

Il nascondino dei bit

Nascondere e trovare dati
La nuova corsa agli armamenti

Mario Pascucci

Convegno CFItaly - Libera Università degli Studi S. Pio V
Roma, 18 giugno 2008

Nascondere o cifrare?

Cifrare

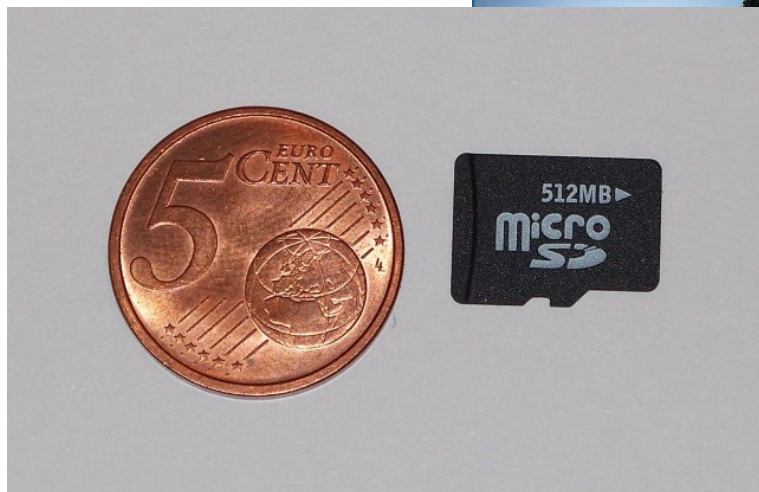
- ✓ Non forzabile in tempi brevi
- ✓ Alla portata di chiunque
- ✗ Facilmente individuabile
- ✗ Bug noti nell'applicazione
- ✗ Sicurezza legata alla robustezza delle chiavi

Nascondere

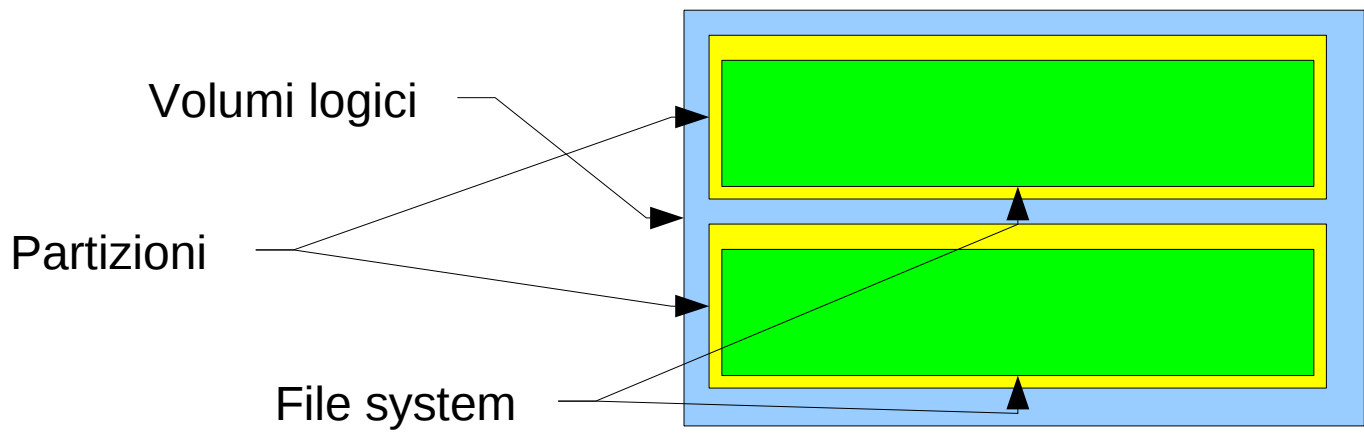
- ✓ Infiniti modi
- ✓ Non si vede, non c'è
- ✓ Applicazioni sviluppate ad hoc
- ✗ Non alla portata di tutti
- ✗ Risultato non certo

Dove nascondere

La risposta più ovvia è: nel computer!
E' l'unica risposta possibile?



Strutture...

