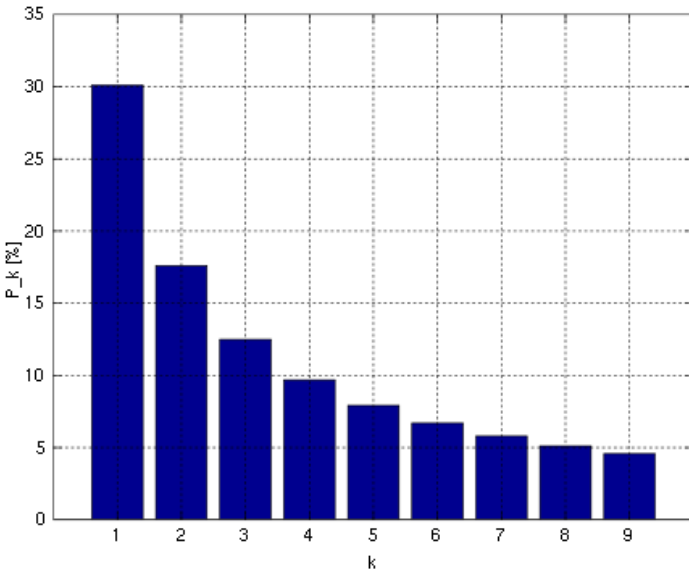
Benford's Law - historical facts

- 1881, Newcomb observed that the first pages of books with logarithmic tables, then heavily used for computation, were a lot more worn out than the last ones.
- Benford observed and abstracted formally this behaviour for random data sets around 1938.
  - Empirical law, a satisfactory explanation of which was provided by Hill (1996).
- This phenomenon can be observed and be of use in multiple domains and types of data sets.

Benford's Law – First Digit Law

- The leading digit n, n in {1, …, 9} in a uniformly and randomly distributed set of data has a probability of occurrence that can be expressed with the equation:

$$P(n) = \ell og_{10}\left(1 + \frac{1}{n}\right) \quad n = 1, \ldots, 9$$



The Law can be extended for other logarithmic bases. For b = 10 the following holds:

n	1	2	3	4	5	6	7	8	9
P(n)	30.1%	17.6%	12.5%	9.7%	7.9%	6.7%	5.8%	5.1%	4.6%

Lead digit distribution examples in natural data sets

col.	title	1	2	3	4	5	6	7	8	9	samples
A	Rivers, Area	31.0	16.4	10.7	11.3	7.2	8.6	5.5	4.2	5.1	335
B	Population	33.9	20.4	14.2	8.1	7.2	6.2	4.1	3.7	2.2	3259
C	Constants	41.3	14.4	4.8	8.6	10.6	5.8	1.0	2.9	10.6	104
D	Newspapers	30.0	18.0	12.0	10.0	8.0	6.0	6.0	5.0	5.0	100
E	Specific Heat	24.0	18.4	16.2	14.6	10.6	4.1	3.2	4.8	4.1	1389
F	Pressure	29.6	18.3	12.8	9.8	8.3	6.4	5.7	4.4	4.7	703
G	H.P. Lost	30.0	18.4	11.9	10.8	8.1	7.0	5.1	5.1	3.6	690
H	Mol. Wgt.	26.7	25.2	15.4	10.8	6.7	5.1	4.1	2.8	3.2	1800
I	Drainage	27.1	23.9	13.8	12.6	8.2	5.0	5.0	2.5	1.9	159
J	Atomic Wgt.	47.2	18.7	5.5	4.4	6.6	4.4	3.3	4.4	5.5	91
	Average	30.6	18.5	12.4	9.4	8.0	6.4	5.1	4.9	4.7	1011
	Probable Error	±0.8	±0.4	±0.4	±0.3	±0.2	±0.2	±0.2	±0.3		

Various applications of Benford's Law

- Hal Varian (1972) proposed its use for detecting fraud in socio-economic data reporting.
- Used widely to detect fraud in transactional data (e.g. Nigrini, 2000 and others), as implemented within audit packages (ACL, IDEA etc.).
- Acceptable in courts of law in the US.
- Used to analyse the 2009 election results in Iran to prove rigging.
- Limitation: The law may be true for a set of items but not for a certain subset of it.

Specific application of Benford's law for steganalysis

- Fu, Shi & Sub
  - examined the byte value distributions in the pixel domain (unsuccessful) as opposed to the Discrete Cosine Transform (DCT) values (that seems to obey Benford's law)
  - generalised the law to apply in detection of watermarked images
- We'll follow up this idea later (approach #2 in this set of slides)

The JPEG format (I)

- The jpeg standard specifies the way of coding and decoding an image. In other worlds, it defines the process of compressing the image into a byte stream and decompressing the byte stream back to form the image.
- The jpeg compression is lossy which means that, during the compression, there is information that will be lost but this will not dramatically affect the final result (depends on the compression rate we will use).

The JPEG format (II)

- The structure of a jpeg image follows the logic of continuous segments .
- Each segment begins with a marker which begins with an ‘FF’ (hexadecimal) byte followed by another byte which indicates the current marker.
- Common jpeg markers can indicate for example the start of the image (0xFF 0xD8), or the Huffman tables (0xFF 0xC4) and the end of image (0xFF 0xD9).

The JPEG format (An example)

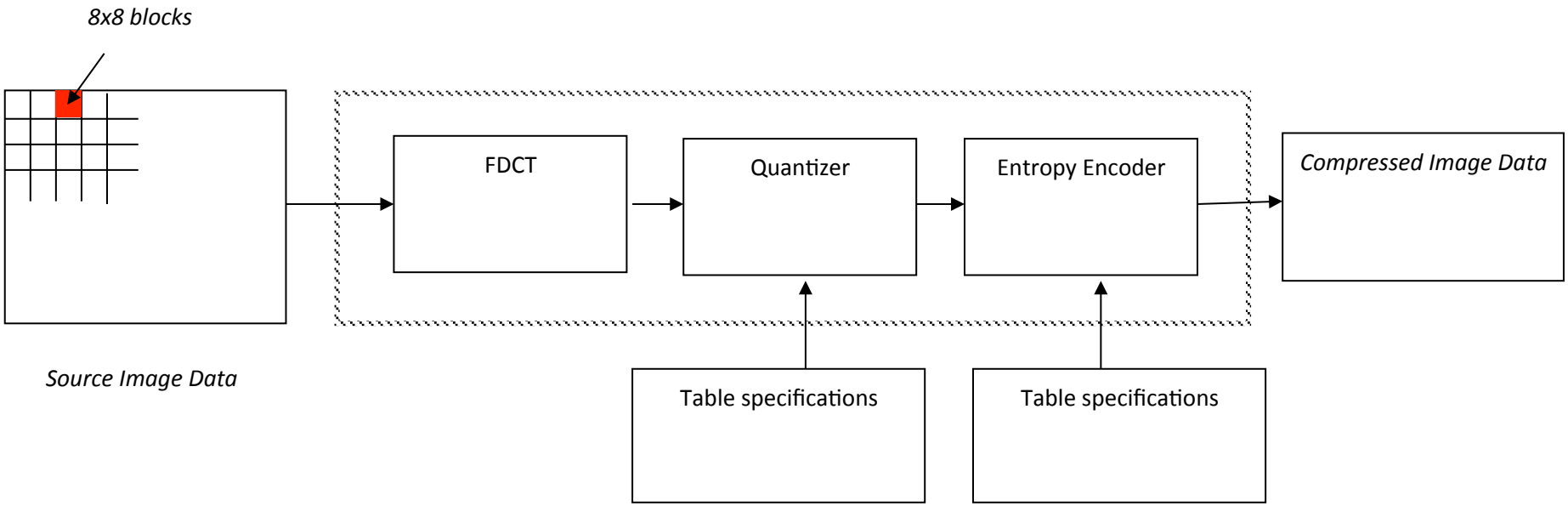
Using a HEX editor we can see how a JPEG image looks like

