Informatica e diritto, XLI annata, Vol. XXIV, 2015, n. 1-2, pp. 345-360

Acquisizione remota di *personal object stores*: questioni di ripetibilità

CORRADO FEDERICI*

SOMMARIO: *1. Introduzione – 2.* Cloud Data Imager *– 3. Accertamenti tecnici non ripetibili – 4. Classificazione delle acquisizioni di evidenze digitali – 5. Ripetibilità delle acquisizioni di* personal object store *– 5.1. Dropbox – 5.2. Microsoft OneDrive – 6. Conclusioni*

1. INTRODUZIONE

L'avvento del *cloud computing* ha realizzato una visione dell'informatica nata negli anni Sessanta del secolo scorso, secondo cui le risorse computazionali vengono vendute a consumo similmente alle utenze domestiche come l'energia elettrica o il gas. Ciò è stato possibile anche grazie a un radicale ripensamento degli approvvigionamenti di materiali e delle modalità di gestione dei *data center* da parte dei pionieri del *cloud* come Amazon o Google, che hanno introdotto livelli di automazione talmente spinti da ridurre il costo orario dei propri servizi a pochi centesimi di dollaro[1]. A tal riguardo, un ruolo centrale è stato giocato dai *personal object stores* - POS, architetture informatiche che consentono la memorizzazione affidabile dei contenuti degli utenti come documenti o video, grazie ad una struttura distribuita in grado di scalare in modo virtualmente infinito utilizzando hardware a basso costo[2]. Se da un lato tutto ciò ha comportato indubbi benefici per l'utente finale che può disporre di hard disk virtuali sempre accessibili mediante Internet, dall'altro ha parimenti aumentato la possibilità che tali aree di *storage* possano essere utilizzate per ospitare *digital evidence* - DE, ossia dati di possibile interesse nel processo penale poiché pertinenti a reati sia prettamente informatici sia "tradizionali". Sfortunatamente per il *digital evidence specialist* - DES le architetture distribuite possono comportare difficoltà insormontabili quando si tratti di ricostruire una visione d'insieme perché ogni *object* (per esempio un video) può essere costituito da numerosi file che vengono sparsi

* L'A. è collaboratore presso il CIRSFID di Bologna.

[1] Secondo le tariffe di Amazon *Elastic Compute Cloud*, il costo orario di una macchina virtuale con una CPU, 2 Giga Bytes di RAM e sistema operativo Linux (t2.small) è di 3 centesimi di dollaro + IVA (Regione Ue Francoforte, 10/2016).

[2] Dropbox, Google Drive e Microsoft OneDrive sono tra le soluzioni più note.

© Edizioni Scientifiche Italiane ISSN 0390-0975 ISBN 978-88-495-3285-2

all'interno di una quantità di server anche molto numerosa. Ciò comporta la quasi certa impossibilità di applicare le classiche procedure della *digital forensics* (DF) per il repertamento fisico delle memorie di massa[3], sia perché può risultare arduo individuare tra tutti i possibili server quali contengano i dati di interesse sia anche per difetti di giurisdizione. È però possibile valutare una differente strategia che consiste nell'eseguire la copia logica[4] di contenuti e metadati[5] operando da remoto. Infatti i POS sono raggiungibili mediante un'interfaccia di programmazione basata su *web service*[6] che consentono ad applicazioni create dall'utente di compiere operazioni sugli oggetti come listare il contenuto di una cartella o visualizzare un documento. Un aspetto di fondamentale interesse per la DF risiede nella proprietà di *immutabilità*. Infatti, a differenza dei normali file per i quali le modifiche comportano l'inserimento o la cancellazione di un certo numero di byte nella posizione desiderata, ogni alterazione di un oggetto[7] causa dietro le quinte il caricamento nel POS di una nuova versione dello stesso e conseguente memorizzazione della versione precedente[8]. Si intuisce che tale meccanismo, a differenza di quanto normalmente accade per i *file system* tradizionali, potrebbe consentire di ricostruire il percorso delle alterazioni occorse ad un oggetto, sia come contenuto sia come linea del tempo, con conseguenti potenziali benefici per il DES in termini di conoscenza del caso.

Lo studio dell'interfaccia di programmazione esposta dalle principali piattaforme di *personal storage*[9] ha dimostrato che essa può essere idonea per la realizzazione di innovativi strumenti di DF in grado di creare su dispositivo locale una copia logica fedele dell'area dati in *cloud*, stante la possibilità

[3] Si tratta della copia sequenziale a basso livello di tutti i settori del reperto, detta anche copia *bit a bit*.