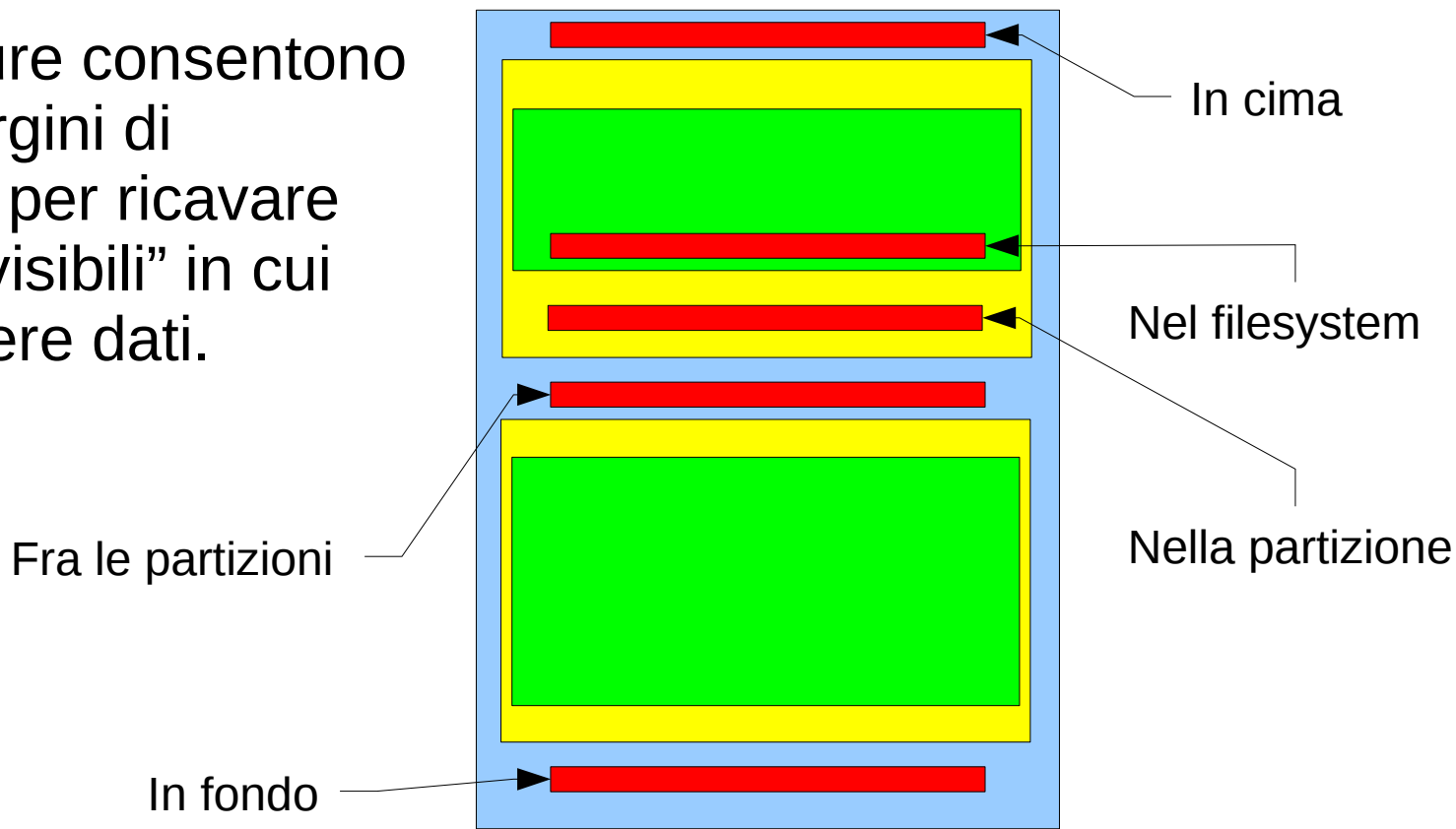


In un comune supporto di memoria i dati sono inseriti in strutture ben definite



...e nascondigli

Le strutture consentono ampi margini di manovra per ricavare spazi "invisibili" in cui nascondere dati.



Idee geniali?

- Banali calcoli dimostrano l'esistenza di spazio inutilizzato
- I supporti possono essere verificati per conoscerne la reale capacità
- Semplici comandi possono mostrare discrepanze fra partizione e filesystem

I dati vengono trovati
in breve tempo

Strani file

Ogni filesystem ha:

- Strutture di organizzazione dello spazio
- Database dei file
- Attributi dei file
- Aree di servizio

Ognuna di queste può essere sfruttata in molti modi per nascondere uno o più file

Alcuni esempi:

- File nascosti
- File cancellati
- Spazio libero
- *Bad clusters*

Troppo strani

- Lo spazio libero è per definizione pronto per essere utilizzato: i dati ivi memorizzati sono ad alto rischio
- Lo stesso accade per i file cancellati: per il sistema operativo è spazio riutilizzabile
- Un file nascosto è semplicemente un file con un particolare attributo: toglierlo è banale come metterlo
- I *bad cluster* sono i settori inutilizzabili a causa dei difetti del supporto fisico. Verificarne il reale stato è banale (S.M.A.R.T. ad esempio)

