MOCM * 20 Settembre 2003 * Pescara

Nuovi progetti di steganografia

Claudio Agosti <vecna@s0ftpj.org>

Parleremo di...

* paradigma di steganografia, concetti principali
* Attacchi ai sistemi di steganografia
* tecniche attuali
* vulnerabilita` delle tecniche attuali
* progetti di steganografia in seguito alle prime esperienze

Concetti principali della steganografia

La steganogafia e` la tecnica in grado di rendere un dato da trasmettere **invisibile** a chiunque non sia in grado di estrarlo.

Per renderlo invisibile si utilizza un “contenitore”, ovvero un dato che apparentemente debba sembrare l'oggetto della trasmissione, quando il vero oggetto e` nascosto al suo interno

Concetti principali della steganografia

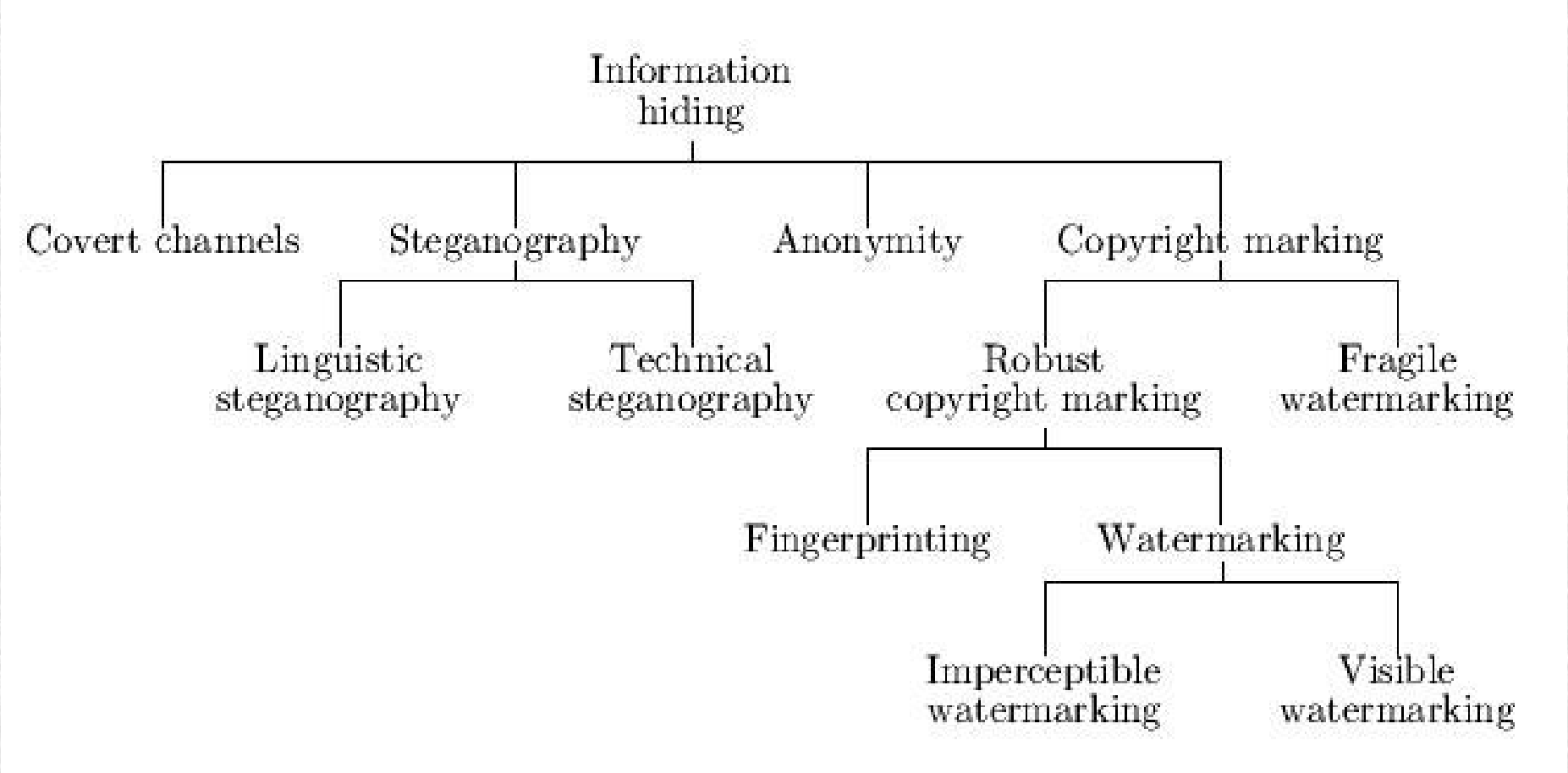
Messaggio + dato di copertura (+ chiave) = dato steganografico