[4] Copia a livello di singolo oggetto o cartella di un *file system* (a differenza della copia *bit a bit* restano esclusi i settori non allocati della memoria di massa).

[5] I "dati relativi ai dati", come la data di modifica di un oggetto, possono contribuire a collocare temporalmente le azioni dell'utente cui si riferiscono e risultare talora più utili dei contenuti stessi.

[6] I WS sono canali di programmazione basati sui protocolli standard di Internet (HTTPS, XML, JSON) attraverso cui è possibile inviare comandi ad una piattaforma e ricevere contenuti, metadati ed informazioni di varia natura.

[7] Si pensi, per esempio, ad un documento di testo all'interno della cartella Dropbox del profilo utente.

[8] Il periodo di mantenimento delle revisioni storiche potrebbe dipendere dal tipo di servizio sottoscritto.

[9] Si veda, ad es.: *Dropbox, Core API - endpoint reference*, 2016.

ISSN 0390-0975 ISBN 978-88-495-3285-2 © Edizioni Scientifiche Italiane

di accedere in modalità protetta da scrittura e recuperare sia i dati presenti sia quelli cancellati[10]. Ambiti di utilizzo ipotizzabili per il repertamento remoto comprendono, ma non esauriscono, gli accertamenti tecnici compiuti durante la fase delle indagini preliminari[11] in cui, per esempio, i consulenti tecnici del pubblico ministero (PM) potrebbero aver recuperato dall'analisi di materiale informatico in uso all'indagato le credenziali di accesso all'area di POS. Lo scopo di questo contributo è pertanto quello di descrivere il funzionamento di *Cloud Data Imager*[12], applicazione progettata per compiere il repertamento forense di servizi POS come Dropbox, Google Drive e Microsoft OneDrive da una postazione di lavoro connessa ad Internet e discutere uno degli aspetti maggiormente di frontiera della DF da un punto di vista normativo: la ripetibilità degli accertamenti nel contesto di un'acquisizione remota. Infatti, nella pratica forense tradizionale, la copia fisica *bit a bit* delle memorie viene normalmente considerata un'attività irripetibile sia perché *shock* di natura meccanica, termica o la semplice connessione del reperto digitale con l'alimentazione di rete potrebbero causarne il danneggiamento fisico sia perché la mancata applicazione delle migliori pratiche della DF come la protezione dalla scrittura potrebbero causare l'alterazione del contenuto informativo del reperto. Tale eventualità di fatto non si pone invece nel caso delle infrastrutture *cloud storage*, che sono progettate per tollerare guasti dell'hardware anche importanti ed hanno livelli di disponibilità anche superiori al 99,9% nel corso dell'anno[13]. Verrà pertanto osservato se la ripetizione dello stesso esperimento di acquisizione di dati dal *cloud* produrrà lo stesso risultato. In caso contrario, nei casi di accertata irripetibilità in senso stretto dovuta al cambiare ad ogni nuovo accertamento del codice di *hash* calcolato sul complesso dei dati acquisiti, verrà valutato se le modificazioni rilevate siano determinanti per classificare l'attività tecnica come irripetibile o se essa non possa essere considerata come *sostanzialmente ripetibile* nel senso che le alterazioni tra due esperimenti successivi possano essere giudicate irrilevanti nell'ambito processuale poiché esse non impattano negativamente sulle fondamentali caratteristiche di affidabilità e sufficienza dell'evidenza digitale.

[10] *Cloud Connector* di *F-Response* ed *Extractor for Clouds* di *Oxygen Forensics* sono due esempi di applicazioni commerciali.

[11] Cfr. artt. 326 ss. c.p.p.

[12] C. Federici, *Cloud Data Imager: A Unified Answer to Remote Acquisition of Cloud Storage Areas*, in "Digital Investigation", Vol. 11, 2014, n. 1, pp. 30-42.

[13] Indisponibilità inferiore a 9 ore annue.