Piece of cake

Nascondere in bella vista

- In alcuni sistemi operativi basta cambiare il nome di un file per cambiarne il tipo (icona e associazione)
- Concatenando due file, il secondo “scompare”
- Inglobare un file più piccolo nella struttura di uno più grande
- Alcuni tipi di file possiedono una struttura a segmenti, in cui è possibile inserire contenuti estranei al file

Tutti questi metodi sono inefficaci

Foremost & co

- Questa semplice applicazione rende inefficaci tutti i sistemi per nascondere dati elencati fino a questo punto
- Esegue una ricerca sull'intero supporto ignorandone la struttura di partizioni e filesystem
- Trova in quale settore inizia un file, cercando “firme” di formati noti
- Con il contorno di altre applicazioni dedicate ad ogni filesystem si può risalire a quale struttura appartiene il settore

Se il file c'è, lo trova.

Antiforensic

- Conoscendo il funzionamento di Foremost si può agire per evadere una scansione effettuata con questo strumento
- Un limite di Foremost è il funzionamento *signature-based*, simile a quello degli antivirus: riconosce solo i file di cui possiede la firma
- Basta cambiare pochi bytes in un file per renderlo “invisibile” a Foremost
- Altra possibilità è usare codifiche molto semplici, come ad esempio Base64 o Rot13, che cambiano profondamente la “firma” di un file

Qualche esempio

- Il file in formato PDF inizia con la stringa “%PDF-” e termina con “%EOF”
- Codificato in Base64 diventano rispettivamente “JVBERi0” e “U9GCg==”, e non è detto che la stringa finale sia sempre la stessa, a causa del tipo di codifica.
- C'è di peggio: il semplice programma Rot13, contenuto nel pacchetto *BSD-games* di molte distribuzioni Linux, opera una codifica aggiungendo 13 al valore di ogni carattere alfabetico, come il cifrario di Cesare.
- Con Rot13 le stringhe diventano: “%CQS-” e “%RBS”

E allora?

- Dovendo esaminare un solo file, non è un problema ricostruirne il formato.
- La dimensione media di un disco moderno è oltre i 100Gb, e sono sul mercato a costi accessibili dischi per PC da 1Tb
- Individuare un file camuffato in mezzo a un milione di altri non è un compito banale.
- In una indagine il tempo è un lusso.

Pura fantasia

- Prendiamo una immagine JPEG
- Applichiamo una codifica Base64, seguita da una codifica Rot13 con una chiave differente da 13
- Inseriamo il file ottenuto in un file Wave, in una posizione arbitraria
- Il file Wave è in un disco da 250Gb, insieme ad altri 200.000 file.

Come trovarlo in tempi umani?

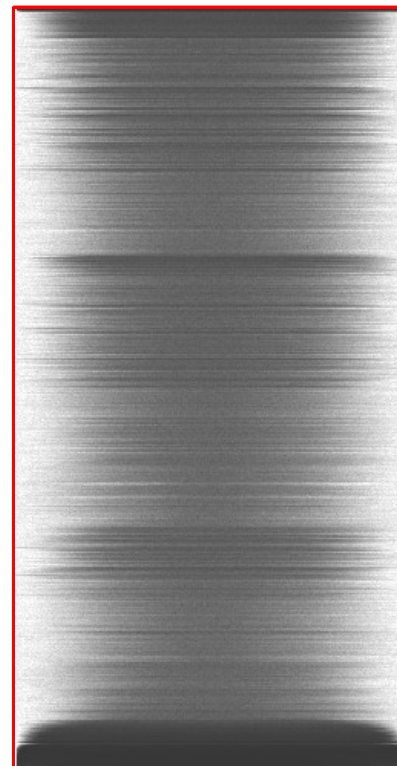
Un barlume di speranza

- Trattare un file come il risultato di una misurazione
- Usare l'analisi delle frequenze e ricavarne uno spettrogramma
- Applicare allo spettrogramma algoritmi derivati dal riconoscimento del parlato per individuare la tipologia di dati contenuti nel file
- Non è una idea nuova: nel 2006 il Mar. Ord. Giuseppe Finizia del RIS ha presentato un documento in cui mostra un software per realizzare lo spettrogramma.

