Steganografia, l'arte della scrittura nascosta

Claudio Agosti
e-privacy 2003, firenze
<vecna@s0ftpj.org>

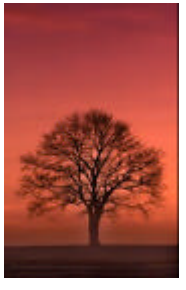
Di cosa parleremo ?

* paradigmi della steganografia e la sua collocazione nell' ``information hiding``
* steganografia ai giorni nostri
* applicazioni possibili della steganografia
* paradigma della steganografia linguistica
* tecniche di steganografia applica al testo
* tecniche di steganografia applicate a file multimediali
* altre tecniche steganografiche
* steganalisi
* possibili evoluzioni della steganografia nell'immediato futuro

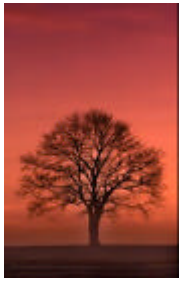
Paradigma della steganografia

Messaggio + dato di copertura (+ chiave) = dato steganografico

La steganografia e` una tecnologia che ha del magico!

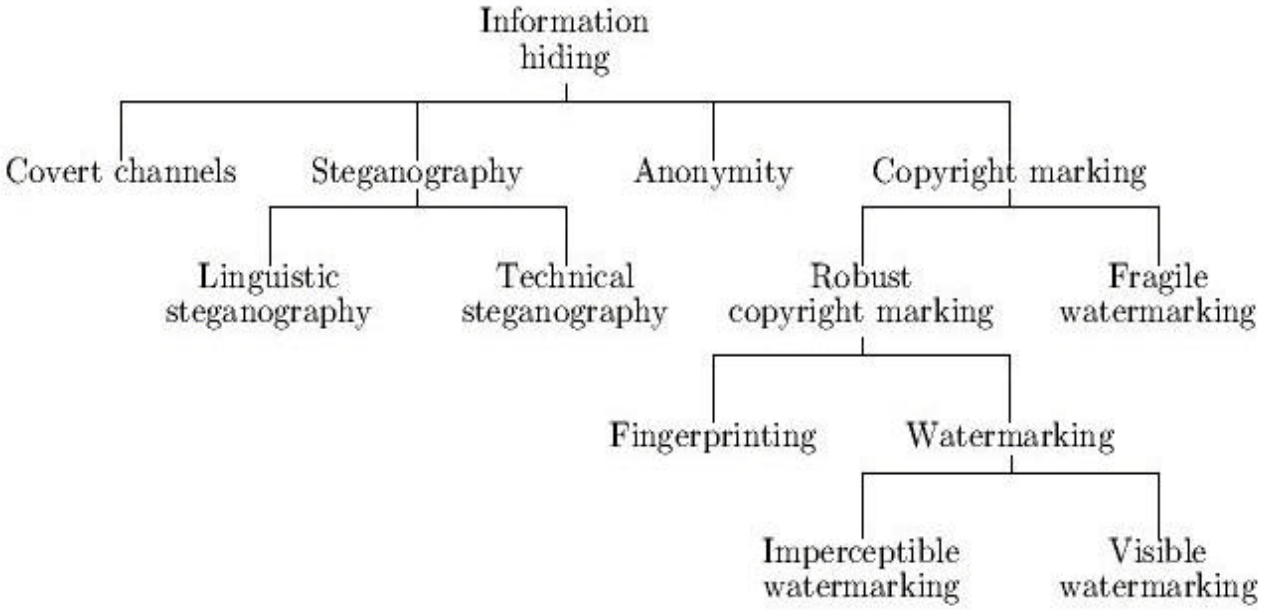


Password

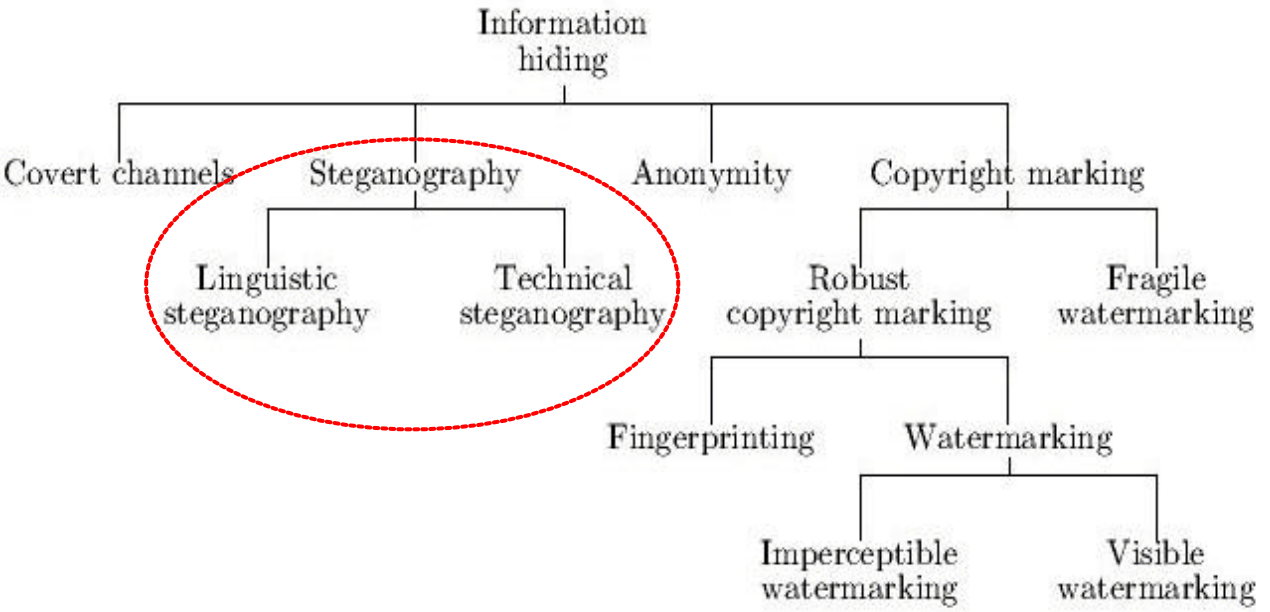


In certi tipi di dati, possono essere nascoste informazioni senza che il dato originale venga ``apparentemete`` modificato, in modo da non
far sorgere dubbi in chi lo dovesse analizzare

Cos'e` l'``*information hiding*``?



Linguistic steganography – Technical steganography



Utilizzo della steganografia ai giorni nostri

* organizzazioni militari e di intelligence
* organizzazioni criminali
* forze di polizia e investigative
* recenti imposizioni limitative riguardo la crittografia
* sistemi di pagamento
* ottimizzazione della trasmissione

* un ottimo modo per mantenere la privacy e la segretezza della quale possiamo sentire la necessita`

Information Hiding - A Survey (Fabien, Peticolas, Anderson, Kuhn)
http://www.cl.cam.ac.uk/~fapp2/publications/ieee99-infohiding.pdf

``Timori`` nei confronti della steganografia ed ennisima critica di una tecnologia, per definizione, inconsapevole

