Selective File Dumper
Un utile strumento Open Source di computer forensics e data recovery.

Da un'indagine informatica nasce l'idea di un tool interattivo per il recupero dei files selezionati per estensione.

Autore: Nanni Bassetti

La digital forensics è una disciplina che raccoglie tutti i sistemi per investigare e scoprire informazioni nascoste sui media digitali, anche quando si pensa di aver cancellato tutto dal proprio computer. Da questo articolo si capisce come con molta semplicità si può dare del filo da torcere all'investigatore informatico.

Dall'articolo imparerai...

- Ad usare il tool

Cosa Dovresti sapere

- Elementi di digital forensics
- Conoscenza dei file system
- Conoscenza dello Sleuthkit

Sull'autore

L'autore lavora dal 1998 come sviluppatore di applicazioni web e system administrator e dal 2006 si occupa di sicurezza informatica e computer forensics.
Ha scritto un libro sulla sicurezza del web ed un thriller informatico “Onphalos”.
Collabora con la procura del Tribunale di Bari per le indagini informatiche.
Inoltre ha scritto un libro sulla Computer Forensics – INDAGINI DIGITALI - http://www.lulu.com/content/1356430
Ha sviluppato un tool di computer forensics insieme a Denis Frati (http://www.denisfrati.it).
Web site: http://www.nannibassetti.com

In Rete

Tools:
Sleuthkit (http://www.sleuthkit.org)
Foremost (http://foremost.sourceforge.net)
Selective File Dumper - (http://sfdumper.sourceforge.net/)
Bash Guide - (http://tldp.org/LDP/abs/html/)
Zenity - (http://freshmeat.net/projects/zenity)

Che cos'è l'Open Source? È una filosofia, è un stato mentale, è la fede nella condivisione delle conoscenze e nel progresso e nel genere umano.

Perché? Perché chi sviluppa per la comunità Open Source non pensa a proteggere il suo lavoro, ma pensa a dare un contributo agli altri e crede nella cooperazione per migliorare la sua idea.

Un altro vantaggio è che, spesso, le idee, i progetti e lo sviluppo corrono sui cavi di Internet, tra persone che non si conoscono fisicamente e sono sparsi in tutto il mondo, non più uffici o orari di lavoro, ma solo passione e cooperazione.

Seguendo questa filosofia, noi, (lo scrivente e Denis Frati), abbiamo sviluppato un nuovo strumento

"**Selective File Dumper**", pubblicato su **Sourgeforce.net** la più grande comunità Open Source del web presso questo URL **http://sfdumper.sourceforge.net**

L'inizio

Stavo facendo un'indagine di computer forensics, usando **Linux Kubuntu 7.10** e tutti tools open source forensi come **Autopsy** e **Sleuthkit**, **Foremost**, ecc.

Dovevo trovare qualcosa probabilmente inclusa in file con estensione DOC, PDF o BMP e l'hard disk in esame era di circa 200 Gb, ho provato con la ricerca per keywords, ma non sapevo esattamente che cosa stavo cercando, perché il mio mandato era generico.

Per questa ragione, avevo bisogno di esportare e salvare, tutti i file DOC, PDF e BMP, eliminati e referenziati manualmente, in una directory chiamata "REPERTI", poi dare tutti questi file all'Autorità Giudiziaria per la valutazione.

Parlando con il mio amico Denis Frati, di questo problema, abbiamo pensato ad un software che potesse estrarre, automaticamente, tutti i files del tipo prescelto, cancellati e referenziati e poi effettuare un carving sulla partizione selezionata, cancellando i file carvati doppioni di quelli cancellati e referenziati ed infine poter fare una ricerca per keywords sull'insieme dei files estratti (referenziati, cancellati e carvati), così nacque Selective File Dumper (SFDumper.sh).

Sfdumper è un tool che permette il recupero dei file, utile quindi all'operatore che intenda lavorare con le infomazioni in essi contenute (il testo del documento, il contenuto del file immagine) e non sulle informazioni correlate al file, quali i metadati, che sono comunque recuperabili per i singoli file, grazie alle informazioni incluse nei log, principalmente il percorso e nome del file e il suo inode.

I software richiesti, per il funzionamento dello script, sono:

Sleuthkit - per il recupero dei file referenziati e cancellati.

Foremost - per il data carving.

Sha256deep – per calcolare il codice hash dei file referenziati, eliminati e carvati.