© Edizioni Scientifiche Italiane ISSN 0390-0975 ISBN 978-88-495-3285-2

2. Cloud Data Imager

Cloud Data Imager[14] è una nuova applicazione forense per l'acquisizione da remoto di POS con le seguenti funzionalità:

- autenticazione dell'utente secondo lo standard OAUTH[15] versione 2.0;
- recupero delle informazioni di account come nome e identificatore univoco;
- accesso in sola lettura;
- navigazione di tipo "esplora risorse di Windows" con elencazione del contenuto di una cartella, comprese, se disponibili, le sottocartelle, gli oggetti cancellati e le revisioni[16] di questi ultimi;
- visualizzazione del contenuto degli oggetti;
- creazione dell'immagine logica[17] di un albero completo di directory in una cartella locale della postazione forense o in un volume virtuale esportato in formato EWF[18];
- calcolo di un codice di *hash* cumulativo (*inner hash*) per tutti i dati ricevuti dalla piattaforma di *cloud storage* e generazione di un elenco di tutti i file scaricati con relativo codice di *hash* (*hash list*);
- produzione di un file di log in cui vengono registrate tutte le attività dell'utente, lo scambio dati con il server remoto e le eventuali condizioni di errore derivanti, per esempio, da interruzione della connettività di rete.

La fig. 1 mostra schematicamente l'architettura a blocchi di *Cloud Data Imager*, mentre la fig. 2 riporta l'area di lavoro relativa alla navigazione di un account denominato "Seminario Tenerife".

Al momento è disponibile l'accesso a Dropbox, Box, Google Drive e Microsoft OneDrive, ma la lista dei provider disponibili è in costante aggiornamento. L'applicazione utilizza una libreria che maschera le differenze di implementazione tra i vari *cloud service provider* (CSP) ed espone un'interfaccia unificata.

[14] Prototipo reperibile all'indirizzo doi.org/10.13140/RG.2.2.10520.19206.

[15] Con il protocollo OAUTH non è necessario configurare le credenziali utente nell'applicazione. Essa invece riceve un "lasciapassare temporaneo" (*access token*) per accedere alle risorse dopo che l'avente diritto è stato identificato e ha esplicitamente autorizzato l'accesso.

[16] La modifica di un oggetto (per esempio un documento) comporta il caricamento della nuova versione e la storicizzazione di quella precedente sotto forma di revisione.

[17] Copia di tutti gli oggetti con i relativi metadati originari.

[18] L'*Expert Witness Format* è uno dei formati principali in cui vengono salvati i file immagine derivanti dalla copia *bit a bit* di una memoria di massa.