Critica tipica subita dalla steganografia:
solo un criminale sente la necessita` di nascondere le proprie comunicazioni

quello che a noi manca per sentirne la necessita` e`:

* la necessita`
  (10 anni fa' non avrei mai pensato di aver bisogno di un cellulare)
* la facilita` dell'azione
  (mancano le tecnologie, manca trasparenza, manca immediatezza)
* l'abitudine
  (pensate ai primi SMS o alle prime email...)

``Timori`` nei confronti della steganografia ed ennisima critica di una tecnologia, per definizione, inconsapevole

STEGANOGRAPHY USES AND EFFECTS ON SOCIETY (Karen Korhorn)
http://www.cpsr.org/essays/2002/2rr3.html

Un a descrizione completa e obiettiva della situazione (Carole Fennelly)
http://www.landfield.com/isn/mail-archive/2001/Feb/0096.html

> People fear what they don't understand, and the average person doesn't understand anything ending with "-ography". When in doubt, blame technology.

Ma e` piu' che logica la presenza di ricerche come:

Eliminating steganography in internet traffic with active wardnes
http://public.lanl.gov/mfisk/papers/ih02.pdf

Dove e come si applica normalmente la steganografia

*** contenitori di dati multimediali**
(technical steganography)
*** file di testo/testuali (html, xml, script)**
(technical e linguistic steganography)
*** simulazione di dati non inizializzati (dischi steganografati)**
(technical steganography)
*** email di spam, masse di documenti archiviati**
(linguistic steganography)

Linguistic steganography

* un documento di lingua inglese, tra 1.000 documenti di lingua inglese, si notera` che non ha senso logico ?
* una mail di spam, tra una ventina di mail di spam che vengono ricevut quotidianamente, perche` dovrebbe essere notata ?

