di Nanni Bassetti

LE PROCEDURE ED I METODI DELLA DIGITAL FORENSICS SONO COSI' IMPORTANTI?

Nanni BASSETTI, laureato in Scienze dell'Informazione a Bari, è libero professionista specializzato in informatica forense. Iscritto all'albo dei C.T.U. presso il Tribunale di Bari. Project manager di CAINE Linux Live Distro forense, Fondatore di CFI – Computer Forensics Italy, la più grande community di computer forensics italiana.

Che impatto avrebbe il non effettuare una copia bit a bit e non calcolare i codici hash sull'attendibilità dei dati che saranno analizzati? Accendere un computer per ispezionarlo prima di acquisire la copia dei dischi, cosa comporta? Spesso sembra che adottare delle procedure sia solo un esercizio di stile, a avolte sembre importante solo trovare le evidenze.

Sappiamo da tempo che esistono delle best practices[1] nella digital forensics, che impongono dei metodi per acquisire ed analizzare le fonti di prova digitali, in Italia è in vigore la Legge 48/2008[2] che genericamente impone delle cautele nelle operazioni suddette, ma a volte ci si imbatte in elaborati tecnici o verbali che sono carenti di alcune metodologie o procedure scientifiche[3]. La domanda da porsi è se queste procedure siano necessarie o siano solo un esercizio accademico, perchè lo scontro è sempre sull'impatto che possono avere sulla veridicità, integrità e correttezza legale e scientifica. Insomma quanto incide il non aver seguito rigorosamente una queste metodologie?
Nella fase d'acquisizione di un reperto digitale si dovrebbe agire cercando di non alterarlo, quindi non operare sui dati contenuti in un dispositivo ma trasportandoli così come sono e per questo si parla di copia bitstream o bit a bit di un dispositivo di memorizzazione (es. Hard disk, pendrive,ecc.) per ottenere questo si devono adottare sistemi di write blocking, ossia blocchi in scrittura sul dispositivo ed infine per garantire l'identicità va calcolato il codice hash[4] della sorgente e della destinazione, se i due codici coincidono allora la copia e l'originale saranno identici.

Che impatto avrebbe il non effettuare una copia bit a bit e non calcolare i codici hash sull'attendibilità dei dati che saranno analizzati?

Immaginiamo una copia fatta solo col copia ed incolla di tutte le cartelle ed i file di un computer e col calcolo dell'hash:
Range valori **soggettivi**: 0 - 10; Impatto = Valore Inf. x Perdita informazioni
Dove per **Completezza** intendiamo la totalità delle informazioni presenti sul disco (metadati di file system, partizioni, registri, ecc.), per **Integrità** intendiamo la genuinità e l'identicità del file copia con quello originale, per **Analisi** il valore e la quantità di informazioni ricavabili da quei dati slegati dal contesto generale in cui erano presenti.

CIR	Valore informazione	Perdita	Impatto totale
Completezza	7	3	21
Integrità	10	0	0
Analisi	6	4	24
			55

In questo caso gli impatti maggiori sono sulla completezza dell'informazione che perde tutto lo spazio non allocato (file cancellati) e altri dati sul file system e sull'analisi che sarà carente di tante informazioni a contorno. Se eliminiamo anche il calcolo dell'hash vediamo che:
Infatti si perde la garanzia d'integrità e genuinità dei file, essi potrebbero non esserci tutti, alcuni potrebbero esser stati alterati, insomma non abbiamo un parametro di controllo per verificare se ciò che è stato preso il giorno dell'acquisizione corrisponde a ciò che stiamo analizzando al momento.

CIR	Valore informazione	Perdita	Impatto totale
Completezza	7	3	21
Integrità	2	8	16
Analisi	6	4	24
			61

1 https://www.swgde.org/documents/Current%20Documents/2014-09-05%20SWGDE%20Best%20Practices%20for%20Computer%20Forensics%20V3-1

2 http://www.parlamento.it/parlam/leggi/08048l.htm - Legge 48/2008 art. 354 CPP – co. 2.

3 https://it.wikipedia.org/wiki/Metodo_scientifico

4 Le funzioni di HASH sono funzioni matematiche, che applicate ad un determinato file o documento creano una stringa alfanumerica, detto codice di HASH di lunghezza predefinita. Ad ogni file corrisponde il suo HASH. Se il file fosse modificato in qualunque sua parte il codice HASH sarebbe differente.

Un altro problema potrebbe essere che l'hash sia stato calcolato ma non riportato su un verbale di operazioni compiute, ma scritto solo sul supporto di memorizzazione che conserva la copia stessa.
In questo caso si avrebbe una minaccia sull'integrità, data dal fatto che non si può affermare se la copia sia la vera copia oppure qualcos'altro scambiato, con un hash ricalcolato, dato che non si ha un termine di paragone, non si può asserire che la copia sia la stessa dell'epoca dell'acquisizione, diversamente avendo il codice hash su un supporto separato, come un verbale, si potrebbe subito controllarne la genuinità.

Vediamo che succede facendo una copia bit a bit con calcolo del codice hash del sorgente e della destinazione:

Ecco spiegato in modo quantitativo come l'impatto su tutti i parametri si azzera e questo può giustificare una procedura forense d'acquisizione, non per dei meri capricci accademici ma per un reale impatto sull'analisi.