Grep - per la ricerca di parole chiave.

Awk - per analizzare l'output del software di cui sopra.

Sed - per analizzare l'output del software di cui sopra.

DD – per fornire il bitstream della partizione prescelta come input per Foremost.

L'idea era chiara, il progetto è basato su un *Linux Bash Script* interattivo, utilizzando tutti softwares Open Source, testati ed accettati dalla comunità di computer forensics, per automatizzare molte operazioni manuali.

L'output è solo su quattro directory:

1) File referenziati

2) File Eliminati

3) File carvati

4) Reports e logs.

Il primo ostacolo nel quale ci siamo imbattuti èstato il corretto riconoscimento del file system e delle partizioni.

Abbiamo utilizzato **mmls** e **fsstat** (Sleuthkit) per determinare se vi fosse un file system sul file immagine o sul dispositivo che si va ad analizzare.

Grazie alla combinazione di questi due potenti strumenti, si possono avere risultati diversi:

1) Non c'è file system o partizione (tra questi riconosciuti dalla Sleuthkit) - è possibile solo il data carving.

2) Vi è una strana partizione a partire dal settore ZERO - potrebbe essere una partizione.

3) C'è il file system e le partizioni ed è possibile scegliere che cosa si vuole analizzare.

Per i casi 2 e 3 abbiamo utilizzato **fls** e **icat** (Sleuthkit) per estrarre tutti i file referenziati e quelli cancellati dalla partizione da noi scelta.

INCOMINCIAMO

Possiamo lanciare lo script in due modi:

sudo sh sfdumper.sh

Oppure

chmod + x sfdumper.sh

. / sdumper.sh

Eseguiamo lo script!