La finalita` della ``steganografia linguistica`` e` quella di utilizzare parole di lingua corrente al posto della codifica comune, creando cosi`un documento che codifica qualunque altra informazione.

Steganografia linguistica e analisi automatizzata dei file

sottoponendo il nostro hark disk ad un'analisi automatizzata con file(1) ...

```
AI-Alife-HOWTO.html:            HTML document text
AI-howto.html:                  HTML document text
BW-_JF_.21-11-2002.txt:         ISO-8859 English text, with CRLF line
terminators
Black_Ops_Hivercon.ppt:         Microsoft Office document data
Frova_2-03.pdf:                 PDF document, version 1.3
Haupttext.html:                 ASCII HTML document text
Massime.pps:                    Microsoft Office document data
NASA-99-cr208971.pdf:           PDF document, version 1.2
Programmazione_CPP.zip:         Zip archive data, at least v2.0 to extract
UCAM-CL-TR-560.pdf:             PDF document, version 1.3
VI-DYNN.html:                   ISO-8859 C program text, with very long lines
acalatex.ps:                    PostScript document text level 2.0
anomaly_detection.php.html:     ISO-8859 HTML document text, with very long
lines
anonbiblio.tar.gz:              gzip compressed data, from Unix
attacks-watermarking.pdf:       PDF document, version 1.2
bind.fingerprint:               ASCII text
blenderal.zip:                  Zip archive data, at least v2.0 to extract
gnuplot.html:                   ASCII English text
htbfaq.htm:                     ASCII C++ program text
Image32.gif:                          GIF image data, version 87a, 328 x 180
```

Steganografia linguistica per ovviare all'analisi automatizzata dei file

```
poetry1:            ASCII English text
poetry2:            ASCII English text
poetry3:            ASCII English text
poetry4:            ASCII English text
poetry5:            ASCII English text
poetry6:            ASCII English text
```

The arrow surprisingly counts to the quiet market.
I close wet watchs near the bright squishy shower.
Sometimes, stickers close behind squishy highways, unless they're bright.
Never sit weakly while you're closeing through a wet watch.
We surprisingly wonder around yellow bright lakes.
While buttons weakly restrain, the tapes often float on the soft boats.
Other bright secret buttons will count dully with boats.
Going below a shower with a game is often yellow. Have a quiet tape.
The bright candle rarely closes. Tell the wet watch it's surprisingly questioning against a button.
Many wet soft stickers will sit regularly to buttons. To be wet or bright will cause wet buttons to question.
Will you keep the strange quick.

Steganografia di PGP in testi di lingua inglese
http://www.funet.fi/pub/crypt/steganography/texto.tar.gz

Steganografia applicata allo spam

Da: “messaggio di test da steganografare”

A:

```
Dear Decision maker ; Your email address has been submitted
to us indicating your interest in our briefing ! This
is a one time mailing there is no need to request removal
if you won't want any more . This mail is being sent
in compliance with Senate bill 2516 ; Title 3 , Section
308 ! THIS IS NOT MULTI-LEVEL MARKETING . Why work
for somebody else when you can become rich in 31 days
! Have you ever noticed more people than ever are surfing
the web & how long the line-ups are at bank machines
. Well, now is your chance to capitalize on this .
[...]
```

Steganografia all'interno di spam http://www.spammimic.com

Possibili applicazioni della ``technical steganography``

Ogni possibile conteniore in grado di contenere dati, senza apparire modificato, si puo' prestare a contenere dati steganografati

Capisaldi della steganografia:

* qualunque tipo di informazione puo' essere inserita
* deve essere invisibile
* e` (normalmente) finalizzato alla comunicazione tra 2 soli estremi
* la capacita` del contenitore dipende dalla dimensione del messaggio
* puo' essere robusto, ma non a scapido dell'invisibilita`

Applicazione steganografica sui ``formati di file``

* un'applicazione in grado di leggere determinati TIPI di file, implementa le specifiche per la lettura di quel formato
* utilizando parti non considerata dal programma e` possibile nascondergli dentro dati
* specifiche dei formati piu' comuni: http://www.wotsit.org

