csafe
Center for Statistics and Applications in Forensic Evidence

Can stego images from a mobile phone stego app be detected?

PI: Jennifer Newman*

Other collaborators: Yong Guan**, Min Wu#

Graduate RAs: Li Lin*, Wenhao Chen**, Stephanie Reinders*

*Department of Mathematics

**Department of Electrical and Computer Engineering

Iowa State University, Ames IA

#Department of Electrical & Computer Engineering

University of Maryland, College Park, MD

Today's presentation

- Present errors in steg detection arising from four different software packages to mobile phone app PixelKnot
- There are no published results of steg detection on app data from mobile phones (that we could find)
- This is an important topic
  - Number of mobile phone apps performing steganography has increased in the past 5-7 years
  - Gives forensic practitioners a preliminary benchmark of some steg detection software packages
  - More practical scenario as it uses mobile phone pictures and apps
  - Recent: Australian government proposes to compel technology companies to provide access to users' messages, regardless of whether they have been encrypted (Guardian)

Four software packages

- Stego Hunt (Wetstone, commercial software)
- StegDetect from DC3 (courtesy of Bill Eber, DC3)
- StegDetect from Dr. N. Provos (available on Internet)
- Academic pattern classifier (standard S.O.T.A. scenario)

- Explain process of detection by each method
- Discuss results, and where and why errors occur
- Encourage questions, feedback, comments

What is steganography?

- *Steganography* is the hiding of a message – "payload" – in an image

Cover

+

Payload

=

Stego

Steganography: "covered writing" (Greek)

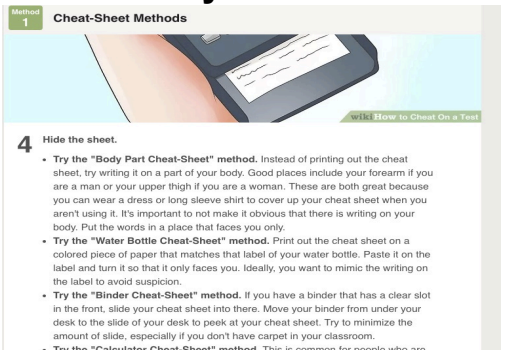
What is steganalysis?

- *Steganalysis* is a *digital image forensic* tool to determine if an image has been altered to contain a hidden message

Cover

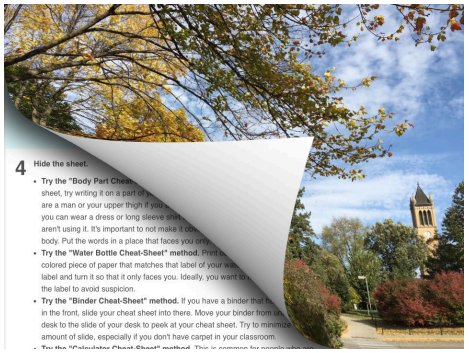
+

Payload



=

Stego



What is PixelKnot?

PixelKnot: Hidden Messages

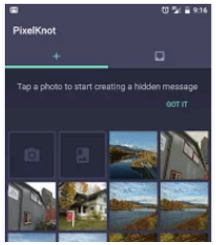
The Guardian Project Communication 866

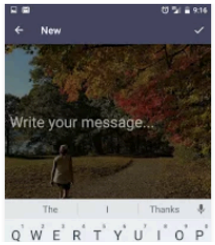
Everyone

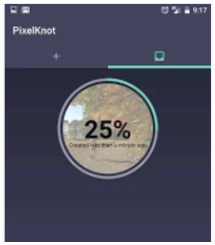
Contains ads

You don't have any devices

Add to Wishlist Install







REVIEWS

3.8

866 total

★ 5 451
★ 4 127
★ 3 102
★ 2 46
★ 1 140

Helpfulness

Homer Slated August 4, 2013

Perfect To those who can't seem to figure out how to decrypt files, try "open as..." then choose PixelKnot, which is registered as a mime type handler for images (doh!). Works for me, and no crashes. :-)

Adamant Orcinus January 17, 2014

Bout time I have been sharing hidden files in public view for years. Mp3s as jpegs and soft ware as .gifs or png's. I'm glad to see others in the covert tip. Good job.