- **FF D8** FF E0 00 10 4A 46 49 46 00 01 01 01 01 2C 01 2C 00 00 FF E1 9A C0 45 78 69 66 00 00 4D 4D 00 2A 00 00 00 08 00 0B 01 0F 00 02 00 00 00 12 … continues (Start of image)
- 01 **FF C4** 01 A2 00 00 01 05 01 01 01 01 01 01 00 00 00 00 00 00 00 00 01 02 03 04 05 06 07 08 09 0A 0B 10 00 02 01 03 03 02 04 03 05 05 04 04 00 … continues (define Huffman tables)
- F6 F7 F8 F9 FA **FF DA** 00 0C 03 01 00 02 11 03 11 00 3F 00 F4 BC D3 AB 42 0A 33 C7 B9 78 EB 59 24 62 BA 22 CC 24 25 25 6A 64 14 94 00 57 35 AC 92 … continues (Start of scan)
- … 32 70 AA 30 4F 5C 75 34 51 51 2F 89 FF 00 5D 0E A9 2D 6C 7F **FF D9** (End of image)

JPEG Codec (JFIF encoding)

- Convert the representation of colours from RGB to YC_bC_r.
- Downsample the chrominance values (usually by a factor of two).
- Transform values to frequencies and use 8x8 pixel blocks.
- Quantization process.
- Zigzag ordering.
- Lossless compression using a variant of Huffman encoding.

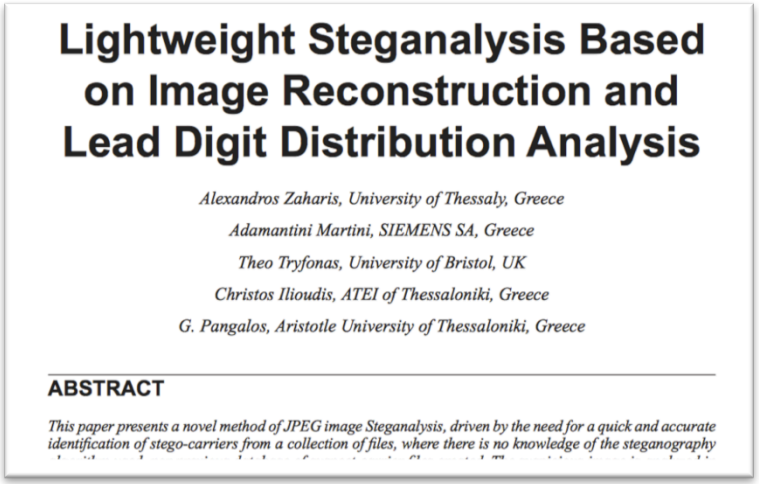
A simplified view of the DCT encoding procedure



*A detailed example will follow during the dissection of our second steganalytic approach.

Our lightweight blind steganalytic approaches for JPEG images

- "Lightweight Steganalysis Based on Image Reconstruction and Lead Digit Distribution Analysis, (IJDCF 2011)"
- "JPEG steganography detection with Benford's Law, (Digital Investigation 2013)"

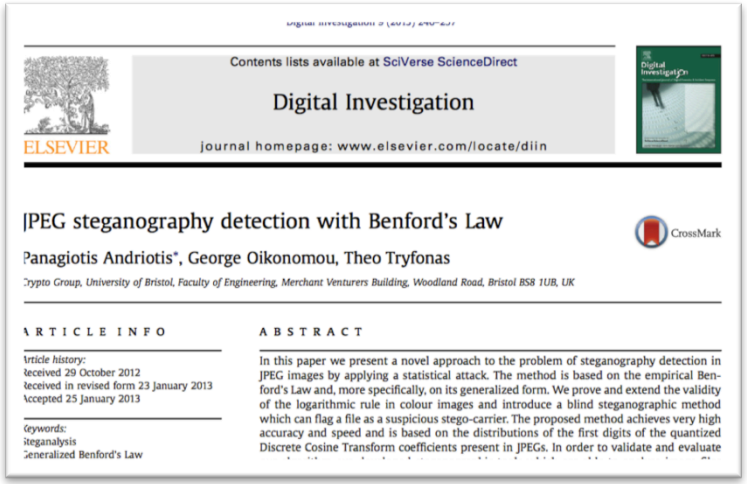


Lightweight Steganalysis Based on Image Reconstruction and Lead Digit Distribution Analysis

Alexandros Zaharis, University of Thessaly, Greece
Adamantini Martini, SIEMENS SA, Greece
Theo Tryfonas, University of Bristol, UK
Christos Ilioudis, ATEI of Thessaloniki, Greece
G. Pangalos, Aristotle University of Thessaloniki, Greece

ABSTRACT

This paper presents a novel method of JPEG image Steganalysis, driven by the need for a quick and accurate identification of stego-carriers from a collection of files, where there is no knowledge of the steganography



Contents lists available at SciVerse ScienceDirect

Digital Investigation

journal homepage: www.elsevier.com/locate/diin

JPEG steganography detection with Benford's Law

Panagiotis Andriotis*, George Oikonomou, Theo Tryfonas

Crypto Group, University of Bristol, Faculty of Engineering, Merchant Venturers Building, Woodland Road, Bristol BS8 1UB, UK

ARTICLE INFO

Article history:
Received 29 October 2012
Received in revised form 23 January 2013
Accepted 25 January 2013

Keywords:
Steganalysis
Generalized Benford's Law

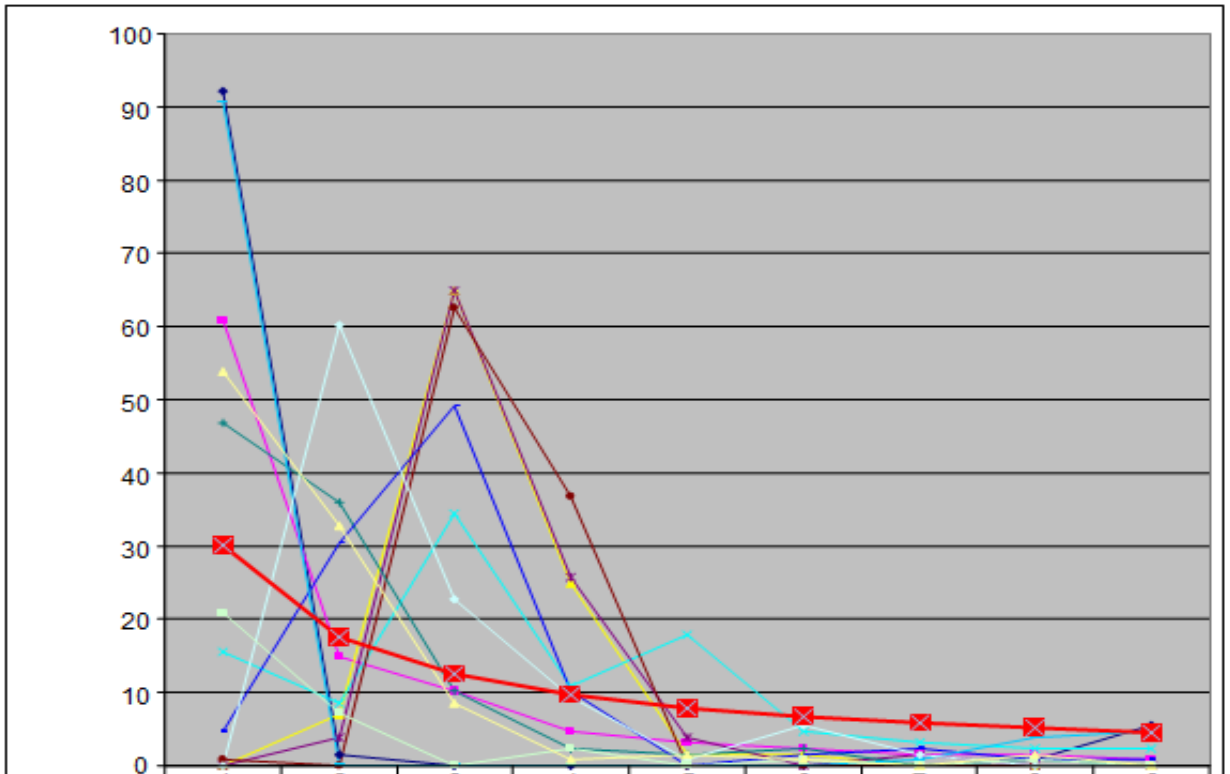
ABSTRACT

In this paper we present a novel approach to the problem of steganography detection in JPEG images by applying a statistical attack. The method is based on the empirical Benford's Law and, more specifically, on its generalized form. We prove and extend the validity of the logarithmic rule in colour images and introduce a blind steganographic method which can flag a file as a suspicious stego-carrier. The proposed method achieves very high accuracy and speed and is based on the distributions of the first digits of the quantized Discrete Cosine Transform coefficients present in JPEGs. In order to validate and evaluate