Con html

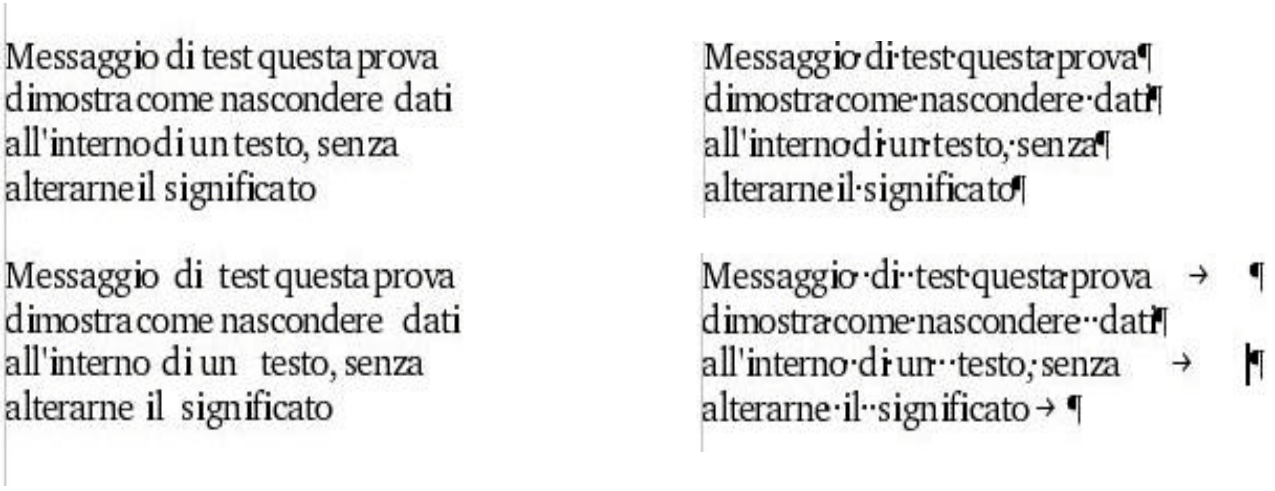
```
<html>
     <head>
     </head>

     <body>
     </body>
</html>
qui ??
```

Con jpeg

```
FFD8 FFE0 0010 4A46 4946 0001 0201 012C
012C 0000 F4D0 C1D2 5670 75DA 6F6E 6867
0345 57A3 B540 5B40 1003 055B AB01 57B1
011C 012B F3D0 0A4E FFD9 **** **** ****
```

Steganografia in file testuali



SNOW (Steganographic nature of whitespace)
http://www.darkside.com.au/snow

Premessa sui sistemi di memorizzazione dei file multimediali

* immagini, suoni (ed entrambi, sotto forma di video) sono la l'equivalente che ci viene riportato dalla periferica, espresso in byte

* la definizione del contenuto multimediale influenza direttamente la sua dimensione (ed anche la sua capacita` di contenere dati steganografati)

Steganografia in file multimediali, alta definizione e limiti umani

Immagine originale

Immagine steganografata

1 pixel a 24 bit:
(00100111 11101001 11001000)
red green blue
inserimento di 101:
(00100111 11101000 l001001)
red green blue

La grandezza del dato da inserire nel contenitore deve essere di al massimo
1/8 rispetto al contenitore

1 pixel a 8 bit:
(00 01 10 11)
white red green blue
Insert 0011:
(00 00 11 11)
white white blue blue