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

Steganografia e “information hiding”



Attacchi ai sistemi di steganografia

* Individuazione del contenitore steganografico tra tanti potenziali contenitori
  (steganalisi)

* estrazione del messaggio steganografato nel contenitore
  (crittanalisi)

* eliminazione del dato nascosto
  (modifica tramite ulteriore steganografia)

Possibili applicazioni della steganalisi

Possedendo in analisi:

* dei potenziali contenitori
* il software di steganografia, e dei poteziali contenitori
* il software di estrazione, e dei potenziali contenitori
* dei potenziali contenitori, tra i quali alcuni apparentemente modificati
* il dato originale e dei potenziali contenitori

Ed ogni possibile combinazione di questi elementi.

Tecniche attuali di steganografia

* steganografia su LSB
  (noise manipoulation, image downgrading)

* generazioni mimiche

* steganografia con tecniche di distorsione

* steganografia su *“palette-based Image”*

* steganografia su dati quantizzati

* steganografia in zone riservate o non utilizzate

* steganografia modulata su in fase

Steganografia con tecniche applicate a LSB/noise manipoulation

01011101
10010110
01001101
01010011
00110111
11101011

Processo steganografico

01011100
10010111
01001101
01010010
00110110
11101011

Generazioni mimiche

* visto che la steganalisi e` effettuata prima in modo automatizzato, (e solo sui positivi effettuata un'analisi dettagliata) si puo' far forza su sistemi che non illuderebbero un uomo, ma sono indistinguibili per una macchina

* nelle funzioni mimiche il dato da nascondere **genera** il dato di copertura, assumento un significato apparente ed un aspetto tipico ed innocente

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.

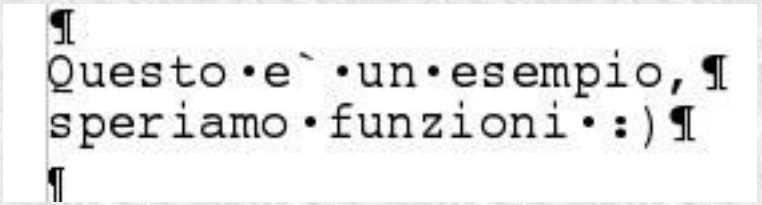
Steganografia con tecniche di distorsione

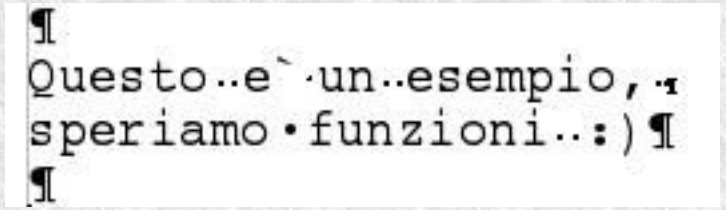
* quando i due capi della comunicazione possiedono il dato di copertura, originale, tramite l'inserimento e l'analisi di differenze e` possibile nascondere informazioni

* l'inserimento dell'informazione viene applicato a caratteristiche al quale il creatore del contenitore ha normalmente accesso.

```
Questo e` un esempio,
speriamo funzioni :)
```

```
Questo e` un esempio,
speriamo funzioni :)
```