The first approach

- Combine the use of Benford's Law for detection of file anomalies as previously, but on byte values (not pixels or DCT output)
  - Work from Karresand (2006) on byte value (eventually pairs) distribution in detection of image file format (and camera make) and
  - Work from Haggerty (2007) on file fingerprinting by byte value

First digit appearance using Byte array representation of common digital files

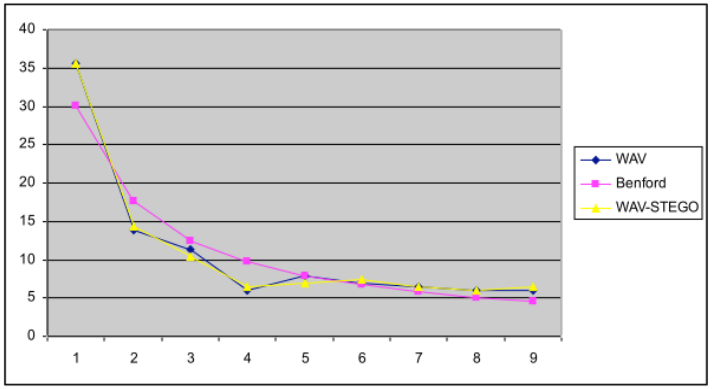


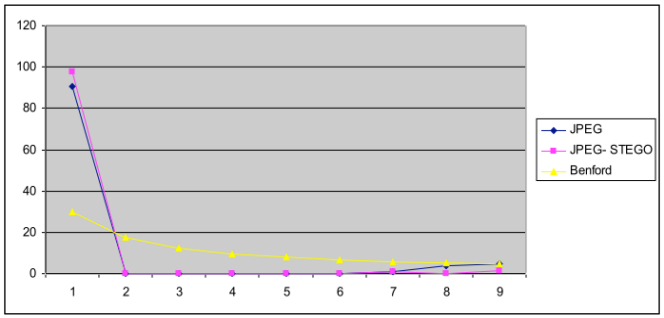
	1	2	3	4	5	6	7	8	9
WMV	92,19	1,56	0	0	0	0	0	0,78	5,47
WAV	61	15	10	5	3	2	2	2	1
PNG	0	7	65	25	2	2	0	0	0
BMP	16	9	34	11	18	5	3	2	2
MP3	0	4	65	26	4	0	2	0	0
ZIP	1	0	63	37	0	0	0	0	0
EXE	47	36	10	2	2	2	0	1	0
PDF	5	30	49	10	0	2	2	1	1
JPEG	91	0	0	0	0	0	1	4	5
GIF	0	60	23	9	1	5	2	0	0
RAR	21	7	0	2	0	1	0	1	0
DOC	54	33	9	1	2	1	0	2	0
Benford	30	18	12	10	8	7	6	5	5

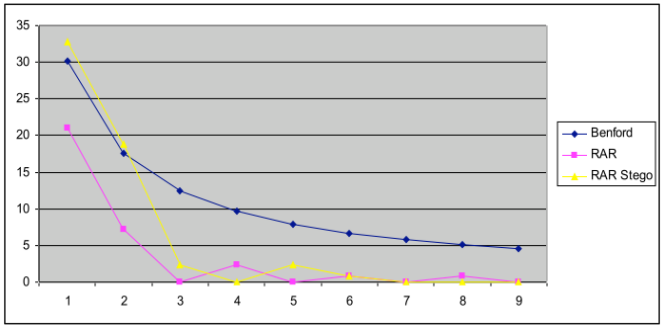
Steganography and alterations of file structure

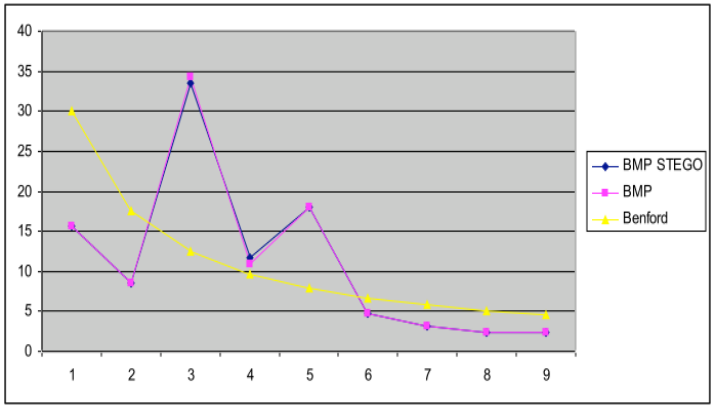
- We observed that the byte array representation's distribution was affected, in relation to the one of the original file types.
- Interestingly:
  - This was measurable for small size input secret files.
  - Increased with the size of the secret file.
  - It was detectable with no dependency of the type of stego algorithm used.

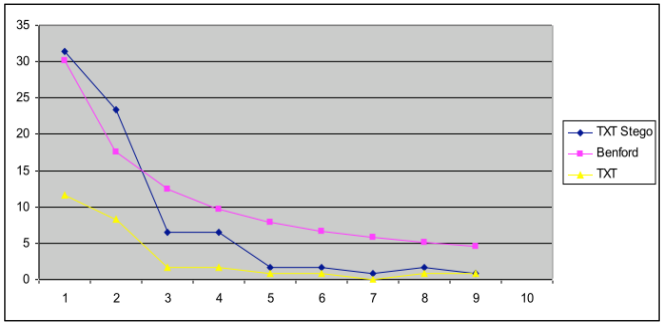
Variation of file structure before and after the application of steganography (payload <= 1kb)