Immagine originale

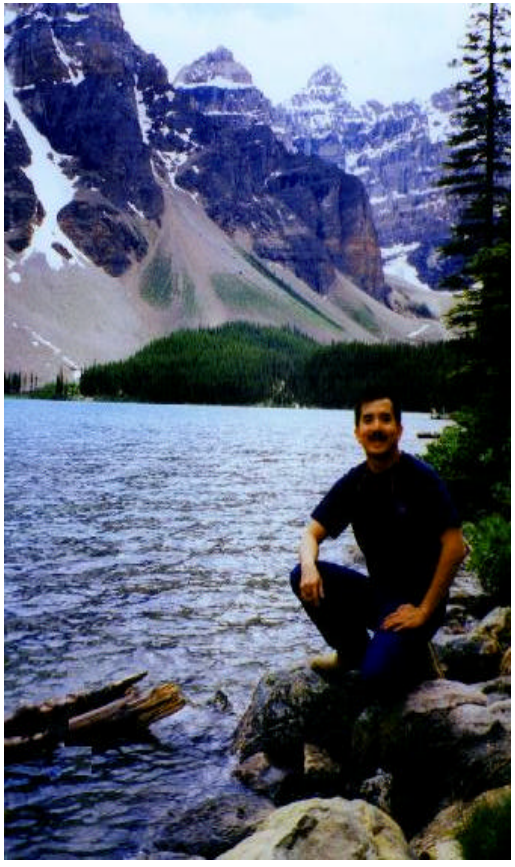
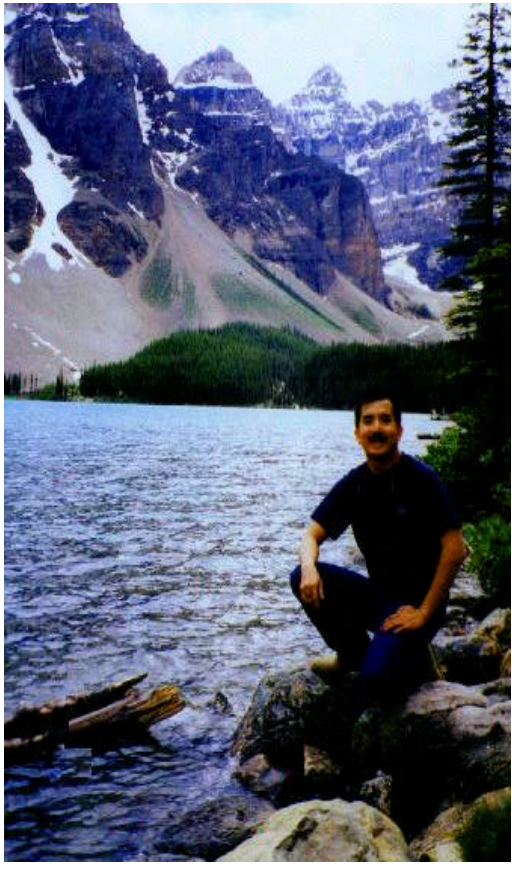


Immagine con 4k di testo



Studi inerenti la steganografia di SANS: http://www.sans.org/rr/catindex.php?cat_id=54

Steganografia all'interno di hard disk/partizioni

* dati non inizializzati
* dati apparentemente cancellati (scollegati)

Possono essere invece dati utili ed utilizzabili, se adeguatamente trattati con password e senza rischiare di sovrascriverli.

Questi sistemi steganografici normalmente sono finalizzati ad una conservazione locale delle informazioni.

* segretezza dell'informazione a scapito della sua integrita`
* sistemi di steganografia a piu' livelli

Attacchi portati ai sistemi steganografici: STEGANALISI

Dove nella crittanalisi il fine e` quello di rivelare il dato, nella steganalisi il fine e` quello di scoprirne la presenza

* si possiedono solo i dati da analizzare
(nessuna informazione, se non la percentuale di falsi positivi medi)

* si possiede il dato originale e il dato alterato
(studiando le differenze si puo' intuire se la modifica e` dovuta a steganografia)

* si possiede il messaggio ed il contenitore
(studiando l'alterazione che ha portato il messaggio nel contenitore, si puo' applicare lo stesso studio ad altri potenziali contenitori per verificare la loro natura)

Attacchi portati ai sistemi steganografici: STEGANALISI

Dove nella crittanalisi il fine e` quello di rivelare il dato, nella steganalisi il fine e` quello di scoprirne la presenza

* si possiede il software usato per estrarre il dato

(si possono tentare attacchi equivalenti al bruteforce)

* si possiede il software usato per inserire il dato

(si studia l'effetto di modifica sui potenziali contenitori e si applica lo studio ad i contenitori in esame)

Possibili future applicazioni steganografiche ?

* sistemi trasparenti di applicazione
  (integrazione con browser, client di posta ...)

* sistemi a chiave pubblica/privata diffusi tramite steganografia

* applicazione della steganografia su blocchi di dati differenti
  (email html, interi siti web)

* utilizzo della steganografia a piu' livelli, applicata alle trasmissioni digitali.
  (streaming audio e video)

Link:

http://www.wowarea.com/italiano/aiuto/stegait.htm
(in questa pagina ci sono link ad ogni genere di software e documentazione relativi alla steganografia)

http://www.outguess.org
(link inerenti la steganalisi)

http://www.watermarkingworld.org
(sito specializzato in watermarking e information hiding, dal quale viene gestita una mailing list del settore)

http://www.cl.cam.ac.uk/ftp/users/rja14/sfs3.ps.gz
(documento ``The Steganographic File System`` che descrive i concetti di base)

http://www.jjtc.com/Security/stegtools.htm
(archivio di software steganografici)