LA CANCELLAZIONE SICURA DEI DATI TRA MITO E REALTÀ

di Paolo Reale

L'utente informatico, anche non particolarmente esperto, è consapevole che sia possibile, attraverso l'utilizzo di opportuni metodi e strumenti, il recupero dei dati cancellati dai supporti di memorizzazione. Non altrettanto si può dire del problema simmetrico: la cancellazione "sicura" dei dati (*wiping* in inglese), ovvero far sì che i dati eliminati siano effettivamente irrecuperabili, con certezza. È un quesito che, necessariamente, si devono porre le organizzazioni private e pubbliche che trattano i dati personali e sensibili dei propri clienti o dei cittadini, oltre a chiunque voglia tutelare la propria *privacy*.
A volte anche chi utilizza e conosce la tecnologia non ha sufficiente chiarezza di cosa è possibile realmente recuperare dopo una cancellazione e cosa no: questo dubbio è indirettamente alimentato anche dalle procedure americane emanate per le forze militari e le agenzie di *intelligence*, come la DoD 5220.22-M[5] del 1995 (ormai obsoleta) e altre, in cui si trova indicazione come in caso di informazioni "*top secret*" l'azione di "*sanitization*" preveda di "disintegrare, incenerire, polverizzare, sminuzzare, o fondere" i supporti. Queste azioni sono in realtà da considerarsi come soluzioni rapide ed efficaci, difficili da sbagliare e alla portata anche di chi non avrebbe le necessarie competenze tecniche per poter effettuare una verifica puntuale dell'avvenuta cancellazione definitiva dei dati, condizione necessaria per tali tipologie di informazione.

Vi è comunque una diffusa percezione del fatto che le Agenzie governative e di *intelligence* avrebbero la capacità e le tecnologie per recuperare dati cancellati, anche se sovrascritti più volte, tramite l'utilizzo di microscopi a forza magnetica (MFM), e/o strumenti più evoluti. Uno dei principali portavoce di questa tesi è Gutmann[1], che sosteneva come "*le organizzazioni di intelligence hanno molta esperienza nel recupero delle immagini*" relative ai diversi strati di sovrascrittura sui dischi magnetici. Questa convinzione si basa, sinteticamente, sul presupposto che quando la cifra "1" viene scritta sul disco l'effetto reale è più vicino ad ottenere un valore di 0,95 se è in sovrascrittura ad un precedente valore "0", e un valore di 1,05 quando un "1" viene sovrascritto ad un precedente "1". **Per questo Gutmann propose un algoritmo di cancellazione (c.d. "metodo Gutmann") basato sulla sovrascrittura di ben 35 diversi *pattern* digitali**, finalizzati alla completa distruzione di ogni possibile traccia precedente. Detto algoritmo si trova implementato in molti *softwares* di *wiping* disponibili in rete, e le sue '35 passate' sono diventate nel tempo una sorta di rituale Voodoo della cancellazione (come detto dallo stesso Gutmann).

Pochi hanno tuttavia intrapreso la strada di una puntuale verifica delle conclusioni di Gutmann, come ha fatto Feenberg[2], che dopo aver accuratamente analizzato tutta la bibliografia citata[1], ed aver approfondito il tema con successive ricerche, conclude che "*le affermazioni di Gutmann appartengono alla categoria delle leggende urbane*", in quanto non supportate da concrete evidenze di effettivo recupero dei dati.

Sembrerebbe poco utile soffermarsi su questi aspetti contraddittori relativi alla possibilità di recupero successivo dei dati sovrascritti, tuttavia è fondamentale comprendere che per molto tempo questi temi sono stati oggetto di posizioni quasi preconcette, e in mancanza comunque di oggettive dimostrazioni sull'effettiva recuperabilità dei dati sovrascritti si deve rilevare come si è comunque mantenuta in essere l'idea che **un alto numero di sovrascritture garantisse l'impossibilità nel recupero dell'informazione originaria. Tale idea è in realtà falsa**, addirittura fuorviante come vedremo meglio nel seguito.