ISSN 0390-0975 ISBN 978-88-495-3285-2 © Edizioni Scientifiche Italiane

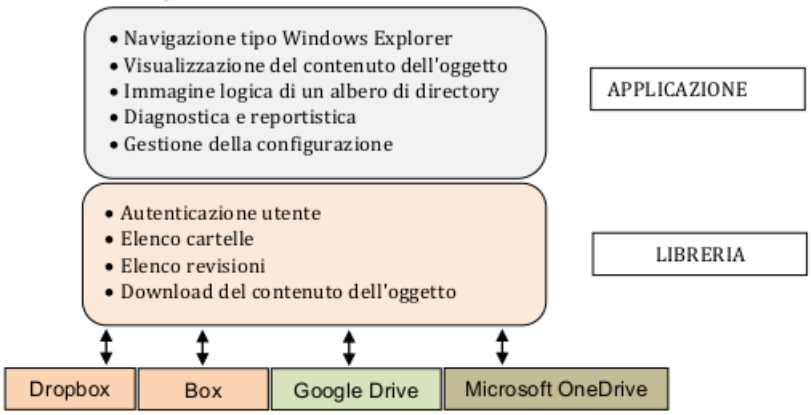

Fig. 1 – Funzioni ed architettura di Cloud Data Imager

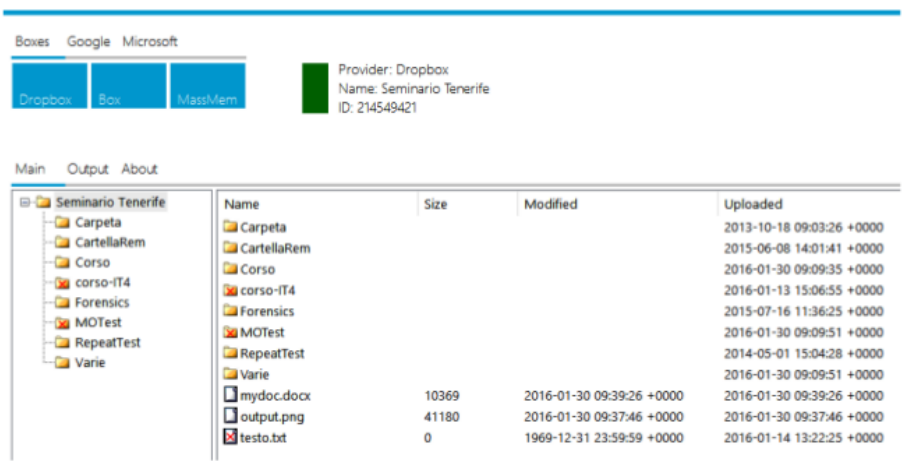

Fig. 2 – Dashboard *di CDI*

© Edizioni Scientifiche Italiane ISSN 0390-0975 ISBN 978-88-495-3285-2

Nome	*Categoria*	*Descrizione della funzione*
getAuthorizeUrlV2	Autenticazione/Autorizzazione	Implementa il protocollo di autenticazione / autorizzazione OAUTH 2.0
getAccessTokenV2	Autenticazione/Autorizzazione	Implementa il protocollo di autenticazione / autorizzazione OAUTH 2.0
getAccountInfo	Informazione	Ottiene il nome utente e l'identificatore univoco
listFolder	Navigazione e immagine logica	Ottiene i metadati di una cartella, compresi quelli relativi ad oggetti e sottocartelle contenute
listFileRevisions	Navigazione e immagine logica	Ottiene i metadati di tutte le revisioni di un oggetto
getFileContent	Navigazione e immagine logica	Opera il download del contenuto grezzo dell'oggetto

Tab. 1 – Lista dei metodi esposti dalla libreria

In tab. 1 sono riportati i metodi esposti dalla libreria e utilizzati dall'applicazione.

3. Accertamenti tecnici non ripetibili

Il codice di procedura penale italiano disciplina queste attività all'art. 360, con riguardo alle funzioni di competenza del PM durante la fase delle indagini preliminari. Quando è necessario compiere accertamenti di natura tecnica concernenti luoghi, persone o cose il cui stato potrebbe essere soggetto a modificazioni, il PM deve dare notizia a tutte le parti del processo del luogo e dell'ora in cui verrà nominato il proprio consulente tecnico a mente dell'art. 359. Le parti vengono anche informate circa la possibilità di farsi assistere da propri tecnici di fiducia che hanno il diritto di partecipare a tutte le successive sessioni tecniche, fare commenti e verificare la correttezza delle operazioni. Di estremo rilievo nei processi di creazione dell'immagine della DE è poi l'art. 117 delle disposizioni di attuazione del c.p.p. che estende il campo di validità dell'art. 360 anche ai casi in cui sia l'esecuzione dell'accertamento stesso ad introdurre alterazioni che altrimenti potrebbero non verificarsi affatto. Per esempio, nel caso di una memoria di massa, se conservata in una busta antistatica e in condizioni ambientali idonee, è lecito attendersi che non si verificheranno modificazioni di sorta in un periodo di tempo ragionevole, ma il fatto stesso di collegarla all'alimentazione elettrica o il mancato utilizzo di un dispositivo di protezione dalla scrittura durante le operazioni

ISSN 0390-0975 ISBN 978-88-495-3285-2 © Edizioni Scientifiche Italiane

di copia potrebbero causarne l'inutilizzabilità in dibattimento. Una procedura condivisa garantisce quindi la tutela degli interessi delle parti sin dalla fase delle indagini preliminari poiché in futuro potrebbe essere impossibile ripetere alcune attività, a causa delle irreparabili modificazioni sopravvenute.

4. Classificazione delle acquisizioni di evidenze digitali

Una corretta collocazione di tutte le possibili fattispecie di acquisizioni riguardanti la DE è fondamentale per valutarne gli effetti nell'ambito del processo penale. Gli accertamenti di cui si è discusso nel precedente paragrafo sono solo una parte delle possibili attività tecniche che possono avere luogo nella fase delle indagini preliminari con differenti livelli di garanzia per la difesa[19]. A mente dell'art. 354 c.p.p. la polizia giudiziaria (PG) può compiere accertamenti urgenti e rilievi, raccogliendo materiale sulla scena del crimine, quando vi sia il rischio che questo possa alterarsi o disperdersi prima che il PM abbia assunto la direzione delle indagini. Una tale situazione comporta l'attenuazione delle tutele per l'indagato rispetto a quanto previsto dall'art. 360 perché al difensore è data la facoltà di essere presente, ma senza il diritto di essere avvisato in anticipo (art. 356 c.p.p.). Il codice stesso comunque si occupa di controbilanciare questo squilibrio stabilendo che le acquisizioni di DE avvengano garantendo la salvaguardia dell'integrità dell'originale e, se possibile, procedendo alla copia immediata e fedele su supporti non riscrivibili. Ancora meno caratterizzati da profili di garanzia sono gli atti compiuti dalla PG su delega specifica del PM a norma dell'art. 370 c.p.p., che tuttavia non prevede la presenza della difesa[20]. Infine, fuori dai casi di urgenza, la necessità di assicurazione delle fonti di prova consente alla PG, su delega o per iniziativa, di avvalersi di consulenti esterni per svolgere atti o operazioni di particolare rilievo tecnico (art. 348 c.p.p.). Purtroppo il codice non fornisce una definizione di tutte queste diverse tipologie di atti, né distingue chiaramente tra accertamenti e rilievi[21]. Questa lacuna è stata tuttavia colmata dalla giurisprudenza e dalla dottrina secondo cui un *rilievo* è un'attività di mera osservazione, identificazione e raccolta di materiale prodromica