Steganografia su *palette-based images*

* l'applicazione di LSB sulla tavolozza (con forte diminuzione delle performance)

* ordinamento arbitrario della tavolozza, e cambio dei riferimenti nell'immagine

Altre tecniche steganografiche

* su dati quantizzati
  potendo predirre una quantita` di dati con certezza, i dati predicibili vengono sostituiti il messaggio

* zone non utilizzate, zone riservate
  in vari formati sono presenti zone di dati non utilizzate, riciclabili per contenere il messaggio

* steganografia modulata in fase
  per steganografare dati all'interno di file audio si puo' nascondere il messaggio come parte del brano, senza sovrascrivere il rumore arbitrariamente

Steganografia in rete, *“covert channel”*

I contenitori possono essere i piu' disparati, si e` visto per ora utilizzare:

* i dati trasmessi all'interno dei ping

* la porta sorgente, o i flag non utilizzati, all'interno di sessioni TCP non stabilite

* i HTTP GET/PUT verso server web

* L'ip sorgente di pacchetti IGMP

Steganografia con tecniche applicate a LSB/noise manipoulation

01011101
10010110
01001101
01010011
00110111
11101011

Processo steganografico

01011100
10010111
01001101
01010010
00110110
11101011

Steganalisi

Vulnerabilita` della tecnica:
Analisi dei bit meno significativi, per la ricerca di pattern ed analisi statistiche.

Generazioni mimiche

* visto che la steganalisi e` effettuata prima in modo automatizzato, (e solo sui positivi effettuata un'analisi dettagliata) si puo' far forza su sistemi che non illuderebbero un uomo, ma sono indistinguibili per una macchina

* nelle funzioni mimiche il dato da nascondere **genera** il dato di copertura, assumento un significato apparente ed un aspetto tipico ed innocente

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.

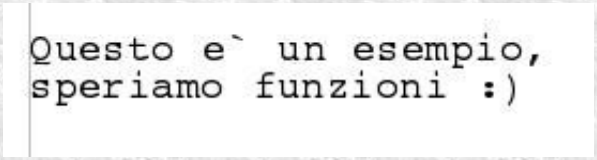
Vulnerabilita` della tecnica:
Utilizzo di sistemi esperti per l'individuazione dell'assenza di significato

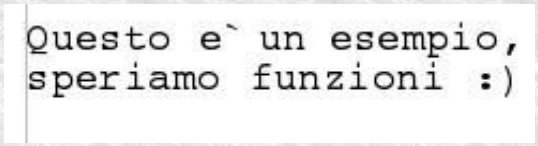
Steganografia con tecniche di distorsione

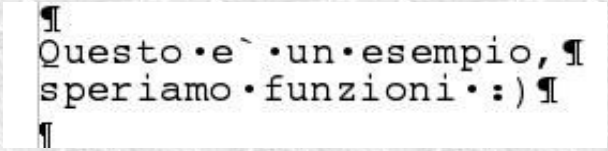
Steganalisi

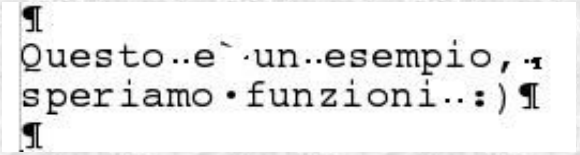
* quando i due capi della comunicazione possiedono il dato di copertura, originale, tramite l'inserimento e l'analisi di differenze e` possibile nascondere informazioni

* l'inserimento dell'informazione viene applicato a caratteristiche al quale il creatore del contenitore ha normalmente accesso.









