

# Le tracce digitali e la fragilità del dato: rilievo, custodia e analisi

Università degli Studi di Milano  
Dipartimento di Informatica e Comunicazione  
M. Monga

COMPUTER FORENSICS, INVESTIGAZIONE PENALE E CRIMINALITÀ  
TECNOLOGICA - L.E.F.T. (Legal Electronic Forensics Team)  
<http://www.avanzata.it/left> - Milano, 23 novembre 2006

## Sintesi dell'intervento

A cura del Magg. CC Ing. Marco Mattiucci - Roma, 29 novembre 2006

La prima parte dell'intervento ha introdotto il concetto di *informazione* e *bit* quale unità minima di rappresentazione della prima. Si è parlato quindi di *codifica* delle informazioni e di *semantica* delle stesse, quest'ultima collegata al *modello* ed alla *interpretazione* sulla base delle quali il dato digitale viene trattato.

In relazione ai segnali fisici codificati secondo *grandezze analogiche* si è visto che segnale misurato e grandezza rappresentativa (informazione analogica) sono fortemente correlati se non omologhi. Nel caso digitale la codifica in dato binario si stacca spesso totalmente dalla natura fisica delle grandezze che costituiscono l'informazione, da cui si è sottolineato che *un minimo errore nei dati può portare ad enormi errori di interpretazione* degli stessi, fatto questo determinante da un punto di vista forense.

Dalla codifica il discorso si è ampliato alle *macchine*, intese come sistemi tradizionalmente solo *hardware* in cui la struttura fisica determinava il tipo di attività eseguibile. In definitiva, tradizionalmente, le macchine interpretavano direttamente una informazione legata alla loro

struttura fisica. Ad oggi le *macchine digitali* sorpassano tale concetto mediante l'uso del *software*, ossia di una parte variabile che ne cambia radicalmente le funzioni eseguibili anche in tempo reale (per la macchine tradizionali il cambio di funzione è legato invece alla modifica dell'hardware). Le macchine digitali, poi, si è visto essere in grado di ospitare al loro interno più *macchine virtuali* che operano in contesti e svolgono funzioni prettamente diversificate tra loro.

Nella seconda parte dell'intervento ci si è calati nella dimensione forense riportando innanzitutto la definizione di *digital evidence* tipicamente anglosassone data dal prestigioso SWDGE (Scientific Working Group for Digital Evidence): *informazione con valore probatorio che sia memorizzata/trasmessa in formato digitale*. Un fatto sottolineato è stato che anche molti atti di PG ad oggi viaggiano in formato digitale per poi essere impiegati in dibattimento, da cui risponderebbero alla definizione di cui sopra.

A questo punto si è analizzato il concetto: “...*questo file è una prova di..*”. La debolezza di questa affermazione si basa ovviamente sul fatto che un file in se stesso può essere interpretato in diversi modi, mediante diversi software, per cui indicare un file senza il sistema operativo, il software creatore, ecc. rappresenta una forte limitazione nel suo impiego probatorio.

Si è quindi discusso il concetto di *valore probatorio* di una digital evidence indicando i seguenti fattori:

- *Autenticità*: bisogna identificare esattamente la provenienza dei dati digitali indicati come prova – questo è generalmente complicato e non sempre possibile, inoltre esistono casi in cui l'assenza di meccanismi di controllo sui dati e la loro correttezza o la loro intrinseca debolezza non rendono neanche ipotizzabile tale forma di garanzia. Sono stati presentati a tale proposito il codice di parità, quello hamming, gli hash ed in particolare le *collisioni dello MD5* (nel caso specifico sono stati presentati due file .ps palesemente diversi con lo stesso codice MD5 associato – per un approfondimento della questione si rinvia al link di questo sito <http://www.marcomattiucci.it/md5.php>).
- *Integrità*: stabilire il metodo di conservazione dei dati una volta acquisiti (ad es. il supporto di memorizzazione);
- *Veridicità*: usare gli strumenti di analisi corretti al fine di avere una interpretazione dei dati repertati valida e completa da presentare in dibattimento.
- *Completezza*: bisogna saper escludere i dati ininfluenti e saper reperire tutti quelli inerenti il caso investigativo. Entrambe le attività sono molto pesanti e complesse, inoltre non esistono strumenti tecnici che consentono una tale completezza in automatico (senza l'ausilio del tecnico). Questo implica che per lo svolgimento delle analisi tecniche

sarebbe consigliabile l'impiego di più strumenti apparentemente equivalenti.

- *Legittimità*: conformità alle leggi italiane dei metodi e degli strumenti.

Sono state presentate allora le tipiche fasi del forensic computing:

- *Identificazione*: rilevare i dati di effettivo interesse;
- *Acquisizione*: prelevare i dati identificati. A tale proposito si è valutata l'acquisizione "live" di dati in macchine attive evidenziando i due aspetti della *irripetibilità* ed *alterazione dei dati* che ne conseguono, nonché la necessità di documentare, anche in automatico e con precisione i passi necessari a tale attività.
- *Analisi*: attività di laboratorio o "live" orientata all'individuazione ed estrazione delle prove dai dati acquisiti.
- *Presentazione*: momento di interfaccia tra il legale ed il tecnico in cui i dati digitali vengono ulteriormente interpretati nel particolare ambito dell'indagine e delle leggi.

La discussione è quindi terminata con due aspetti importanti quali:

- *L'ammissibilità degli strumenti e dei risultati*: sebbene perfino il NIST (National Institute of Standards and Technology) abbia cercato di testare ufficialmente i più impiegati tool informatico forensi l'unico risultato è stato un test completo solo per le versioni più obsolete e dopo diversi anni di lavoro. Al contrario per le versioni più attuali non è stato possibile neanche completare un piano di testing da cui il controllo completo non risulta possibile neanche da un punto di vista teorico, data la varietà delle situazioni investigative che possono presentarsi.
- Il *Trusting Computing* come base per il forensic computing: ipotizzando una macchina digitale C, un utente U ed un proprietario P della C è possibile stabilire le seguenti regole mediante meccanismi di crittazione: (a) *P decide cosa si può fare con C* – (b) *U può fare con C solo quello che ha precedentemente deciso P*. Questa modalità consentirebbe di controllare in maniera sicura lo svolgimento delle indagini tecniche.