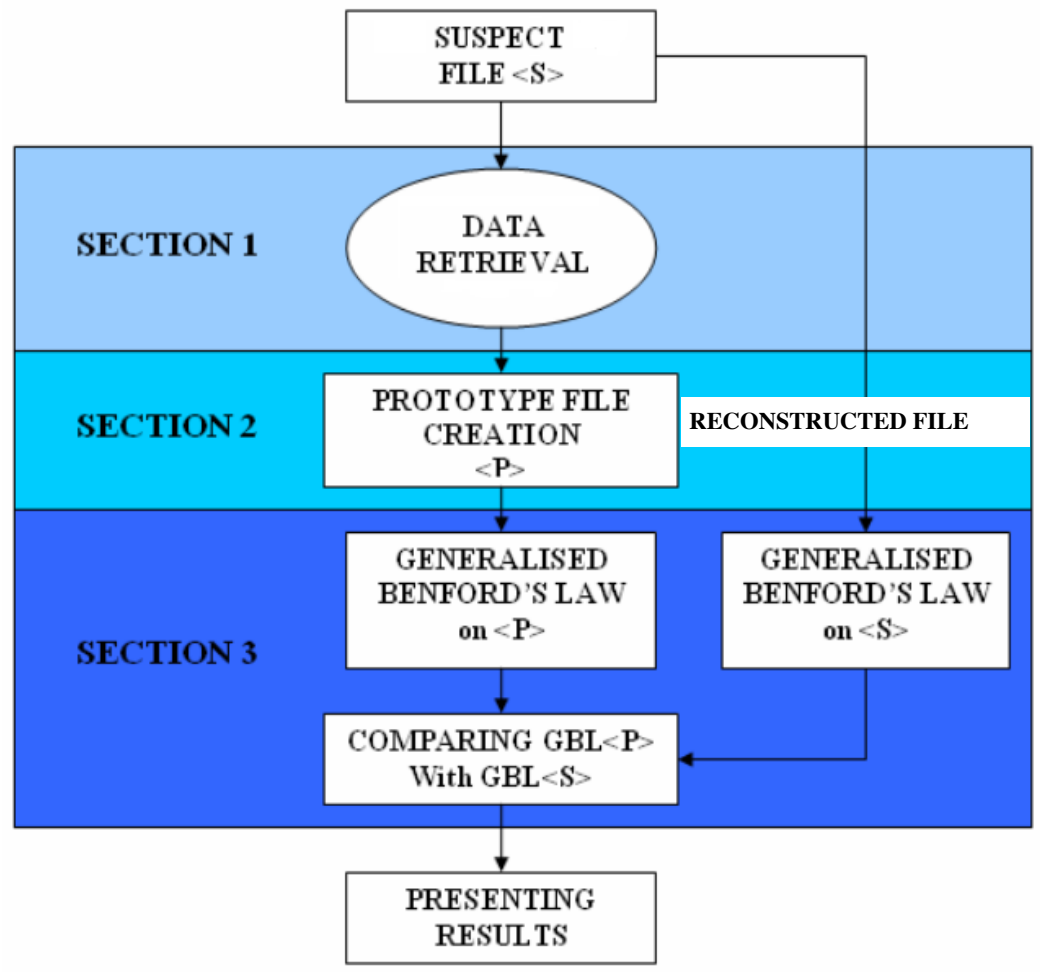


A key idea: Cover file generic reconstruction

- Generic reconstruction is a process whereby a file with similar properties to the original one is reconstructed from the stegocarrier.

- Properties refer to:
  - Image quality.
  - File structure.
  - Content.
- Procedures that may change those could be :
  - Format alteration.
  - Copying reproduction (e.g. JPEG).
  - Use of stego tools.

Steganalysis method and proof of concept (for JPEG/MS Paint): Ben-4D

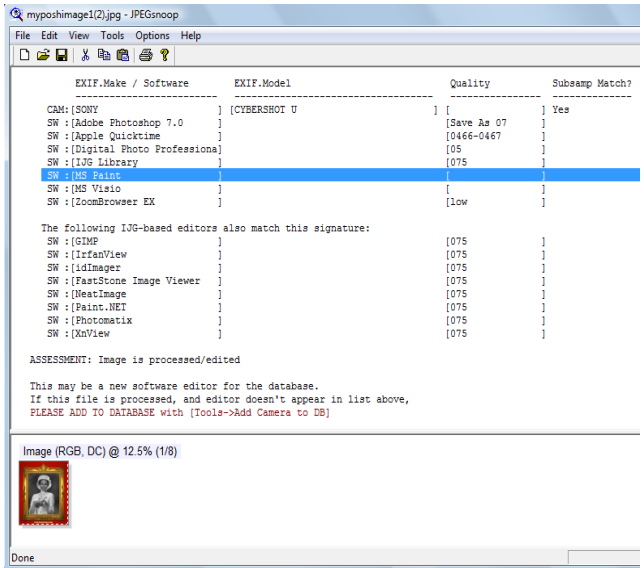


Similarity Threshold

- Predefined constant value identified experimentally after applying Generalised Benford's Law on large numbers of reconstructed files.
- This value is encoding-specific, so MS Paint has a certain Similarity Threshold while Photoshop 9 has a different one.

File reconstruction

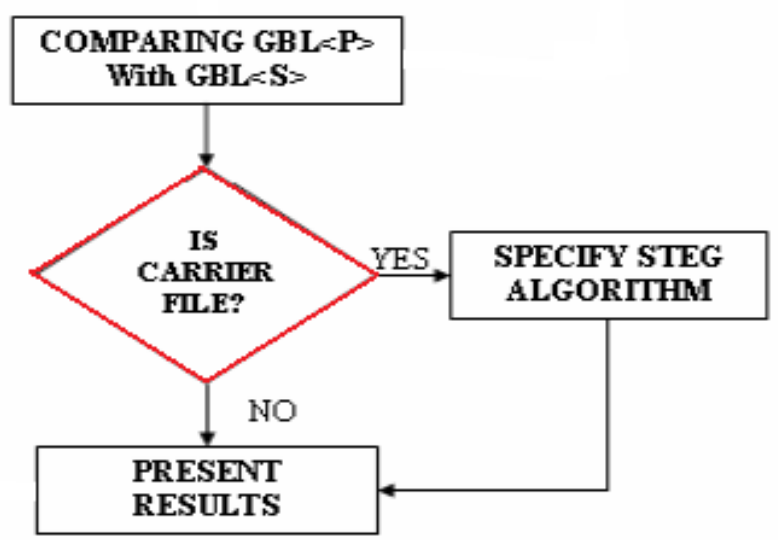
```
Precision=8 bits
Destination ID=0 (Luminance)
  DQT, Row #0:   8   6   5   8  12  20  26  31
  DQT, Row #1:   6   6   7  10  13  29  30  28
  DQT, Row #2:   7   7   8  12  20  29  35  28
  DQT, Row #3:   7   9  11  15  26  44  40  31
  DQT, Row #4:   9  11  19  28  34  55  52  39
  DQT, Row #5:  12  18  28  32  41  52  57  46
  DQT, Row #6:  25  32  39  44  52  61  60  51
  DQT, Row #7:  36  46  48  49  56  50  52  50
  Approx quality factor = 74.75 (scaling=50.51 variance=0.81)

Precision=8 bits
Destination ID=1 (Chrominance)
  DQT, Row #0:   9   9  12  24  50  50  50  50
  DQT, Row #1:   9  11  13  33  50  50  50  50
  DQT, Row #2:  12  13  28  50  50  50  50  50
  DQT, Row #3:  24  33  50  50  50  50  50  50
  DQT, Row #4:  50  50  50  50  50  50  50  50
  DQT, Row #5:  50  50  50  50  50  50  50  50
  DQT, Row #6:  50  50  50  50  50  50  50  50
  DQT, Row #7:  50  50  50  50  50  50  50  50
  Approx quality factor = 74.74 (scaling=50.52 variance=0.19)
```