Vulnerabilita` della tecnica:
Analisi statistica al layer adeguato, notando un'anomalia troppo ripetitiva e troppo unica

Steganografia in rete, *“covert channel”*

Steganalisi

I contenitori possono essere i piu' disparati, si e` visto per ora utilizzare:

* i dati trasmessi all'interno dei ping

* la porta sorgente, o i flag non utilizzati, all'interno di sessioni TCP non stabilite

* i HTTP GET/PUT verso server web

* L'ip sorgente di pacchetti IGMP

Vulnerabilita` della tecnica:
le trasmissioni sono assolutamente inusuali, in alcuni casi lasciano spazio ad identificativi che potrebbero causare un loro filtro preventivo

Finalita` dei futuri progetti di steganografia

* essere invulnerabili alla steganalisi, almeno a livello teorico.
  (attualmente, quasi ogni tecnica di steganografia potrebbe aver la sua contromisura atta a mostrarne l'utilizzo su contenitori steganografati)

* riuscire ad essere trasparente all'utente, e fornire i propri pregi potendo utilizzare i software comunemente utilizzati

* studiare sistemi di steganografia ottimizzata, sfruttando piu' livelli di inserimento, in modo da poter usare il meno possibile ogni tipo di contenitore e far sembrare le singole analisi falsi positivi

Progetto n.1: steganografia su un intero sito web

* il progetto e` finalizzato a sfruttare piu' sistemi di steganografia, per nascondere un solo dato (cifrato e gestito tramite GPG o altri sistemi a chiave pubblica)

* utilizzando vari software gia` esistenti,
  outguess (JPEG)
  gifshuffle (GIF)
  snow (file di testo, applicabile a .html)

il progetto mira a suddividere in modo poco invasivo il dato da nascondere, in modo da non abusare della capacita` del contenuto, fornendo al tempo stesso un sistema di steganografia in supporto ad un sistema di cifratura a chiave pubblica.

Progetto n.2: steganografia su IP applicata a sessioni stabili

* I sistemi di steganografia applicate alle trasmissioni di rete (*covert channel*), sfruttano i piu' disparati sistemi, ma han due lacune:

  - il traffico che generano e` sempre e comunque, inusuale
  - non vi e` un invio effettivo di dati ne l'utilizzo di alcun protocollo di sessione

* e` necessario garantire:
  - la credibilita` della connessione (e non fornire discriminanti che potrebbero farla escludere con un firewall), comprendendo sia l'uso di un protocollo coerente, che l'invio di dati, che la simulazione di interattivita` ecc...

Progetto n.2: steganografia su IP applicata a sessioni stabili

- il campo dell'ID nell'header IP, serve per IDENTIFICARE i singoli pacchetti, normalmente ogni stack tiene un contatore e lo usa in modo incrementale, ma alcune implementazioni (linux+GRSec e OpenBSD), la utilizzano random.

  Questo crea l'opportunita` di sfruttare quel campo come contenitore di dati:

```
14:14:26.150156 216.239.59.99.www > portatile.32963: . [tcp sum ok]
8098:8098(0) ack 2099 win 31740 [tos 0x10]  (ttl 50, id 50871, len 40)
14:14:26.256673 216.239.59.99.www > portatile.32964: . [tcp sum ok]
582:582(0) ack 1040 win 31740 [tos 0x10]  (ttl 50, id 61454, len 40)
14:14:26.282356 216.239.59.99.www > portatile.32964: P 582:1962(1380)
ack 1040 win 31740 [tos 0x10]  (ttl 50, id 61898, len 1420)
14:14:26.284187 216.239.59.99.www > portatile.32964: P 1962:2302(340)
ack 1040 win 31740 [tos 0x10]  (ttl 50, id 61900, len 380)
14:14:26.284223 portatile.32964 > 216.239.59.99.www: . [tcp sum ok]
1040:1040(0) ack 2302 win 12420 (DF) (ttl 64, id 8455, len 40)
```

Link:

http://vecna.itapac.net

home page personale, ricerche, progetti, software...

http://e-privacy.firenze.linux.it

E-privacy, convegno annuale sulla privacy digitale,
dal punto di vista legale e tecnico.

E` possibile trovare le slide della mia ultima presentazione:

“Steganografia, l'arte della scrittura nascosta”

nel quale vengono spiegate le tecniche principali,
mostrati link ad alcuni software ed il loro effetto,
ricerche e discussioni in merito.

http://www.s0ftpj.org

Gruppo tramite il quale espongo i miei lavori.