(rif.: <http://www.marcomattiucci.it/GFinizia.art.01.v1.0a.pdf>)

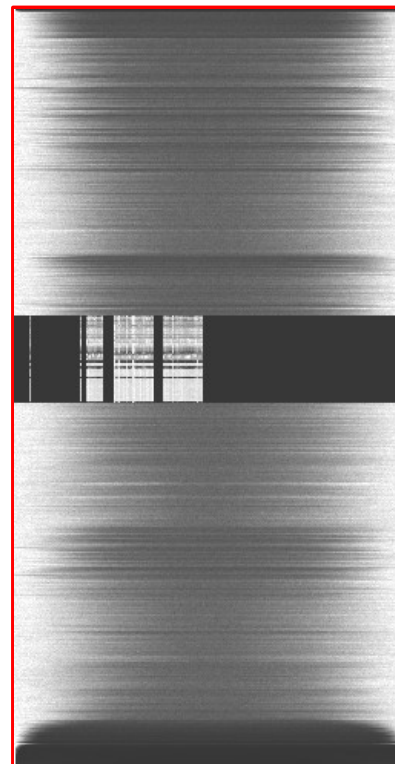
Possibilità

- Ho realizzato un semplice programma che produce due tipi di spettrogramma: totale e segmentato
- L'output è una immagine PNG a livelli di grigio
- In figura i due tipi di spettrogramma di un file Wave



Possibilità (2)

- Inserendo un file PDF codificato in Base64 lo spettrogramma cambia vistosamente
- Foremost individua solo il file Wave “contenitore”



Arriva Internet

- La presenza di numerosi servizi web di condivisione e collaborazione ad accesso gratuito (il c.d. Web 2.0) cambia radicalmente il problema
- Condividere una foto su Flickr è una operazione compiuta abitualmente da migliaia di persona al giorno
- Scrivere un articolo su Wikipedia è alla portata di chiunque
- Scrivere e gestire un blog è banale

E' tutto bellissimo.