CIR	Valore informazione	Perdita	Impatto totale
Completezza	10	0	0
Integrità	10	0	0
Analisi	10	0	0
			0

Accendere un computer per ispezionarlo prima di acquisire la copia dei dischi, cosa comporta?
In questo caso si possono avere perdite di dati dovute alla cancellazione accidentale o di sistema, che lavorando va a sovrascrivere aree di memoria che potevano essere recuperate, inoltre il sistema potrebbe aggiornare le date e gli orari dei file, potrebbe bloccarsi, potrebbe scaricare dalla rete qualcosa (se non si è provveduto a sconnettere il computer), potrebbe concludere un'installazione o azionare qualcosa che doveva iniziare al prossimo riavvio, insomma ci sono tante incognite che potrebbero alterare anche in modo importante il reperto informatico, quindi anche in questo caso, l'evitare di dare la classica "occhiata" al PC non è un capriccio.

I dispositivi mobili, se trovati accesi vanno isolati da qualsiasi rete, l'ideale sarebbe usare buste di Faraday o almeno inserire la modalità aereo e lasciarli accesi in carica (power bank), poichè non si sa se ad un'eventuale riaccensione si potrebbe incorrere in richieste di codici d'accesso o altre problematiche, se invece lo si trova spento meglio lasciarlo in quello stato fino al laboratorio dove si procederà all'analisi. Anche in questo caso non è un vezzo teorico, ma l'impatto sull'integrità e disponibilità dei dati può essere notevole. Esempio: un telefono lasciato connesso in rete potrebbe essere bloccato o resettato da remoto da qualcuno.

L'acquisizione di pagine web, va fatta cercando di garantire il più possibile la genuinità digitale delle stesse, la data certa, l'origine e la destinazione, il browser usato, ecc. ecc. Questo perchè una mera stampa, come per le e-mail, non assicura che la pagina sia stata alterata o in che data era visibile con quei contenuti o magari solo da quella connessione o solo con un determinato browser o vi fossero altri elementi nel codice sorgente. La tracciabilità e la catena di custodia di ogni azione intrapresa durante le fasi d'acquisizione è importante, perchè a distanza di mesi o anni, bisogna capire cosa è stato fatto esattamente, da dove provengono quei dati, chi, come e quando li ha riversati, altrimenti si potrebbero perdere delle informazioni a contorno che possono spiegare la presenza degli stessi.

Infine, anche estrapolare solo dati parziali e non immersi nel loro "ecosistema", non è una buona pratica, pensiamo ad una serie di file copiati e masterizzati su dischi ottici o copiati su pendrive o hard disk esterni, questi potranno fornire informazioni solo sui loro contenuti ed i loro metadati interni, ma non ci saranno altre informazioni sulla loro storia nel dispositivo d'origine, come anche quando non si incrociano i dati tra pendrive, telefoni, hard disk esterni, ecc.. Con il computer "madre" al quale erano connessi, che potrebbero ribaltare completamente la storia della creazione, modifica e consultazione ed origine dei dati presenti su questi dispositivi esterni.

Insomma, spesso sembra che adottare delle procedure sia solo un esercizio di stile, si pensa che se comunque si trovano le evidenze non importa se queste derivano da un'acquisizione o un'analisi che non ha seguito le regole della digital forensics, perchè sono evidenze che si autosostengono, questa errata percezione e senso di sicurezza può creare degli equivoci ed errori anche gravissimi e sollevare dubbi e si sa che "*in dubio pro reo*" oppure, senza adeguate contestazioni, qualche innocente potrebbe pagare salato.
Anche nella fase di analisi, le regole di base sono quelle di giustificare ogni passaggio con dei dati certi, come bibliografia ufficiale, esperimenti ripetibili, utilizzo di più software d'analisi per il confronto delle risultanze, richieste di chiarimento ai produttori di software ed hardware, questo per evitare di scrivere cose "a naso" o "ad esperienza" o "a memoria".
Anche il concetto più semplice va verificato, va giustificato, pure quando sembra una cosa scontata e di comune conoscenza, perchè potremmo incorrere in errori senza saperlo o per superficialità.
Se si afferma qualcosa, anche di banale, va dimostrata, che sia la data di creazione di un file, l'origine di un indirizzo IP o un funzionamento di un Social Network, va sempre cercata la fonte ufficiale che attesti ciò che affermiamo ed in mancanza di tale fonte, si può anche percorrere la via sperimentale ed il reverse engineering, documentando tutto ai fini di ripetibilità.

Quindi sono importanti o no le procedure della digital forensics? Sì, lo sono, ma forse non sono percepite così importanti, come il fumo di sigaretta, che si sa che fa male, ma non è percepito come un pericolo reale ed imminente, forse perchè non vi è un'interlocuzione chiara tra i due mondi, quello giuridico e quello informatico, non vi è cultura uniforme tra gli informatici forensi, quindi si agisce spesso con superficialità e si chiude un occhio su alcune mancanze, quasi fossero dei dettagli ininfluenti, che però in mano ad un buon consulente informatico forense, potrebbero diventare delle armi potenti, ma anche qui l'effetto dell'arma potrebbe non esser capito ed ignorato, perchè troppo raffinato, troppo sottile, troppo tecnico, troppo improbabile o "fantascientifico", da invalidare quei quattro file presi "alla carlona" dal computer dell'indagato.

La soluzione? La cultura, la cultura e la cultura, bisogna creare cultura uniforme in tutti gli attori della Giustizia e della Tecnica, quindi dalla P.G., agli avvocati, magistrati, tecnici informatici, solo così si parlerà un linguaggio comune e si potrà riservare il giusto rispetto e cautela che meritano le evidenze digitali. ©