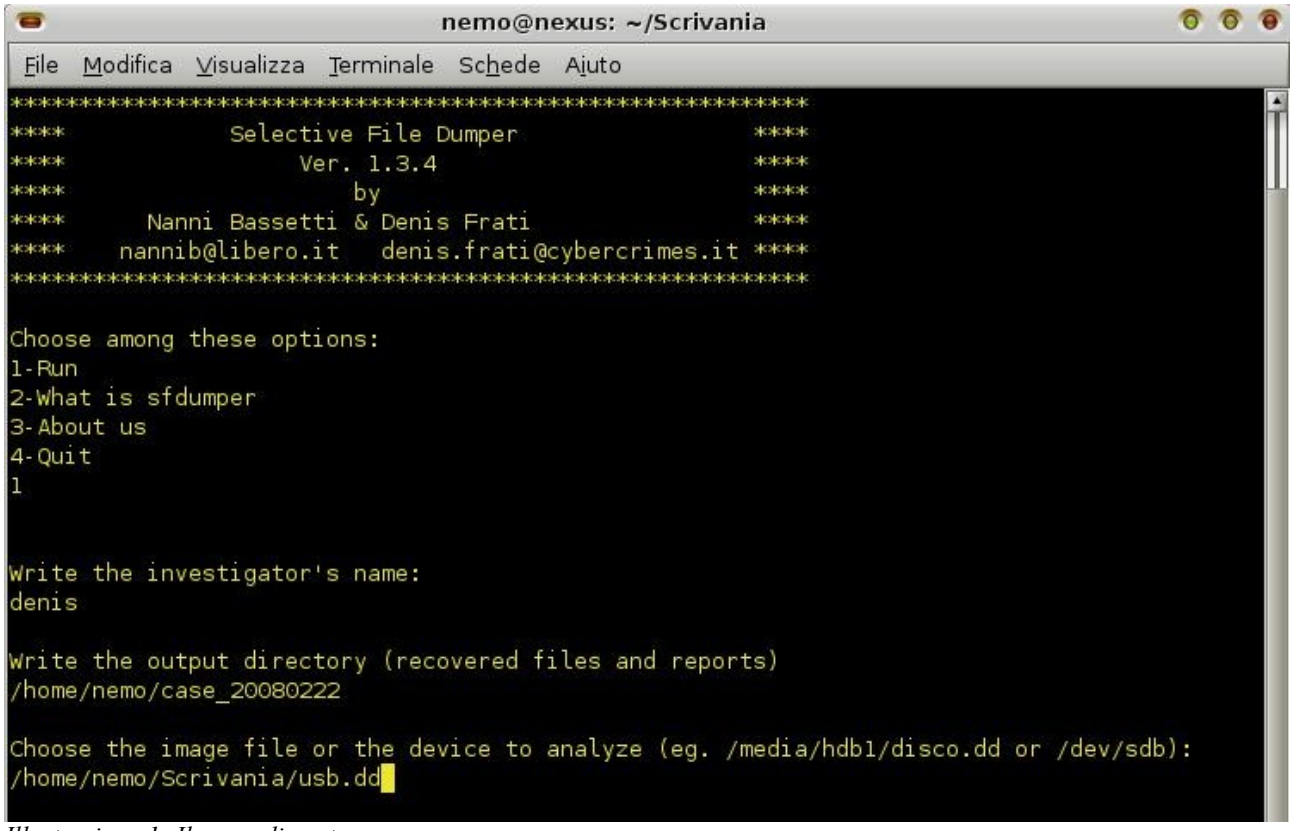

Illustrazione 1: Il menu di partenza

In primo luogo, dobbiamo inserire il "*Investigator's name*", poi bisogna indicare la directory di output per i files recuperati e per i reports.

In secondo luogo si deve indicare il file di immagine o il dispositivo (ad esempio / dev / sdb), a questo punto, si potrebbe avere un output simile a questo (Figura 2):