[19] S. Fasolin, *La copia di dati informatici nel quadro delle categorie processuali*, in "Diritto Penale e Processo", 2012, n. 3, pp. 372-380.

[20] *Ibidem*.

[21] S. Sottani, *Rilievi e accertamenti sulla scena del crimine*, in "Archivio Penale", 2011, n. 3, pp. 777-784; F. Casasole, *Le indagini scientifiche nel processo penale*, Roma, Dike Giuridica, 2013.

© Edizioni Scientifiche Italiane ISSN 0390-0975 ISBN 978-88-495-3285-2

all'*accertamento*, il quale invece comporta una meditata valutazione dei dati, con conseguente produzione di un responso[22]. La riconducibilità delle acquisizioni di DE a mero rilievo è stata più volte sostenuta in pronunciamenti della Corte di Cassazione secondo cui:

1. l'estrazione di dati da un elaboratore è un'operazione puramente meccanica che può essere riprodotta all'infinito e non comporta profili di irripetibilità[23];
2. la copia di un file da un computer sotto sequestro non implica alcuna forma di valutazione tecnico scientifica[24].

Da quanto precede si può pertanto concludere che la difesa non avrebbe necessariamente diritto di essere informata in anticipo circa l'esecuzione di queste tipologie di operazioni. A conclusioni diametralmente opposte è giunta al contrario la dottrina, secondo cui ogni interazione uomo macchina dovrebbe assumere la veste dell'accertamento irripetibile e i documenti digitali dovrebbero sottostare ai principi generali di cautela applicabili a tutti i reperti, in base ai quali se le operazioni tecniche possono causare alterazioni è necessaria che esse vengano compiute in regime di contraddittorio delle parti[25]. La natura ripetibile o irripetibile degli atti di acquisizione della DE è sostanzialmente legata alla possibilità che questa subisca modificazioni. Pertanto, per i media digitali intrinsecamente protetti dalla scrittura (come alcune tipologie di CD o DVD) o nel caso di operazioni *post mortem*[26] condotte da soggetti esperti con le tutele tecniche assicurate dalle *best practices* della DF, le acquisizioni di memorie di massa potrebbero essere considerate come ripetibili. Tali garanzie consistono nell'utilizzare:

- dispositivi fisici o logici di protezione dalla scrittura (*write blockers*);
- gruppi di continuità per minimizzare il rischio che interruzioni nell'erogazione della corrente elettrica possano provocare effetti non prevedibili sui reperti;

[22] E. Aprile, *Le indagini tecnico-scientifiche: problematiche giuridiche sulla formazione della prova penale*, in "Cassazione penale", 2003, n. 12, pp. 4034-4042.

[23] Cass., sez. I, 30 aprile 2009, Corvino in CED Cass. n. 244454.

[24] Cass., sez. I, 5 marzo 2009, Aversano Stabile n. 14511.

[25] P. Tonini, *Il documento informatico: problematiche civilistiche e penalistiche a confronto*, in "Corriere giuridico", 2012, n. 3, pp. 432-439.

[26] Per analisi *post mortem* si intende l'analisi forense condotta dopo lo spegnimento del dispositivo, per analisi *live* si intende l'analisi della memoria con il dispositivo acceso, che comporta l'acquisizione e l'analisi del contenuto della RAM, dei processi, delle connessioni di rete, degli ultimi file aperti, dei registri di sistema, ecc.

ISSN 0390-0975 ISBN 978-88-495-3285-2 © Edizioni Scientifiche Italiane

– codici di *hash*, firme digitali e marche temporali fornite da autorità di certificazione a garanzia di integrità, non ripudio e collocazione temporale certa degli atti.

Ciononostante, per scongiurare il rischio di invalidazione di un reperto sia perché esiste comunque un rischio residuo di danneggiamento sia perché l'indagato potrebbe poi disconoscerne i contenuti, il PM può comunque considerare l'accertamento come irripetibile anche in vista del giorno in cui i reperti dovessero essere restituiti all'avente diritto.