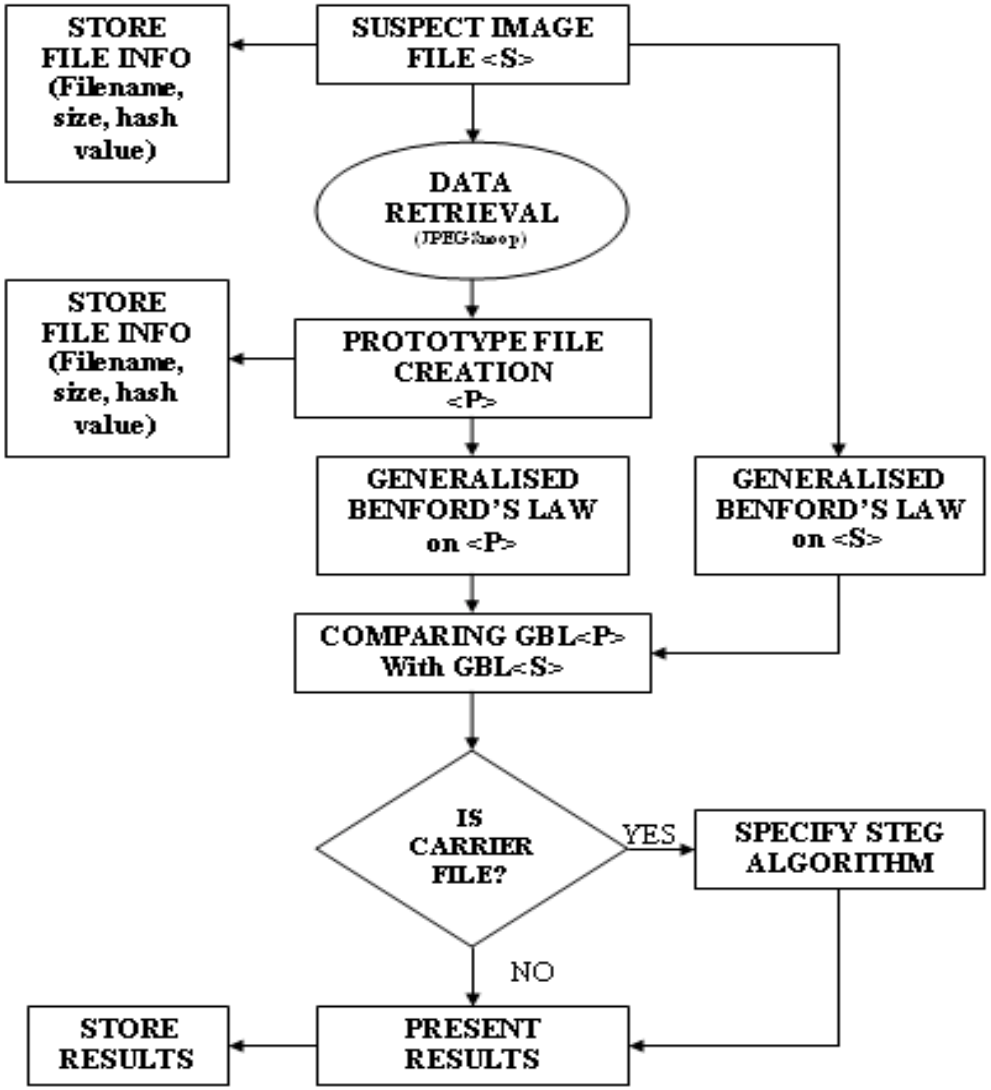
Improvement of detection rate by considering stego tools features

- Signatures/rules for the intended stego tool recognition:
  - Atypical or corrupted Huffman tables (JPHSWin).
  - Significant size difference of stegocarrier and reconstructed file (Camouflage, Invisible Secrets).
  - Specific headers manipulation (Invisible Secrets).
  - Issues with file termination (Camouflage).

- Embedding these rules into the detection method leads to improvement of the False – Positive detection rate from 15% to 0.1%.



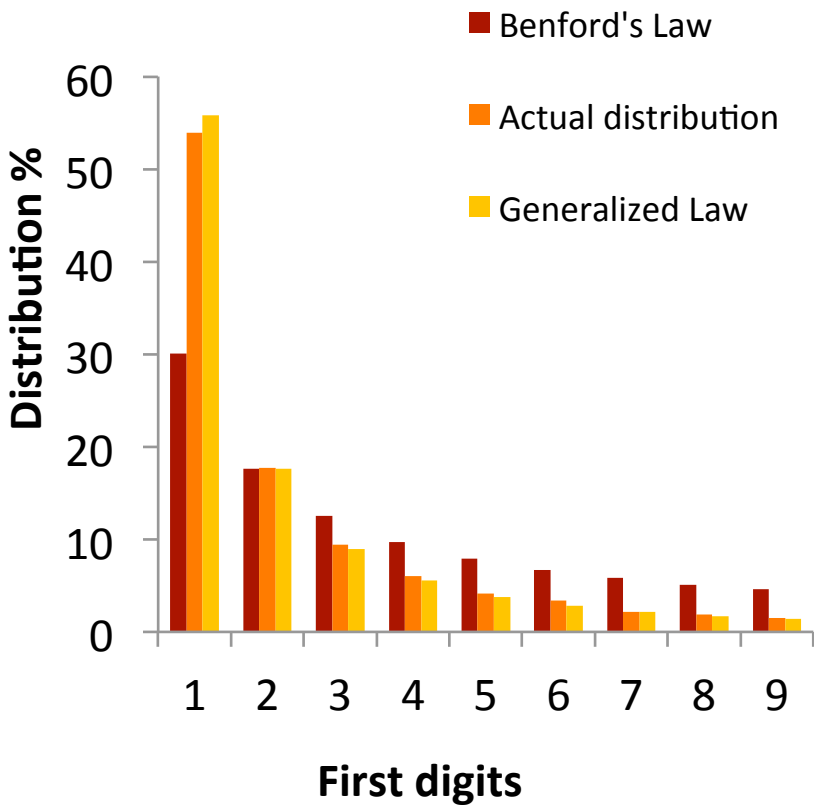
Flowchart of the proposed forensic tool



Another approach

- Fu et al. worked on the distribution of first digits of DCT coefficients, but only on the luminance component of pictures. They only used grey scale JPEG images in their study.
- We extended their work to chrominance and applied it comprehensively. We investigated if their model (generalised Benford's Law: gBL) applies to colour JPEG images and examined the artifacts various steganographic algorithms leave on the DCT coefficients space.

The basic concept (Fu et al.)



Generalized Benford's Law for grey scale JPEG images (Fu et al.)

$$p(n) = N \cdot \log_{10}\left(1 + \frac{1}{s + n^q}\right), \quad \mathrm{n} = 1, 2, \ldots, 9$$

Quality Factor	*Model Parameters*			*Goodness-of-fit (SSE)*
	N	*q*	*s*	
100	1.456	1.47	0.0372	7.104e-06
90	1.255	1.563	-0.3784	5.255e-07
80	1.324	1.653	-0.3739	3.06838e-06
70	1.412	1.732	-0.337	5.36171e-06
60	1.501	1.813	-0.3025	6.11167e-06
50	1.579	1.882	-0.2725	6.05446e-06

JPEG Compression example (I)

- The image consists of pixels, each pixel has usually three bytes that represent the three basic colours Red, Green, Blue (RGB).
- Convert these pixel values from RGB to YC_bC_r which is another colour space that has three components. Y represents the brightness of an image and is called luminance while C_b and C_r represent colours and they are called chrominance. It is known that the human eye can recognize the difference in the luminance of an image more easily than the chrominance coefficients.
- Chroma subsampling, reduction of the chrominance coefficients by a factor of two.
- The next phase after downsampling (or subsampling) is the division of each of the channels (Y, C_b, C_r) to 8x8 blocks.
- Each of these blocks is then converted to a frequency domain representation using a transformation which is the type-II DCT (Discrete Cosine Transform).

$$G_{u,\upsilon} = \sum_{x=0}^{7}\sum_{y=0}^{7}\alpha(u)\alpha(\upsilon)g_{x,y}\cos\left[\frac{\pi}{8}\left(x+\frac{1}{2}\right)u\right]\cos\left[\frac{\pi}{8}\left(y+\frac{1}{2}\right)\upsilon\right]$$

where u is the spatial frequency (horizontally) for the integers $0 \le u < 8$ and

υ is the spatial frequency (vertically) for the integers $0 \le \upsilon < 8$, such that

$$\alpha(u) = \begin{cases} \frac{1}{\sqrt{8}} & u = 0 \\ \frac{1}{2} & u \neq 0 \end{cases} \quad \text{and} \quad \alpha(\upsilon) = \begin{cases} \frac{1}{\sqrt{8}} & \upsilon = 0 \\ \frac{1}{2} & \upsilon \neq 0 \end{cases}.$$

$g_{x,y}$ is the value of the pixel located at position (x, y) and

$G_{u,\upsilon}$ the DCT coefficient at position (u, υ).