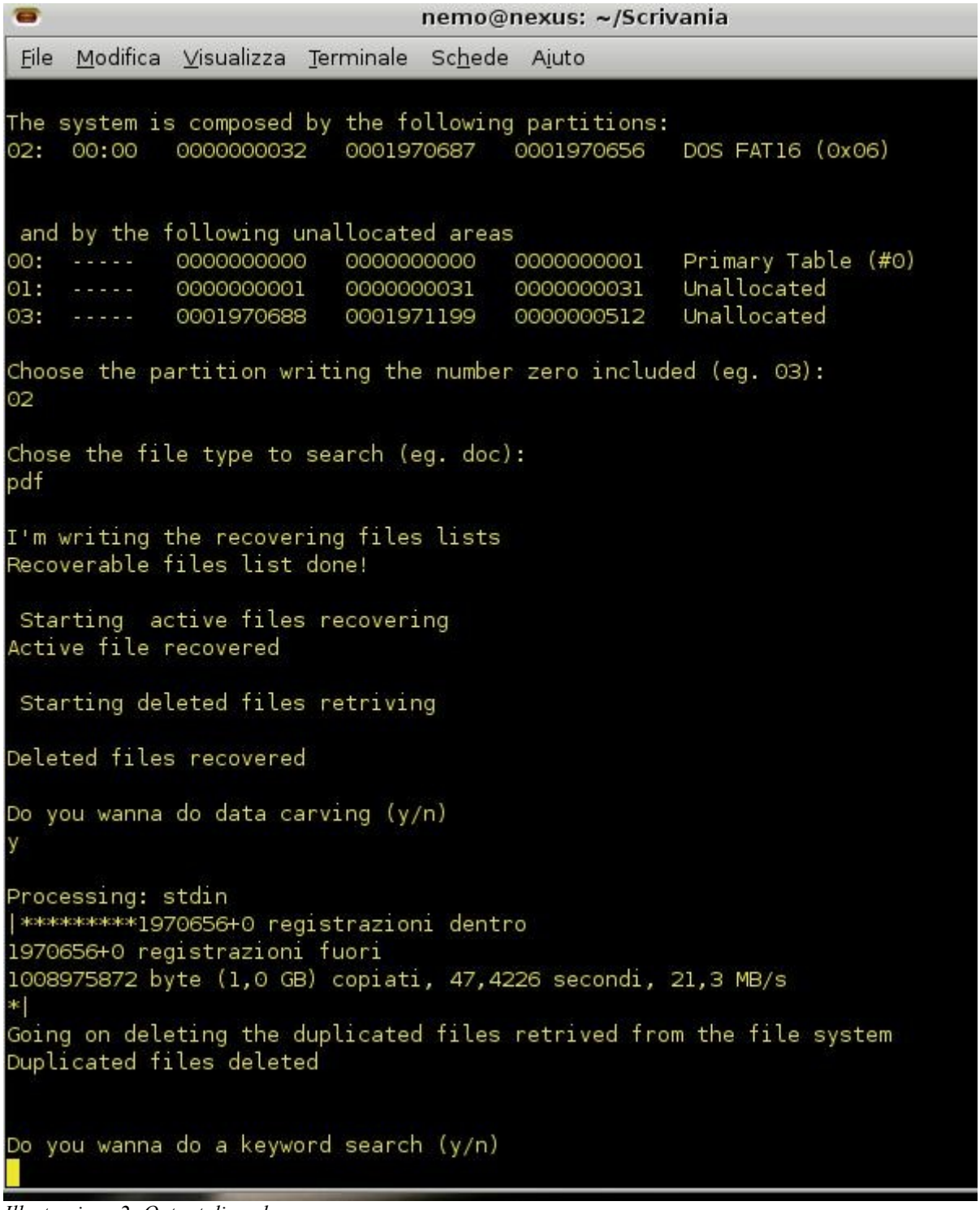

Illustrazione 2: Output di mmls

Illustrazione 2 – Output di mmls

Quando scegliamo una partizione provvederemo, ad esempio, ad indicare il proprio numero. "**02**", allora dobbiamo scrivere il tipo di file di cui abbiamo bisogno per es. doc, jpg, pdf, ecc.

Nella directory di output vedremo una directory nominata in questo modo:

Partition_NumberOfThePartitionChoosed__fileType

Ad esempio: **partition_02__pdf**

Oppure

In caso non vi sia il file system, si tratterà di un file immagine *raw* o dispositivo senza file system (es. formattato) , quindi vedremo una directory nominata in questo modo:

Image__fileType

Ad es.: **Image__pdf**

All'interno di queste directory si troveranno altre directory:

Deleted_files_recovered, **referenced_files_recoverd**, **carvingFileType** (es. carvingpdf), Report

E 'evidente cosa contengono….

Illustrazione 3: Directories

I nomi dei file saranno composti in questo modo:

Inode_path_filename corrispondente, ma gli spazi saranno "_" e il slashes diverranno "-".

Ad esempio: **1994280_My_Documents-story.doc**

Dopo aver indicare il tipo di file si può scegliere se si vuole fare un data carving, in caso affermativo lo script utilizzerà i potenti **Foremost** e **dd**.

Lo script, utilizzando il linux pipe "|", fornirà in input a Foremost il bistream solo della partizione prescelta:

dd if = $ imm conv = noerror, sync | soprattutto-t-o $ 2 $ carv_dir;

Oppure

dd if = $ imm conv = noerror, sync | soprattutto-c $ dir_output "tmp / carving_config.txt"-o $ carv_dir

Dove $imm è la variabile che rappresenta la partizione e $carv_dir rappresenta la directory di uscita del data carving.

Lo script, automaticamente, crea un file di configurazione per Foremost chiamato "*carving_config.txt*" nella "tmp" directory all'interno della directory di output, che abbiamo scelto, in modo da poter utilizzare le estensioni build-in di Foremost o, se il tipo di file non è tra queste, è possibile utilizzare il file di configurazione, che, inoltre, è estensibile manualmente, permettendo così di aggiungere sempre nuovi tipi di file.

Se abbiamo scelto di fare il data carving, avremo tutti i file carvati, ma tra questi ci saranno anche i files referenziati ed eliminati, che non servono e quindi dovranno essere eliminati, per lasciare solo quelli presenti nello spazio non allocato.

Per questo motivo, lo script esegue lo **Sha256deep** su tutti i filese referenziati, cancellati e carvati, al fine di confrontare i codici hash dei file ed eliminare i file carvati che hanno lo stesso hash di quelli referenziati e di quelli cancellati.

L'algoritmo Sha256 hash è stato scelto in quanto, attualmente, non sono stati riscontrati problemi di collisione.

Fare il carving è intelligente e utile, in quanto così si è grado di scoprire i file rinominati, per esempio, se un file è stato nominato *terrorist_plan.doc* ed è stato rinominato in *terrorist_plan.gfx*, quando lo script eseguirà la ricerca di tipo di file "doc ", non lo troverà né nei file di referenziati né nei file eliminati, ma a Foremost non sfuggirà, poiché guarda i file headers e non le estensioni, quindi riconoscerà il formato del file anche se esso è stato rinomnato.

Dopo la richiesta per il carving, lo script chiede se si vuole effettuare una ricerca per parola chiave, in questo caso si è utilizzato il meraviglioso strumento **grep**, tramite il quale lo script cercherà la parola chiave scelta tra i files referenziati, cancellati e carvati e scriverà i nomi dei files, che contengono la parola chiave, in un file di registro denominato *key_found.txt*.