In relazione alla discussione che seguirà, è necessario chiedersi, nel caso in cui un accertamento non sia completamente ripetibile, se tutte le alterazioni sopravvenute siano rilevanti per stabilirne *tout court* la irripetibilità[27]. Per esempio, la perdita di informazione inerente la data di modifica di un file dovuta al ripristino di un precedente *backup* potrebbe non risolversi in dibattimento nel rifiuto del reperto quando dovesse rilevare il contenuto del file stesso e non i suoi metadati. A tal riguardo un accertamento potrebbe essere giudicato *sostanzialmente ripetibile* quando le intervenute modificazioni potrebbero essere non rilevanti nel contesto processuale poiché non impattano su principi di rilevanza, affidabilità e sufficienza della DE accettati in tutti i tribunali del mondo.

5. Ripetibilità delle acquisizioni di personal object store

Anche se i principali standard internazionali relativi al trattamento della DE[28] non affrontano il repertamento di dati da piattaforme di *cloud storage*, si è comunque portati ad assumere che esse possano essere considerate come sistemi non interrompibili, raggiungibili solo mediante una rete e per i quali non è possibile parlare di acquisizione *in toto* a causa delle enormi dimensioni. Questo potrebbe indurre a ritenere di trovarsi al cospetto di sistemi soggetti a continue modificazioni, per i quali ha senso considerare esclusivamente un'acquisizione parziale nelle forme dell'accertamento irripetibile[29], ma ciò non è detto che si debba aprioristicamente applicare in tutti i casi. In realtà non consideriamo le attività a basso livello come la cattura della memo-

[27] S. Fasolin, *op. cit.*

[28] ISO/IEC 27037:2012, *Guidelines for identification, collection, acquisition, and preservation of digital evidence*, ISO/IEC 27042:2015, *Guidelines for the analysis and interpretation of digital evidence* e ISO/IEC 27043:2015, *Incident investigation principles and processes*.

[29] S. Aterno, M. Mattiucci, *Cloud Forensics e nuove frontiere delle indagini informatiche nel processo penale*, in "Archivio Penale", 2013, n. 3, pp. 863-876.

© Edizioni Scientifiche Italiane ISSN 0390-0975 ISBN 978-88-495-3285-2

ria volatile di un elaboratore o la copia *bit a bit* poiché non ancora consentite dalle interfacce di programmazione pubblicate dai CSP. Siamo piuttosto interessati agli oggetti e alle cartelle come strutture logiche in grado di rivelarsi, a certe condizioni, inalterabili nel tempo. In ogni caso, la prima considerazione in termini di ripetibilità consiste in un'accurata analisi del rischio e nella conseguente rimozione di tutti i possibili veicoli di modificazione degli oggetti presenti in un'area di *personal storage*. Ciò potrebbe realizzarsi per esempio quando il CSP fosse in grado di garantire:

- l'utilizzo di un account appositamente creato con accesso esclusivo;
- nel caso di utilizzo delle credenziali dell'avente diritto, l'accesso esclusivo della postazione di lavoro dell'analista forense;
- in alternativa, la rimozione dei diritti di scrittura.

Dovrebbe inoltre essere assicurata la sospensione della politica di eliminazione definitiva degli oggetti cancellati e delle revisioni storiche, come normalmente potrebbe avvenire nel caso di servizi non a pagamento[30]. Quando non sia pertanto possibile escludere l'accesso incontrollato di terze parti o l'eliminazione automatica dei contenuti pregressi, l'acquisizione remota dovrebbe essere considerata come attività irripetibile e come tale andrebbe compiuta in regime di contraddittorio tra le parti. Per contro, in presenza delle richiamate garanzie, considerando le intrinseche qualità di resilienza delle piattaforme di *cloud*, l'acquisizione compiuta con strumenti come *Cloud Data Imager* potrebbe essere considerata ripetibile quando l'*inner hash* non cambia ad ogni esperimento di creazione dell'immagine di una cartella o, per contro, sostanzialmente ripetibile se la variazione dell'*inner hash* fosse legata al variare di metadati ritenuti irrilevanti. Il punto di partenza dell'indagine consiste nell'organizzare un semplice test nel quale un gruppo di file di varia natura viene caricato in una cartella nella disponibilità di account appositamente creati per tutti i quattro CSP supportati da *Cloud Data Imager*, ma non prima che per ognuno di essi sia stata calcolata l'impronta di *hash* MD5 e SHA1[31]. In seguito, viene creata l'immagine della cartella e per ogni CSP è

[30] Per gli account gratuiti (es. Dropbox) tipicamente gli oggetti cancellati e le revisioni storiche vengono mantenuti per 30 giorni prima di essere definitivamente eliminati dal provider.