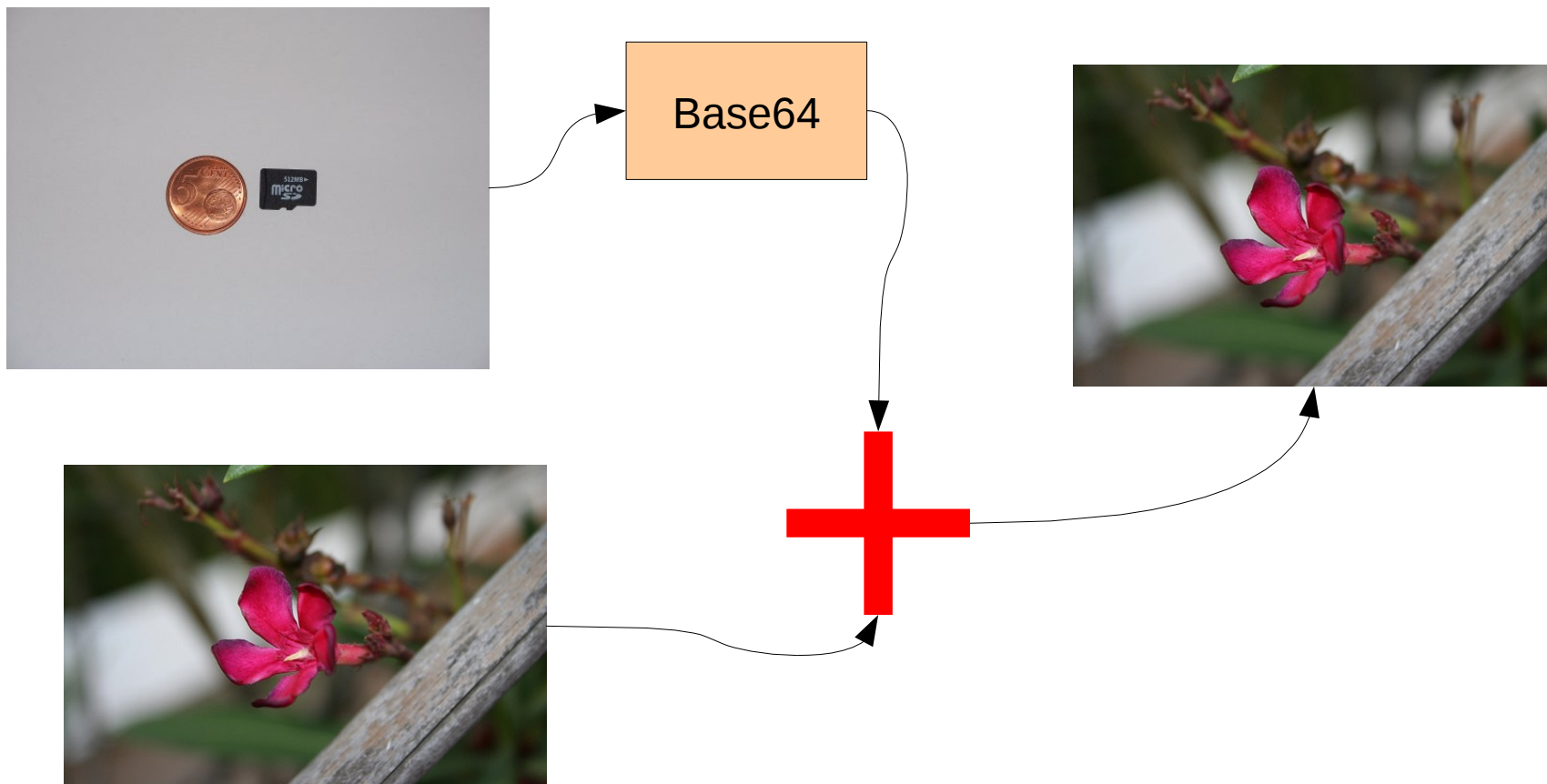
Un album di foto

- Ho codificato l'immagine con la microSD in base64
- Presa una differente immagine ho “concatenato” i due file e li ho caricati su Flickr, in un account gratuito.
- Poi ho scaricato la foto, verificando che fosse identica a quella inviata.
- Ho separato i due file, trasformato il file in Base64 nell'originale JPEG
- Risultato: il secondo file immagine è perfettamente integro.

L'immagine è qui:

<http://www.flickr.com/photos/ilpettegolo/2581219722/>

Un album di foto (2)



Own3d B1og

- Il numero di blog è in continua crescita