Nel tentativo di dirimere la questione, è intervenuto nel 2008 l'articolo "*Overwriting Hard Drive Data: The Great Wiping Controversy*", di Wright e altri[3], che nell'introduzione spiega come l'approccio di Gutmann sulla possibilità di recupero dalle sovrascritture "si può dimostrare falso e che, di fatto, vi è una distribuzione basata sulla densità magnetica" tale che il differenziale statistico nei modelli di scrittura è troppo grande per consentire il recupero di dati sovrascritti. In particolare viene smentita l'idea[1] che "*ogni traccia contiene l'immagine di tutto quello che è stato scritto, ma che il contributo di ciascun strato diventa progressivamente più piccolo*" se si procede a ritroso nel tempo. Questo è un *"fraintendimento della fisica del funzionamento dei drive e della magneto-risonanza. Non vi è infatti alcuna componente di tempo e l'immagine non è a strati."*

Viene dunque sfatato il mito relativo alle modalità di scrittura dei dati su un disco rigido, in cui la scrittura digitale corrisponde ad un'operazione digitale: l'unità scrive livelli analogici che hanno valori probabilistici, compresi in un certo *range*. Come risultato, non c'è differenza nella scrittura di un valore 0,90 o 1,10 del potenziale magnetico: a causa di fluttuazioni di temperatura, umidità, ecc. il valore scritto sul supporto varierà per ogni singola passata del ciclo di scrittura. In altre parole, non c'è modo di determinare se un valore, ad esempio, di "1,06" sia dovuto ad una scrittura precedente o ad una fluttuazione della temperatura. Inoltre il flusso magnetico su un disco decade lentamente nel tempo, distorcendo ulteriormente i risultati e aumentando il livello di incertezza di un eventuale recupero. In conclusione quindi il lavoro di Wright[3] dimostra sperimentalmente che **i dati cancellati con sovrascrittura non possono essere recuperati, neppure con l'uso di un microscopio o altri metodi conosciuti**. Le stesse conclusioni erano peraltro già evidenziate dal CMRR, che nel documento[4] rilevava come nelle "*unità di oggi, sovrascritture multiple non sono più efficaci di una singola sovrascrittura*".
Stabilito quindi che un ciclo di sovrascrittura risulti già efficace, la domanda da porsi è come realizzare questa operazione

in modo da essere certi che ogni zona del disco sia interessata dall'azione.
La sovrascrittura effettuata da uno strumento *software* esterno al supporto stesso può, infatti, fallire in alcune circostanze specifiche, oppure non cancellare i blocchi riallocati (blocchi difettosi), o non rilevare partizioni ulteriori. Per questo, dietro specifica richiesta proprio del CMRR, è stato introdotto il comando "*Secure Erase*" (SE) allo *standard* aperto ANSI per il controllo del disco rigido: tale comando consiste sostanzialmente nel metodo della sovrascrittura del disco, ed è implementato all'interno del *firmware* stesso, rendendolo così immune da qualunque attacco di *software* maligno o altre *utility*. È implementato su tutte le interfacce ATA prodotte dopo il 2001.

Quanto sopra sicuramente non esaurisce il problema più generale, che riguarda l'approccio complessivo alla cancellazione sicura per ogni tipo di supporto di memorizzazione e quali siano i criteri e le soluzioni da adottare. In quest'ottica, il documento più attuale ed esaustivo è quello prodotto dall'organismo americano NIST 800-88[(6)].

Il NIST 800-88 innanzitutto definisce quali siano le tipologie di cancellazione (*sanitization*), da adottare sulla base di una valutazione dell'effettivo grado di riservatezza del dato, del rischio correlato, dei costi legati al metodo scelto e dell'impatto ambientale:

1. **Scarto (*disposal*)**: ove non siano presenti dati ritenuti in qualche modo importanti, è contemplato anche il semplice smaltimento come rifiuto;
2. **Pulizia (*clearing*)**: si tratta di un livello medio di cancellazione, destinato ad evitare il recupero dei dati attraverso gli strumenti software tipicamente utilizzati a questo scopo. Quasi sempre è sufficiente la sovrascrittura dello stesso.
3. **Eliminazione (*purging*)**: si tratta di un livello di protezione più elevato, orientato a proteggere le informazioni più sensibili da un possibile tentativo di recupero in laboratorio. È possibile eventualmente contemplare l'utilizzo di un *degausser* ovvero uno strumento *hardware* in grado di produrre un campo magnetico variabile in intensità e direzione, in modo da eliminare i dati (dai supporti di tipo magnetico evidentemente) e rendere inutilizzabile la memoria. Tipicamente questa modalità rende anche inservibile il drive oggetto dell'azione di cancellazione.
4. **Distruzione (*destroying*)**: la distruzione del supporto è indubbiamente la soluzione più estrema, ma anche inevitabile per alcuni tipi di dispositivi, come quelli ottici (CD, DVD). Le azioni suggerite sono, in dipendenza dal tipo di memoria, la triturazione, la polverizzazione, la fusione e l'incinerazione.