Tutte le operazioni e le liste dei file sono riportati in alcuni file nella directory "Report".

Alla fine lo script chiederà se si vuole fare un'altra analisi, in questo modo è possibile cercare diversi tipi di file e in diverse partizioni, fino a quando noi non si decide di terminare lo script.

La versione GUI

Abbiamo anche sviluppato una versione GUI (Graphics User Interface) di Selective File Dumper, utilizzando le librerie grafiche di Zenity.

"*Zenity is a tool that allows you to display Gtk+ dialog boxes from the command line and through*

shell scripts.

It is similar to gdialog, but is intended to be saner.

It comes from the same family as dialog, Xdialog, and cdialog, but it surpasses those projects by having a cooler name." – tratto dal sito web di Zenity.

Questa versione è stata realizzata al fine di fornire un'interfacci più amichevole, ma necessita di un ulteriore software di terzi, che è appunto Zenity.

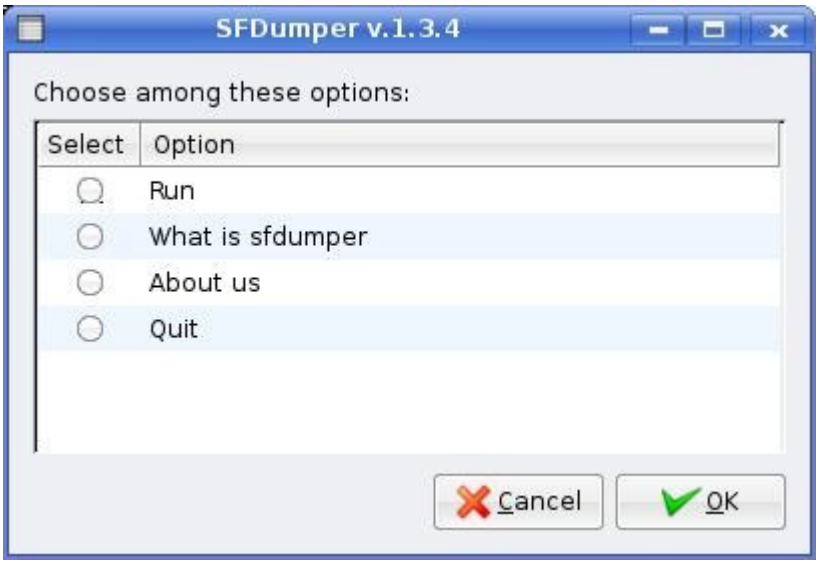

Illustrazione 4: SFDumper GUI Menu

Il vantaggio della versione SHELL rispetto a quella GUI è che la prima può lavorare direttamente dal Linux **runlevel 2**, che significa che funzionerà anche sulle distro con sola riga di comando e senza ambiente grafico di lavoro.

Illustrazione 5: Finestra di dialogo per la ricerca di keywords

Conclusioni

Questo è stato un duro lavoro, nato su un effettivo bisogno, perché non abbiamo mai sentito parlare di un software che faccia tutte queste cose, almeno nel mondo Open Source e queste sono operazioni relativamente semplici, ma erano tutte manualmente, prima di **Selective File Dumper**, in questo modo speriamo, che ciò aiuterà molti operatori di computer forensics e al recupero dei dati.

Non usando **sfdumper**, è molto difficile ottenere gli stessi risultati con il solo ausilio della riga di comando, quindi abbiamo sviluppato questo script interattivo, che con poche domande, in grado di fare molte operazioni.

"*The last but not the least* " è che lo scrivente e Denis Frati non si sono mai incontrati, il nostro progetto e la nostra amicizia è nata su Internet e abbiamo lavorato utilizzando solo e-mail, Skype, il

telefono cellulare-

Questo modo di lavorare rappresenta la filosofia Open Source, la quale è orientato verso la condivisione della conoscenza e la delocalizzazione, permettendo a tutti di partecipare ad un progetto.

Noi speriamo di avere molti feedback e suggerimenti dagli utenti e dagli sviluppatori, al fine di migliorare il nostro lavoro e le nostre future idee.

CONTATTI

Selective File Dumper - http://sfdumper.sourceforge.net

Nanni Bassetti - nannib@libero.it - http://www.nannibassetti.com

Denis Frati - denis.frati@cybercrimes.it - http://www.denisfrati.it