ADDITIONAL INFORMATION

Updated	Installs	Current Version
February 17, 2017	100,000 - 500,000	1.0.1
Requires Android	**Content Rating**	**Permissions**
4.2 and up	Everyone Learn more	View details
Report	**Offered By**	
Flag as inappropriate	The Guardian Project	

Developer
Visit website
Email support@guardianproject.info

- Android app
- Uses sophisticated embedded algorithm (F5)
- Downloaded many times
- Higher rating than the "Will You Marry Me?" app (2.7/5)

Steg Detection using Stego Hunt

- Stego Hunt is commercial software by WetStone
  - “leading software tool for discovering the presence of data hiding activities”
  - ”generate case specific reports for management or court presentation”
  - “identify suspect carrier files: program artifacts, program signatures, statistical anomalies” using hash tables, file signatures, and statistics
- Has 10 possible detection responses for a given scanned file

Steg Detection using DC3's StegDetect

- DC3 StegDetect: courtesy of Bill Eber, DoD Cyber Crime Center (DC3)
- A software program that can be applied to many different types of files
  - We applied only to image files

Steganalysis using Provos' StegDetect

- This StegDetect is NOT same as DC3's StegDetect
- Completely different computer program, developed by Dr. Neil Provos (now at Google)
- Can only accept jpg file images as input to scan
- Detects stego images output by steg embedding programs:
  - jsteg, jphide, and outguess 0.13b (all stegos are jpeg file format)
- Identifies the most likely program used to embed, if detected as stego

Benchmarking Summary Table

				Results of passing through:		
File Type	#	Cover or Stego	Embedding algorithm	Stego Hunt	DC3 StegDetect	Provos StegDetect
PNG Images	2090	Cover	(none)	1304 Carrier Anomalies[1]	0 suspicious	N/A[2]
JPG Images	1606	Cover	(none)	0 anomalies	0 suspicious	380/1606 = 24%
JPG Images	4818	Stego	PixelKnot	0 anomalies	0 suspicious	1160/4818= 24%
JPG Images	421	Stego	F5 exec.[3]	399/421 Carrier Anomalies[1]	421/421: marked as F5[4]	223/421 = 53%
PNG Images	10	Stego	Camouflage	10/10: data appended past EOF	10/10: marked as Camouflage[5]	N/A[2]

File Type	name	Stego Hunt	DC3 StegDetect	Provos StegDetect
executable	F5	0 anomalies	0 suspicious	N/A[2]
executable	Camouflage	0 anomalies	0 suspicious	N/A[2]
source code	F5 (Java)	0 anomalies	0 suspicious	N/A[2]

[1] Carrier Anomaly (shows inconsistent data structure)

[2] Provos StegDetect only works on JPG image format files

[3] F5 executable code, downloaded from Internet, run on a Windows machine

[4] DC3 did not extract embedded message for F5

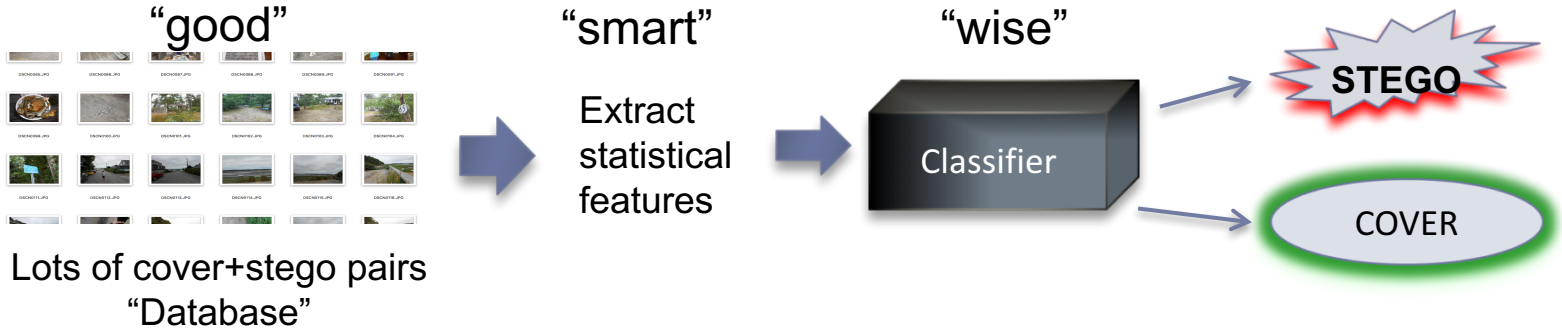
[5] DC3 also extracted password and embedded message for Camouflage