JPEG Compression example (II)

Let A be a 8x8 block as follows

$$\begin{bmatrix} 154 & 123 & 123 & 123 & 123 & 123 & 123 & 136 \\ 192 & 180 & 136 & 154 & 154 & 154 & 136 & 110 \\ 254 & 198 & 154 & 154 & 180 & 154 & 123 & 123 \\ 239 & 180 & 136 & 180 & 180 & 166 & 123 & 123 \\ 180 & 154 & 136 & 167 & 166 & 149 & 136 & 136 \\ 128 & 136 & 123 & 136 & 154 & 180 & 198 & 154 \\ 123 & 105 & 110 & 149 & 136 & 136 & 180 & 166 \\ 110 & 136 & 123 & 123 & 123 & 136 & 154 & 136 \end{bmatrix}$$

- In order to compute the DCT coefficients we must substitute those values with new ones that are centered on zero.
- For this reason if the values are in the range of [0, 255] we will subtract them from 128 which is the mid-point of this range; this achieves a range between [-128, 127].
- The new matrix now is matrix M (as seen above).

$$\begin{bmatrix} 26 & -5 & -5 & -5 & -5 & -5 & -5 & 8 \\ 64 & 52 & 8 & 26 & 26 & 26 & 8 & -18 \\ 126 & 70 & 26 & 26 & 52 & 26 & -5 & -5 \\ 111 & 52 & 8 & 52 & 52 & 38 & -5 & -5 \\ 52 & 26 & 8 & 39 & 38 & 21 & 8 & 8 \\ 0 & 8 & -5 & 8 & 26 & 52 & 70 & 26 \\ -5 & -23 & -18 & 21 & 8 & 8 & 52 & 38 \\ -18 & 8 & -5 & -5 & -5 & 8 & 26 & 8 \end{bmatrix}$$

JPEG Compression example (III)

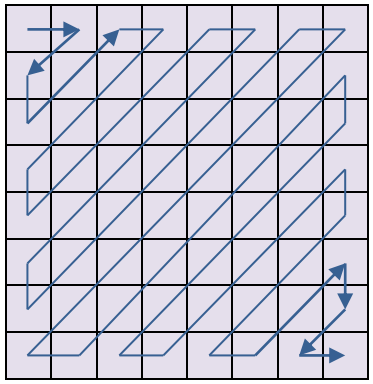
After DCT application:

$$D = \begin{bmatrix} 162.3 & 40.6 & 20.0 & 72.3 & 30.3 & 12.5 & -19.7 & -11.5 \\ 30.5 & 108.4 & 10.5 & 32.3 & 27.7 & -15.5 & 18.4 & -2.0 \\ -94.1 & -60.1 & 12.3 & -43.4 & -31.3 & 6.1 & -3.3 & 7.1 \\ -38.6 & -83.4 & -5.4 & -22.2 & -13.5 & 15.5 & -1.3 & 3.5 \\ -31.3 & 19.7 & -5.5 & -12.4 & 14.3 & -6.0 & 11.5 & -6.0 \\ -0.9 & -11.8 & 12.8 & 0.2 & 28.1 & 12.6 & 8.4 & 2.9 \\ 4.6 & -2.4 & 12.2 & 6.6 & -18.7 & -12.8 & 7.7 & 12.0 \\ -10.0 & 11.2 & 7.8 & -16.3 & 21.5 & 0.0 & 5.9 & 10.7 \end{bmatrix}$$

For example, the element in place (0, 0) upper-left side will be calculated by equation

$$K(0,0) = \text{integer round}\frac{D(0,0)}{Q(0,0)} = \text{integer round}\frac{162,3}{16} = 10.$$

$$K = \begin{bmatrix} 10 & 4 & 2 & 5 & 1 & 0 & 0 & 0 \\ 3 & 9 & 1 & 2 & 1 & 0 & 0 & 0 \\ -7 & -5 & 1 & -2 & -1 & 0 & 0 & 0 \\ -3 & -5 & 0 & -1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$



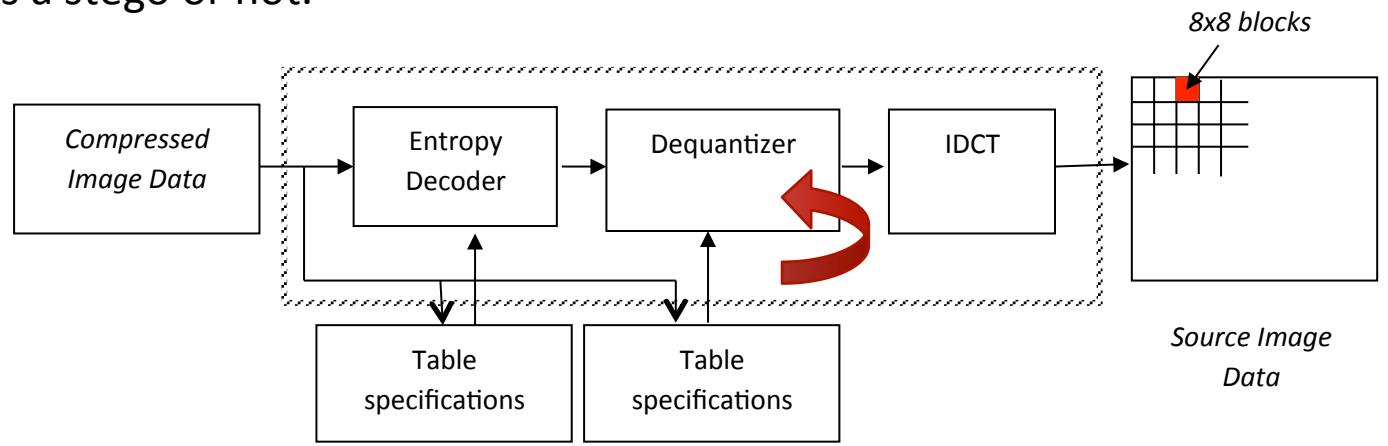
- The top left coefficient has the largest magnitude and it is called DC coefficient. The other 63 entries are called AC coefficients. Only AC coefficients will be used for steganalysis.
- The following phase is the quantization step. This is the lossy part of the compression.