- Scrivere in un blog è estremamente semplice, altrettanto non può dirsi per l'amministrazione: gran parte dei blogger “fai da te” non ha le competenze necessarie
- A partire dalla seconda metà del 2007 il numero di blog violati o compromessi è in continua crescita
- L'intrusione, nella maggioranza dei casi, si limita a iniettare link nascosti alle solite rivendite di pillole azzurre, ma vi sono anche siti che tentano di iniettare malware attraverso il browser.

Own3d B1og (2)

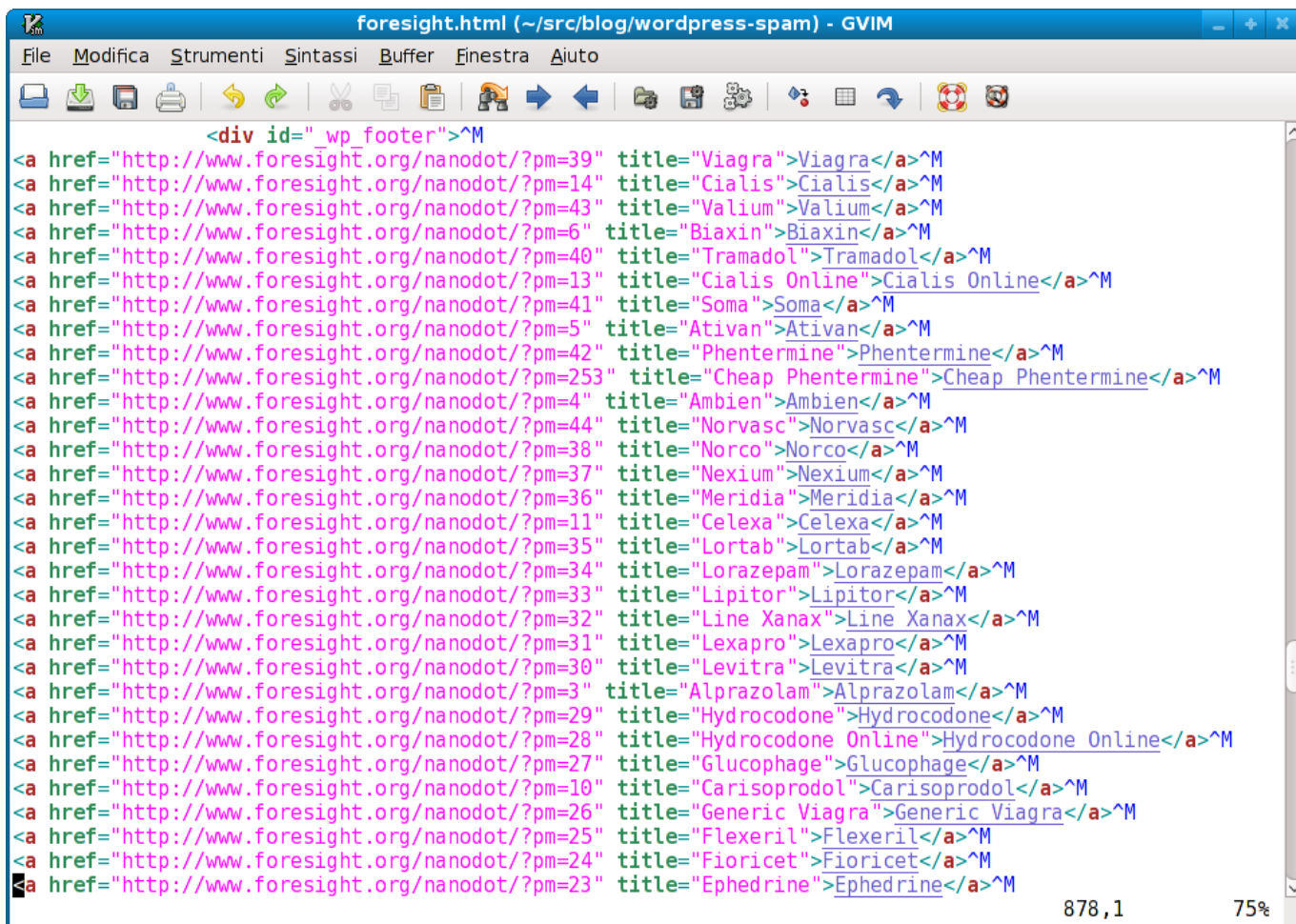
Il proprietario del blog non si accorge di nulla, ma il codice malevolo iniettato nel sito agisce in modo molto sottile:

- I link vengono inseriti solo se a visitare il blog è uno spider di Google, Yahoo, MSN, Live, ecc.
- Nel blog viene creato un amministratore “abusivo”, invisibile dal pannello di controllo. L'amministratore “abusivo” sopravvive ad un aggiornamento del software
- Se un visitatore arriva da Google seguendo uno dei link, viene deviato verso il sito della farmacia online
- Se è un visitatore abituale, il redirect non si attiva

Solo per i tuoi occhi

- Nel blog compromesso viene iniettato del codice PHP, opportunamente offuscato, nascosto in file sepolti in directory poco usate ed attivato come plugin.
- All'arrivo di un visitatore, viene controllata la sua impronta di navigazione, ossia *User Agent*, *Referer* ed eventuali *Cookies*.
- Nel caso di Google, lo *User Agent* dello spider contiene la stringa "Googlebot".
- Il codice malevolo si attiva quando rileva appunto specifici *User Agent* e "farcisce" opportunamente la pagina di link.

La farcitura di link



The screenshot shows a GVIM editor window titled "foresight.html (~/.src/blog/wordpress-spam) - GVIM". The editor displays a list of HTML anchor tags, each with a href pointing to a specific post on foresight.org and a title for a pharmaceutical product. The list includes:

- [Viagra](http://www.foresight.org/nanodot/?pm=39 "Viagra")
- [Cialis](http://www.foresight.org/nanodot/?pm=14 "Cialis")
- [Valium](http://www.foresight.org/nanodot/?pm=43 "Valium")
- [Biaxin](http://www.foresight.org/nanodot/?pm=6 "Biaxin")
- [Tramadol](http://www.foresight.org/nanodot/?pm=40 "Tramadol")
- [Cialis Online](http://www.foresight.org/nanodot/?pm=13 "Cialis Online")
- [Soma](http://www.foresight.org/nanodot/?pm=41 "Soma")
- [Ativan](http://www.foresight.org/nanodot/?pm=5 "Ativan")
- [Phentermine](http://www.foresight.org/nanodot/?pm=42 "Phentermine")
- [Cheap Phentermine](http://www.foresight.org/nanodot/?pm=253 "Cheap Phentermine")
- [Ambien](http://www.foresight.org/nanodot/?pm=4 "Ambien")
- [Norvasc](http://www.foresight.org/nanodot/?pm=44 "Norvasc")
- [Norco](http://www.foresight.org/nanodot/?pm=38 "Norco")
- [Nexium](http://www.foresight.org/nanodot/?pm=37 "Nexium")
- [Meridia](http://www.foresight.org/nanodot/?pm=36 "Meridia")
- [Celexa](http://www.foresight.org/nanodot/?pm=11 "Celexa")
- [Lortab](http://www.foresight.org/nanodot/?pm=35 "Lortab")
- [Lorazepam](http://www.foresight.org/nanodot/?pm=34 "Lorazepam")
- [Lipitor](http://www.foresight.org/nanodot/?pm=33 "Lipitor")
- [Line Xanax](http://www.foresight.org/nanodot/?pm=32 "Line Xanax")
- [Lexapro](http://www.foresight.org/nanodot/?pm=31 "Lexapro")
- [Levitra](http://www.foresight.org/nanodot/?pm=30 "Levitra")
- [Alprazolam](http://www.foresight.org/nanodot/?pm=3 "Alprazolam")
- [Hydrocodone](http://www.foresight.org/nanodot/?pm=29 "Hydrocodone")
- [Hydrocodone Online](http://www.foresight.org/nanodot/?pm=28 "Hydrocodone Online")
- [Glucophage](http://www.foresight.org/nanodot/?pm=27 "Glucophage")
- [Carisoprodol](http://www.foresight.org/nanodot/?pm=10 "Carisoprodol")
- [Generic Viagra](http://www.foresight.org/nanodot/?pm=26 "Generic Viagra")
- [Flexeril](http://www.foresight.org/nanodot/?pm=25 "Flexeril")
- [Fioricet](http://www.foresight.org/nanodot/?pm=24 "Fioricet")
- [Ephedrine](http://www.foresight.org/nanodot/?pm=23 "Ephedrine")

The editor interface includes a menu bar (File, Modifica, Strumenti, Sintassi, Buffer, Finestra, Aiuto) and a toolbar with various editing tools. The status bar at the bottom right shows "878,1" and "75%".

Sa dov'è una farmacia?

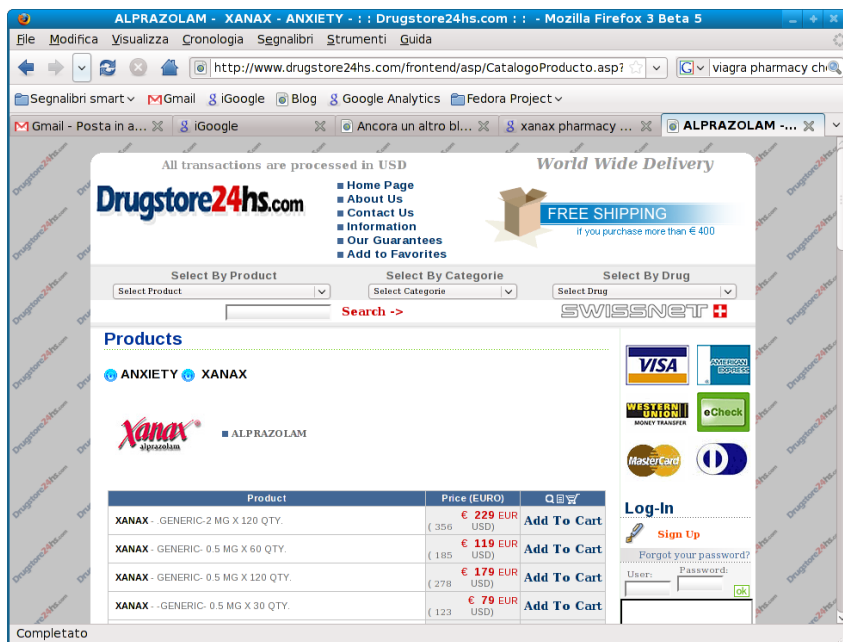
- Cercando in Google i soliti farmaci, si ottengono link appartenenti apparentemente ad uno dei blog compromessi:

[Buy Line Xanax from Google Approved Pharmacy](#) - [[Traduci questa pagina](#)]
Consultant **pharmacists** are regulated separately from physicians. Specifically, the legislation Buy Line **Xanax** Buy Line **Xanax** the Buy Line **Xanax** of **pharmacy** ...
www.bbgeeks.com/?google-approved=78 - [Pagine simili](#) - [Salva risultato](#)