Notes:

- Stego Hunt identifies Camouflage stegos, and anomalies in most F5 stegos.
- DC3 identifies Camouflage stegos, and identifies and extracts messages in all F5 stegos.
- Neither Stego Hunt nor DC3 identifies PixelKnot stegos.
- Provos has high False Alarm Rate (24%) and high Missed Detection Rate (75%).
- Provos detects half of 421 as F5 stegos (219/421); rest identified as Outguess and jphide (4/421).

Can we use a pattern classifier to detect PixelKnot?

- Can we use a pattern classifier to detect PixelKnot?



- We used our *StegoDB Database* to train classifiers
  - Reverse-engineered PixelKnot to batch create images in database
- Error calculated is average of false alarm rate + missed detection rate
  - Depends on many factors (training, testing data, features, etc.)
  - Does not identify blunders (wrongly labeled data, programming mistakes)
  - Does not identify lies in the training phase (data that is deliberately labeled incorrectly)

Examples of Cover & Payload in PixelKnot: 5%

Cover jpg image

Message Text (2 words)

Their eyes

Stego jpg image, 5% payload size

Examples of Cover & Payload in PixelKnot: 10%

Cover jpg image

Message Text (20 words)

That yet looks on me, or would know me Ariel,
Fetch me the hat and rapier in my ce

Stego jpg image, 10% payload size

Examples of Cover & Payload in PixelKnot: 5%

Cover jpg image

Message Text (35 words)

And 'twixt the green sea and the azured vault
Set roaring war: to the dread rattling thunder
Have I given fire and rifted Jove's stout oak
With his own bolt; the strong-based promontory

stego jpg image with 5% payload size

Examples of Cover & Payload in PixelKnot: 10%

Cover jpg image

stego jpg image with 10% payload size

Message Text (85 words)

And ye that on the sands with printless foot
Do chase the ebbing Neptune and do fly him
When he comes back; you demi-puppets that
By moonshine do the green sour ringlets make,
Whereof the ewe not bites, and you whose pastime Is to make midnight mushrooms, that rejoice.To hear the solemn curfew; by whose aid,
Weak masters though ye be, I have bedimm'd
The noontide sun, call'd forth the mutinous winds, And 'twixt the green sea and