$$Q_{50} = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

$$K(u,v) = \text{integer round}\frac{D(u,v)}{Q(u,v)}$$

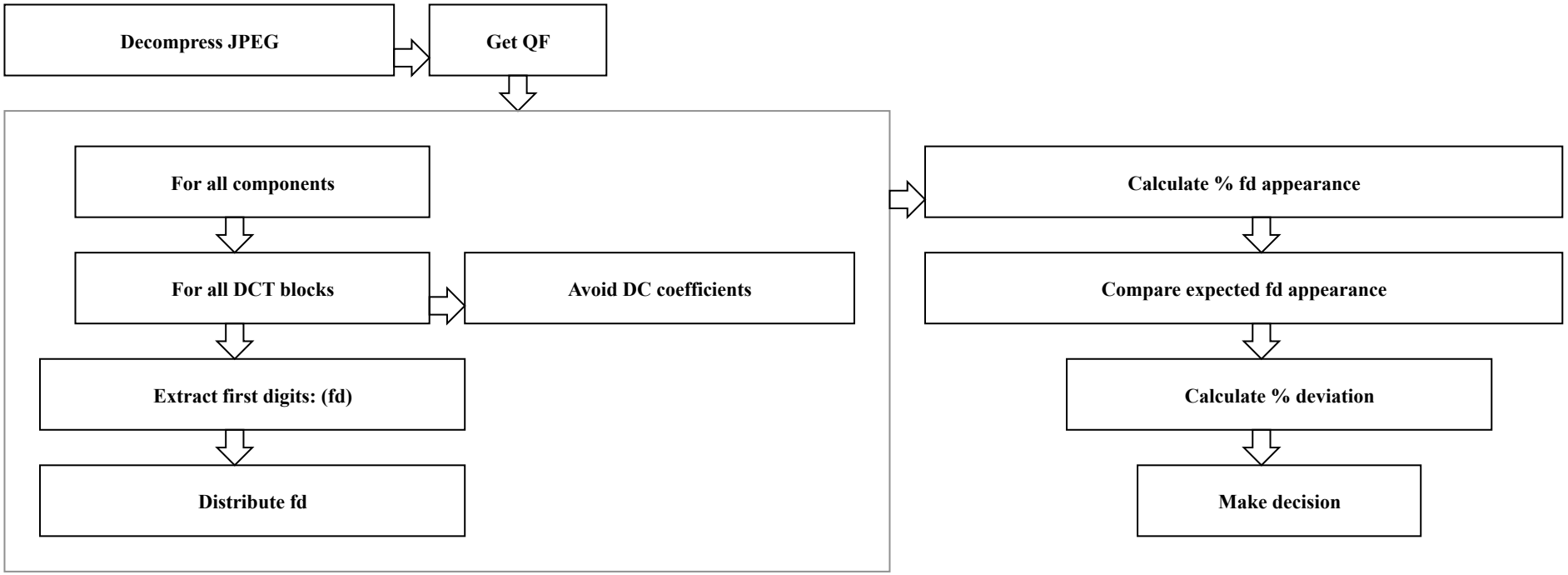
StegBennie Algorithm

- After decompressing the image we read the metadata and find the compression quality factor.
- We are looking at the DCT blocks (8x8) that constitute the image and extract the first digit of each coefficient. For example, if the first row of an 8x8 block of coefficients is [211 22 12 6 1 0 0 0], the first digits are [x 2 1 6 1 x x x] (211 is the DC coefficient and it is excluded and also the zeros are not taken into consideration).
- We calculate the % percentage of appearance of each leading digit. Then we estimate the first digits expected distribution and finally compare the deviations between the expected and the calculated distributions.
- Any deviation between the expected and the estimated distributions will help to decide if the image is a stego or not.



StegBennie Algorithm

Flowchart



Quantised DCT coefficient-based analysis

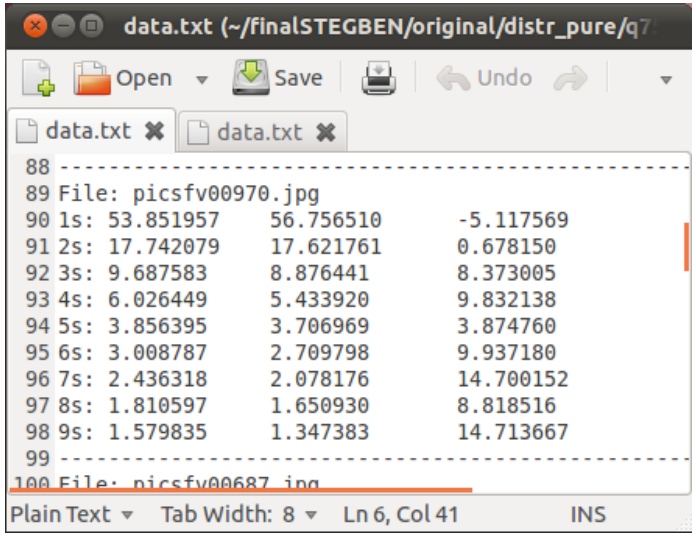
- Using the gBL with different N, q, s for colour images

$$p(x) = N \cdot \ell og_{10}\left(1 + \frac{1}{s + x^{q}}\right)$$

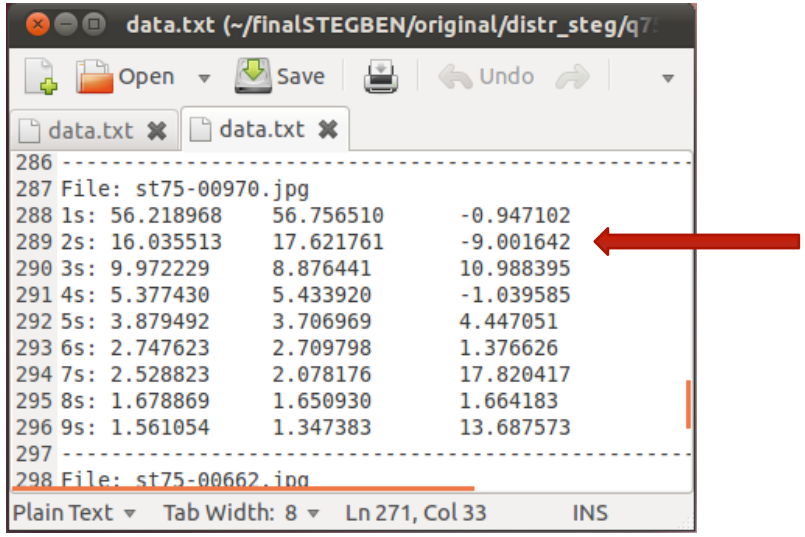
Quality Factor	*Model Parameters*			*Goodness-of-fit (SSE)*
	N	*q*	*s*	
100	1.608	1.605	0.0702	5.129e-06
90	1.25	1.585	-0.405	7.235e-07
80	1.344	1.685	-0.376	3.007e-06
75	1.396	1.731	-0.3549	3.986e-06
70	1.434	1.766	-0.339	4.455e-06
60	1.514	1.843	-0.3114	5.464e-06
50	1.584	1.909	-0.2875	5.119e-06