[31] MD5 e SHA1 sono funzioni crittografiche di *hash* che implementano un algoritmo matematico in grado di trasformare dei dati di lunghezza arbitraria in una stringa binaria di dimensione fissa chiamata valore di *hash*. Gli algoritmi usati sono unidirezionali, quindi difficili da invertire, e ciò permette alle funzioni crittografiche di *hash* di trovare ampio utilizzo negli ambiti di sicurezza informatica e di *digital forensics* come nelle firme digitali, nell'autenticazione delle copie forensi oppure come per la crittografia delle credenziali private degli utenti nelle applicazioni web.

ISSN 0390-0975 ISBN 978-88-495-3285-2 © Edizioni Scientifiche Italiane

semplice verificare che i valori contenuti nelle *hash list* prodotte da CDI coincidono invariabilmente, anche a distanza di tempo, con quelli calcolati per ogni file prima dell'inserimento nelle piattaforme. In altri termini, ripetute chiamate alla funzione della libreria *getFileContent* produce sempre lo stesso risultato in termini di contenuto informativo. Poiché quindi i contenuti sono invariabili ciò che potrebbe cambiare in successivi esperimenti potrebbero essere i metadati associati ad ogni oggetto o cartella. È quindi necessario verificare se più chiamate alle funzioni *listFolder* e *listFileRevisions*, ripetute nel tempo, producano lo stesso risultato o se si verifichino alterazioni tra i metadati acquisiti. In quest'ultimo caso è facile comprendere come i valori di *inner hash* cambierebbero ad ogni esperimento di creazione dell'immagine di una cartella. È stato perciò predisposto un ulteriore semplice esperimento nel quale, utilizzando la relativa applicazione desktop ufficiale, è stato creato all'interno di una cartella del profilo utente denominata *360* un solo file *versione_orginale.txt* con alcuni caratteri di testo. Il contenuto è stato successivamente modificato al fine di generare una revisione. Nel seguito verranno esposti i risultati per i provider Dropbox e Microsoft ottenuti catturando le risposte del server all'invocazione delle funzioni di libreria *listFolder* (all'atto dell'ingresso nella cartella *360*) e *listFileRevisions* (mediante pressione del tasto destro del mouse sul file *versione_orginale.txt* per visualizzare le eventuali revisioni).

5.1. Dropbox

Il meccanismo autorizzativo di Dropbox prevede l'accesso in lettura e scrittura ad una specifica cartella o, in alternativa, all'intera area dati. Ciò si traduce in un sistema di protezione grossolano poiché, per esempio, non è possibile impostare programmaticamente un accesso in sola lettura. CDI comunque garantisce l'immodificabilità dei dati dal momento che la libreria utilizza esclusivamente comandi di lettura HTTP GET[32]. Per gli account gratuiti la possibilità di recuperare gli oggetti cancellati o le versioni storiche è consentita entro trenta giorni dalla data di eliminazione o modifica, mentre per le sottoscrizioni a pagamento non esistono limiti temporali. L'invocazione di *listFolder* viene tradotta nel comando GET del protocollo HTTP contenente nell'indirizzo la parola chiave */metadata* ed il percorso della cartella

[32] Il comando GET dell'HTTP (*Hyper Text Transfer Protocol*) viene utilizzato per leggere il contenuto di una risorsa collocata ad uno specifico indirizzo URL (*Uniform Resource Locator*).

© Edizioni Scientifiche Italiane ISSN 0390-0975 ISBN 978-88-495-3285-2

da sfogliare. In base alla documentazione pubblicata dal provider la risposta inviata dal server in formato JSON[33] contiene i parametri mostrati in tab. 2.