Examples of Cover & Payload in PixelKnot: 40%

Cover jpg image

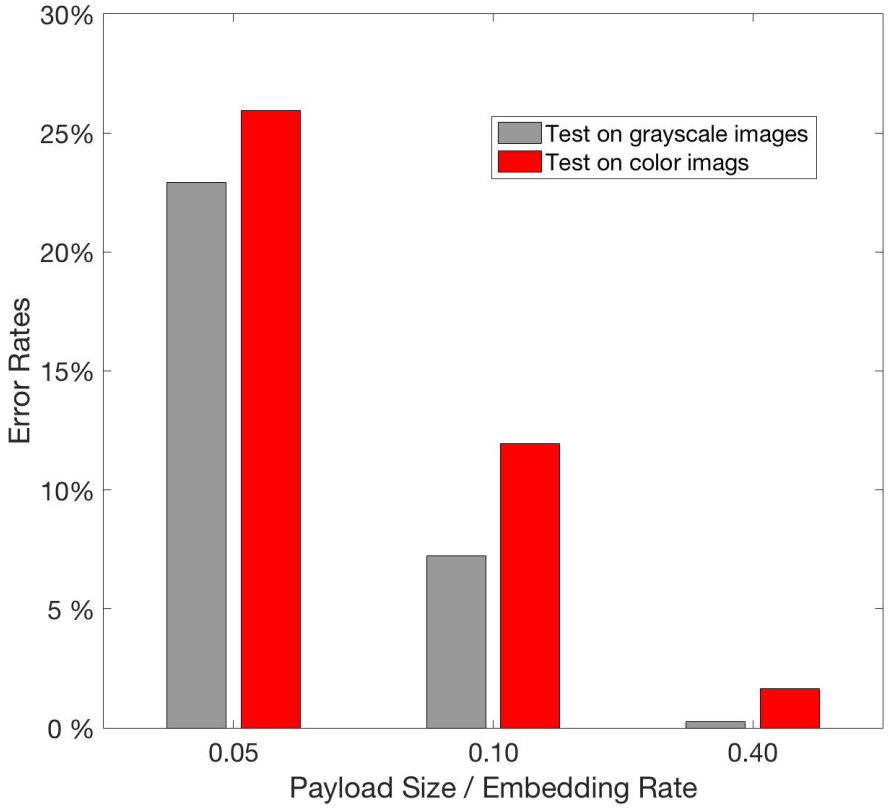
stego jpg image with 40% payload size

Message Text (350 words)

When he comes back; you demi-puppets that
By moonshine do the green sour ringlets make,
Whereof the ewe not bites, and you whose pastime
Is to make midnight mushrooms, that rejoice
To hear the solemn curfew; by whose aid,
Weak masters though ye be, I have bedimm'd
The noontide sun, call'd forth the mutinous winds,
And 'twixt the green sea and the azured vault
Set roaring war: to the dread rattling thunder
Have I given fire and rifted Jove's stout oak
With his own bolt; the strong-based promontory
Have I made shake and by the spurs pluck'd up
The pine and cedar: graves at my command
Have waked their sleepers, oped, and let 'em forth
By my so potent art. But this rough magic
I here abjure, and, when I have required
Some heavenly music, which even now I do,
To work mine end upon their senses that
This airy charm is for, I'll break my staff,
Bury it certain fathoms in the earth,
And deeper than did ever plummet sound
I'll drown my book.
A solemn air and the best comforter
To an unsettled fancy cure thy brains,
Now useless, boil'd within thy skull! There stand,
For you are spell-stopp'd.
Holy Gonzalo, honourable man,
Mine eyes, even sociable to the show of thine,
Fall fellowly drops. The charm dissolves apace,
And as the morning steals upon the night,
Melting the darkness, so their rising senses
Begin to chase the ignorant fumes that mantle
Their clearer reason. O good Gonzalo,
My true preserver, and a loyal sir
To him you follow'st! I will pay thy graces
Home both in word and deed. Most cruelly
Didst thou, Alonso, use me and my daughter:
Thy brother was a furtherer in the act.
Thou art pinch'd fort now, Sebastian. Flesh and blood,
You, brother mine, that entertain'd ambition,
Expell'd remorse and nature; who, with Sebastian,
Whose inward pinches therefore are most strong,
Would here have kill'd your king; I do forgive thee,
Unnatural though thou art. Their understanding
Begins t

Error Rates on PixelKnot classifier

- Trained on grayscale images
  - Less computationally expensive
- Tested on gray and tested on color
- Error rate: average of false alarms and missed detections
- Plot shows that the testing on color vs. grayscale images is not wildly different
  - Color has higher error because classifier did not train on color, only gray
- Our analysis shows that pattern classifiers with reasonable errors can be trained to detect stego images that are produced by mobile phone app PixelKnot



Pros and cons of different detector types

- Advantages of pattern classification
  - General detection method, does not depend on signature of stego file
  - Can be retrained to include other types of stego if necessary
- Disadvantages of pattern classification
  - Could be time consuming to retrain to include other stego classes
- Advantages of signature-based
  - Very fast to execute
  - Can identify suspicious files in addition to stego images (executables, etc.)
- Disadvantages of signature-based
  - Need to update regularly
  - Signatures can change and then will not be detected

Conclusions

- Performance of existing steg detection software packages
  - Fast and can detect steg-related files other than images
  - Possible to extract message under certain circumstances
  - Not always reliable: can't detect stegos if signature is not in package
  - Some have high FAR and MDR
  - Database of hashed signatures can become outdated, and signatures missed

- Pattern classifiers can be trained on any data
  - PixelKnot stegos can be detected by classifiers, but not by software programs such as Stego Hunt, or either StegDetects (no signature)
- Our analysis shows that pattern classifiers with reasonable errors can be trained to detect stego images that are produced by mobile phone app PixelKnot

Questions?

- Do you try to detect steg in your labs?
- What is your experience in detecting steg?
- We welcome any feedback/discussion/suggestions.