Classification of images

- Differences in deviations from the expected distributions
- Pure image

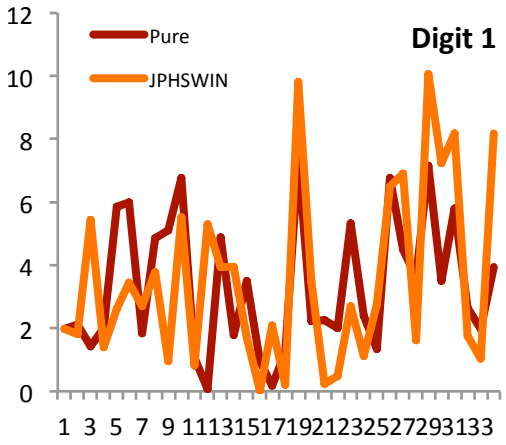


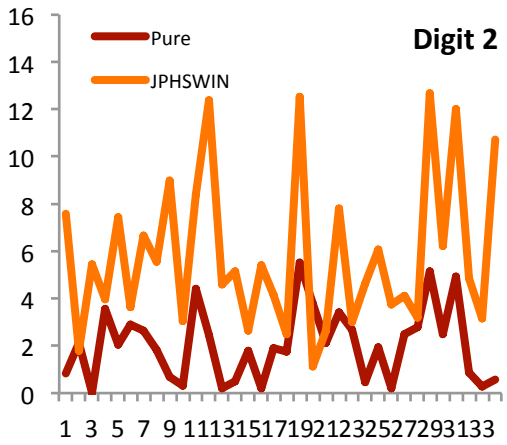
Stego image

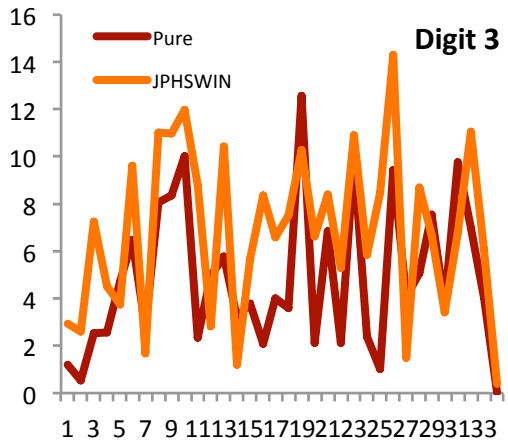


Classification model

- Classification Threshold

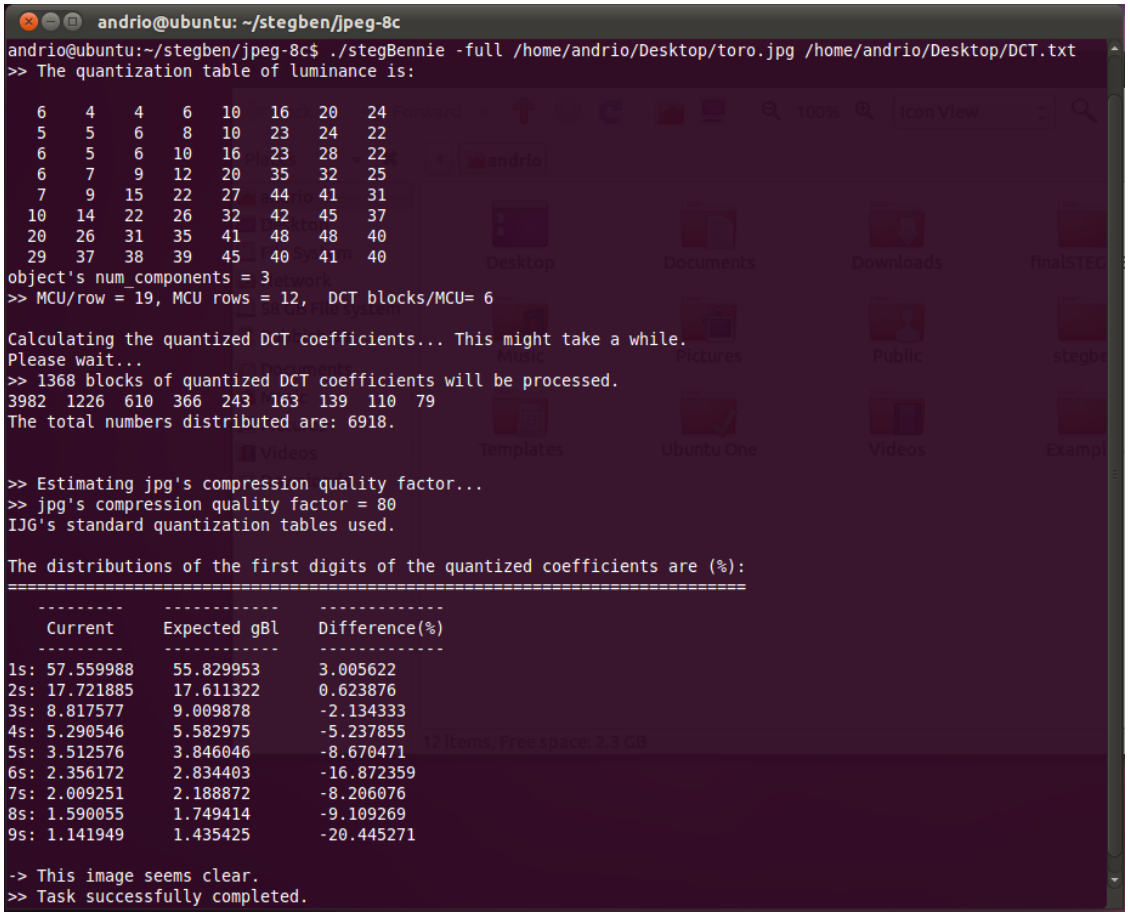






- Other digit deviations behave like digit 3 deviations

StegBennie in action



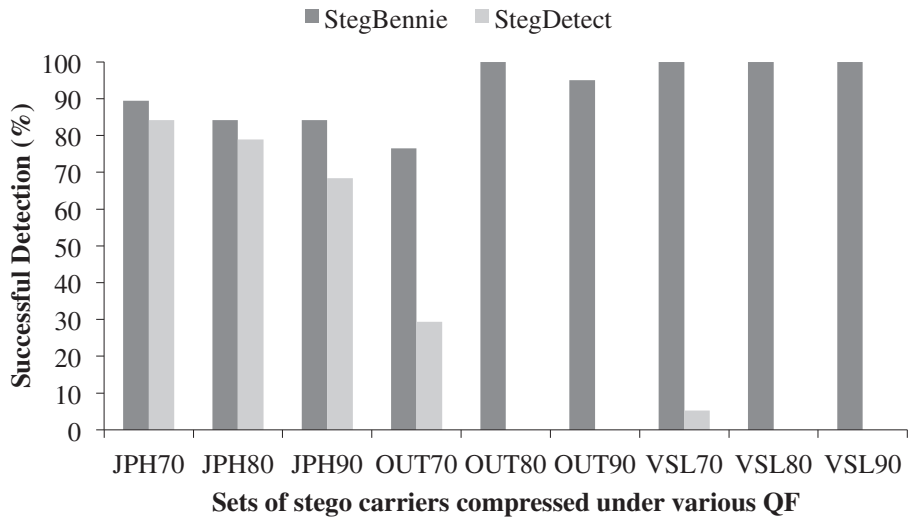
Results and comparison against other steganalytic tools

FPR on real data.

QF	Resolution	Examined images	FPR (%)
Normal QF = 70	small	9	11.11
	1Mp	10	0
	3Mp	9	11.11
	5Mp	8	12.5
	wide1Mp	9	0
High QF = 80	small	10	10.0
	1Mp	9	11.11
	3Mp	10	20.0
	5Mp	8	37.5
	wide1Mp	10	0
Fine QF = 90	small	10	30
	1Mp	19	15.79
	3Mp	9	44.44
	5Mp	10	30.0
	wide1Mp	11	27.27

Hit rates for real data.

QF	Resolution	JPHSWIN	Outguess	Vsl
Normal QF = 70	small	88.89	77.78	100.0
	1Mp	90.0	75.0	100.0
	3Mp	55.55	75.0	100.0
	5Mp	33.33	87.5	100.0
	wide1Mp	66.67	50.0	100.0
High QF = 80	small	100.0	100.0	100.0
	1Mp	66.67	100.0	100.0
	3Mp	50.0	100.0	100.0
	5Mp	50.0	71.43	100.0
	wide1Mp	60.0	100.0	100.0
Fine QF = 90	small	100.0	100.0	100.0
	1Mp	66.67	90.0	100.0
	3Mp	55.55	87.5	100.0
	5Mp	40.0	100.0	100.0
	wide1Mp	72.73	90.0	100.0



Further work - Ben-4D

- Minimization of data loss during reconstruction (use of lossless transcoding).
- Support for detection of more steganography tools.
- Other types JPEG coding.
- Support for other popular image formats (BMP, GIF).
- Message extraction would also be desirable.
- http://sourceforge.net/projects/ben4dstegdetect/

Further work - StegBennie

- Consider the effect of the size of the embedded data and measure its impact on the overall validity of the method.
- Fu et al. (2007) also investigated the distributions of first digits of the coefficients of the blocks of the JPEG images before the quantization step (during the compression of the image). These adhere to the original Benford's Law quite well. Future development should consider this (it will probably provide the opportunity to ascertain the deviations of the distributions of the first digits of the block coefficients before quantization) of the JPEG image.
- Open source dissemination via ForToo website: www.fortoo.eu

Sources

- Andriotis, P., Tryfonas, T., Oikonomou, G., Spyridopoulos, T., Zaharis, A., Martini, A. and Askoxylakis, I. (2013), "On Two Different Methods for Steganography Detection in JPEG Images with Benford's Law", Proc. of 7th Intl. Scientific Conf. on Security and Protection of Information (SPI 2013)

- Panagiotis Andriotis, George Oikonomou, Theo Tryfonas.
  JPEG Steganography Detection with Benford's Law.
  Digital Investigation, Vol. 9, pp. 246-257, 2013.

- A Zaharis, A Martini, T Tryfonas, C Ilioudis, G Pangalos.
  Lightweight Steganalysis based on Image Reconstruction & Lead Digit Distribution Analysis.
  International Journal of Digital Crime and Forensics, Vol. 3, pp. 29-41, 2011.

- A Zaharis, A Martini, T Tryfonas, C Ilioudis, G Pangalos.
  Reconstructive Steganalysis by Source Bytes Lead Digit Distribution Examination.
  Digital Forensics and Incident Analysis - WDFIA 2011, pp. 55-68, 2011.

Thank you!

www.andrio.eu

This work has been supported by the European Union's Prevention of and Fight against Crime Programme "Illegal Use of Internet" - ISEC 2010 Action Grants, grant ref. HOME/2010/ISEC/AG/INT-002 and the Systems Centre of the University of Bristol.