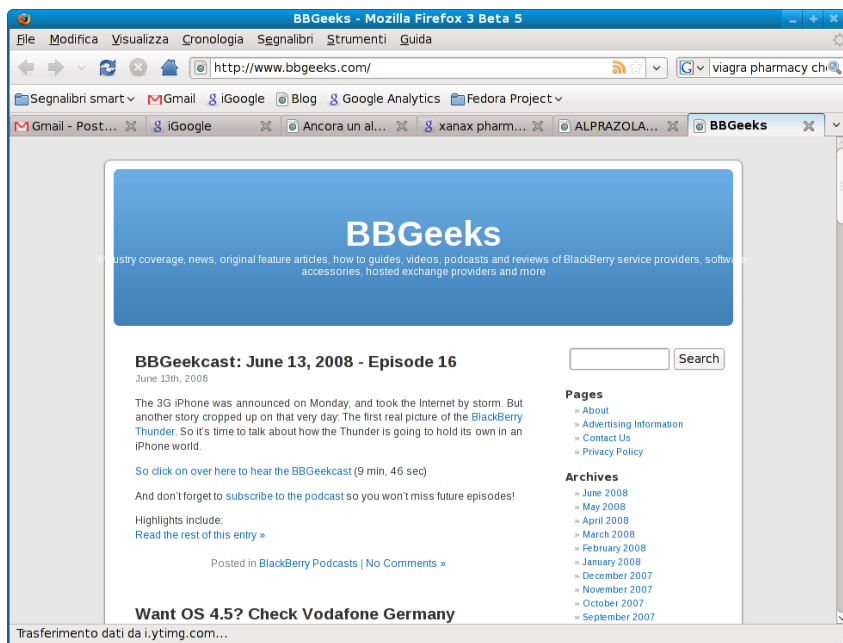
- Cliccando sul link si finisce in prima istanza nel blog compromesso, per essere immediatamente rediretti su un sito che vende farmaci online

Ce l'hai davanti!

- Questo è il sito che ci si trova davanti, dopo due salti
- Se si digita il link direttamente, il redirect non avviene. Se si clicca sul link è il proprietario del blog o un suo visitatore abituale, finisce sul blog, come ci si aspetta.



Io non la vedo...



- Il codice annidato nel blog controlla se il visitatore ha impostato un *Cookie*, quello di sessione del blog: se c'è è un visitatore abituale o il proprietario. Per non destare sospetti, si disattiva.
- Se il *Cookie* non c'è, controlla il *referer*, ossia da dove arriva il visitatore. Se viene da Google, devia verso la farmacia, altrimenti mostra il blog.

Che c'entra?

- Per ottenere questo risultato, sono state impiegate tecniche di *data hiding* molto sofisticate.
- Il codice è iniettato anche nel database, sotto forma di falsa opzione di configurazione o di storico delle news, offuscato con una codifica Base64 e invertito.
- E' pensato per nascondersi sia ai visitatori che al proprietario. I suoi effetti cambiano in funzione delle “impronte di navigazione” del visitatore.
- L'intrusione sopravvive ad un upgrade: il blog mostrato è aggiornato all'ultima versione.

C'entra, c'entra...

- Basti immaginare un codice simile che però non si limiti a iniettare link, ma a mostrare ad esempio un album di foto “speciali”, o un documento particolare.
- Il contenuto nascosto è mostrato ad esempio solo se il visitatore ha una certa stringa nello *User Agent*.
- Oppure, se ha un certo Cookie impostato ad un particolare valore: in questo caso non rimane traccia neanche nei log del web server.
- Altrimenti il visitatore vedrebbe solo un blog che parla del gruppo musicale del momento.

Che ci vuole?

- Chi ha compromesso il blog potrebbe sfruttarlo per un po', poi cancellare tutte le tracce e chiudere baracca e burattini. Per confondere le acque potrebbe installare la versione con i link alle pillole azzurre.
- Il proprietario non sospetterebbe nulla. Non ci sarebbero neanche i link su Google.
- Nei log del web server non ci sarebbe quasi nulla.

Quanti strumenti consolidati di Computer Forensic sono ancora utilizzabili?

Competenza

- Un approccio basato solo su strumenti e potenza di calcolo è inutile
- Con lo spostarsi in Rete di molte attività è sempre più facile per un malintenzionato trovare siti vulnerabili.
- Certamente stiamo parlando di persone con competenze al di sopra della media, e quindi non comuni. Le organizzazioni criminali, però, non mancano certo di denaro per assoldare Black Hat reperiti nel mondo dell'underground.
- Le conoscenze per l'analisi in un simile frangente sono estesissime. Troppe per una sola persona.

Spaventati?

E' un buon segno. Vuol dire che qualcosa sono riuscito a trasmettere!

Grazie per l'attenzione!

Per contattarmi: <http://www.ismprofessional.net/pascucci/>

Oppure: mpascucci@gmail.com

Realizzata con 