Va puntualizzato, tuttavia, come l'evoluzione e la disponibilità di supporti con nuove caratteristiche e/o tecnologie possa generare situazioni in cui le modalità descritte di cancellazione non siano completamente valide. Infatti, **nei dischi a stato solido (SSD) il comando di *Secure Erase* fallisce in alcuni drive per una non corretta implementazione dell'algoritmo da parte del produttore**, lasciando addirittura i dati (per alcuni drive) praticamente intatti sul disco. Le tecniche di cancellazione basate sui *softwares* di *wiping* sono invece tendenzialmente efficaci (ma non sempre), mentre quelle per la cancellazione del singolo *file* falliscono sistematicamente. A questi risultati giunge un lavoro recentemente pubblicato da Wei e altri[(7)]: **è dunque errato assumere, a priori, che le tecniche di cancellazione utilizzate nei dischi rigidi possano funzionare nello stesso modo e con la stessa efficacia per i dispositivi a stato solido**. Ciò dipende da diversi fattori, tra cui *in primis* la presenza, nei drive SSD, di un livello di gestione intermedio, il *Flash Translation Layer* (FTL), che virtualizza di fatto il livello *hardware*. Nello sperimentare soluzioni possibili per la cancellazione sicura, il *team* indirizza comunque verso i produttori la necessità di implementare direttamente sul supporto SSD le funzionalità per realizzare in modo efficace le operazioni di *wiping*, sul singolo *file* come sull'intero disco.

In conclusione, la sicurezza dei dati è diventata nel tempo uno dei problemi più spinosi per le aziende, gli Enti governativi ed i professionisti che trattino informazioni personali o sensibili, sia nelle fasi di esercizio operativo che nella loro successiva distruzione. In particolare la cancellazione sicura (*sanitization*) deve garantire l'effettiva irrecuperabilità delle informazioni riservate, e per questo il riferimento più esaustivo sul tema è il documento NIST 800-88[(6)] che, pur costituendo un valido supporto anche al privato che voglia tutelare al meglio la propria *privacy*, è probabilmente eccessivo: in questa situazione, è molto probabile che i normali programmi di *wiping*, disponibili anche gratuitaente su Internet, o ancor meglio la funzionalità di *Secure Erase* disponibile nei dischi fissi di uso comune, costituiscano rimedio ampiamente efficace alla soluzione del problema.

In tutti i casi, sebbene sia provata l'inutilità di procedure di cancellazione basate su sovrascritture multiple dei supporti, è importante ricordare che:

- non può mai essere sottovalutata la fase di verifica delle operazioni effettuate, a garanzia della qualità dell'azione svolta,
- deve essere posta attenzione nei casi in cui si utilizzino nuove tipologie di dispositivi, specie se basati su nuove tecnologie, perché non vale il principio di estendere automaticamente il patrimonio di esperienza maturato nella gestione degli altri dispositivi noti. ©

NOTE

1. "*Secure Deletion of Data from Magnetic and Solid-State Memory*", Peter Gutmann, Sixth USENIX Security Symposium Proceedings, San Jose, California, July 22-25, 1996.
2. "*Can Intelligence Agencies Read Overwritten Data?*", Daniel Feenberg, National Bureau of Economic Research, 2003 (con aggiornamenti fino al 2011)
3. "*Overwriting Hard Drive Data: The Great Wiping Controversy*", Wright, Kleiman, Sundhar, 2008, ICISS 2008: 243-257.
4. "*Tutorial on Disk Drive Data Sanitization*", Gordon Hughes e Tom Coughlin, settembre 2006.
5. "*U. S. Department of Defense*. 5220.22-M *National Industrial Security Program Operating Manual*", gennaio 1995.
6. "*Guidelines for Media Sanitization*", Recommendations of the National Institute of Standards and Technology, NIST Special Publication 800-88, settembre 2006.
7. "*Reliably Erasing Data From Flash-Based Solid State Drives*", Wei, Grupp, Spada, Swanson, FAST '11: 9th USENIX Conference on File and Storage Technologies, 2011. ◊