Parametro	*Descrizione*
size	Descrizione della dimensione dell'oggetto in formato testo
bytes	La dimensione dell'oggetto in bytes
path	Il percorso in forma canonica dell'oggetto o della cartella
is_dir	Indica se l'elemento è una cartella o un oggetto
is_deleted	Indica se l'elemento è stato cancellato
rev	Identificatore univoco della revisione storica di un oggetto
hash	Impronta di *hash* calcolata sul contenuto della cartella
thumb_exists	Indica se l'oggetto è un'immagine convertibile in una miniatura
icon	Il nome dell'icona usata per mostrare il tipo di oggetto
read_only	Indica se l'utente ha accesso in sola lettura ad una cartella condivisa
modified	La data/ora dell'ultimo caricamento dell'oggetto su Dropbox
client_mtime	Data/ora di modifica del file (come da sistema operativo *client*)
root	Cartella radice (può essere l'intero dropbox oppure una sottocartella

Tab. 2 – Elenco dei metadati inviati da Dropbox per il comando listFolder

La fig. 3 mostra la risposta del server di Dropbox al comando GET, estrapolata dal file di log registrato da CDI. Si evince che *360* è una sottocartella della radice modificata nella piattaforma il 31 gennaio 2016 alle 06.52.43 UTC. *360* contiene un solo file di 41 byte, avente percorso */360/versione_orginale.txt*, id di revisione pari a "a614336529" ed è stato aggiunto a Dropbox il 31 gennaio 2016 alle 08.00.17 UTC. Il booleano *is_deleted* non compare perché viene inviato dalla piattaforma solo per gli oggetti cancellati.

L'invocazione di *listFileRevisions* viene tradotta nel comando GET del protocollo HTTP contenente nell'indirizzo la parola chiave */revisions* ed il percorso del file di cui si cercano eventuali revisioni. La risposta del server di Dropbox contiene i metadati indicati come in tab. 2 relativi a due oggetti. Come visualizzato in fig. 4 l'elemento più in alto è la versione attuale del file *versione_orginale.txt* (ossia la revisione più recente già descritta in fig. 3), mentre il secondo è la versione precedente di 30 bytes aggiunta a Dropbox il 31 gennaio 2016 alle 06.53.40 UTC ed avente identificatore di revisione inferiore ed uguale a "a514336529" . Si noti che il percorso è per entrambe */360/versione_orginale.txt*. Dall'osservazione delle figg. 3 e 4 si evince come non sia ragionevole attendersi che qualche valore corrispondente ai parametri indicati nella tab. 2 possa cambiare nel tempo ed infatti chiamate ripetute

[33] Il *Java Script Object Notation* è un formato testuale per lo scambio di dati tra *client* e server.

ISSN 0390-0975 ISBN 978-88-495-3285-2 © Edizioni Scientifiche Italiane

```
{
  "read_only":false,
  "hash":"67fa708dac16f50d0fb1a747c40562db",
  "revision":164,
  "bytes":0,
  "thumb_exists":false,
  "rev":"a414336529",
  "modified":"Sun, 31 Jan 2016 06:52:43 +0000",
  "size":"0 bytes",
  "path":"/360",
  "is_dir":true,
  "modifier":null,
  "root":"dropbox",
  "contents":[
    {
      "rev":"a614336529",
      "thumb_exists":false,
      "path":"/360/versione_originale.txt",
      "is_dir":false,
      "client_mtime":"Sun, 31 Jan 2016 08:00:17
      +0000",
      "icon":"page_white_text",
      "read_only":false,
      "modifier":null,
      "bytes":41,
      "modified":"Sun, 31 Jan 2016 08:00:17 +0000",
      "size":"41 bytes",
      "root":"dropbox",
      "mime_type":"text/plain",
      "revision":166
    }
```


Fig. 3 – Risposta di Dropbox alla funzione listFolder

alle funzioni *listFolder* per la cartella *360* e *listFileRevisions* per il file *versione_orginale.txt* ottengono invariabilmente gli stessi dati in risposta. Ciò è anche confermato dalla ripetizione dell'esperimento di creazione dell'immagine della cartella *360* anche a distanza di giorni o settimane con il risultato di ottenere sempre lo stesso valore di *inner hash*. Possiamo pertanto concludere che, a condizione di escludere modifiche dei dati compiute da terze parti o da politiche del CSP, le acquisizioni remote dal provider Dropbox sono ripetibili.

© Edizioni Scientifiche Italiane ISSN 0390-0975 ISBN 978-88-495-3285-2

```
{
  "rev":"a614336529",
  "thumb_exists":false,
  "path":"/360/versione_originale.txt",
  "is_dir":false,
  "client_mtime":"Sun, 31 Jan 2016 08:00:17
  +0000",
  "icon":"page_white_text",
  "read_only":false,
  "modifier":null,
  "bytes":41,
  "modified":"Sun, 31 Jan 2016 08:00:17 +0000",
  "size":"41 bytes",
  "root":"dropbox",
  "mime_type":"text/plain",
  "revision":166
},
{
  "rev":"a514336529",
  "thumb_exists":false,
  "path":"/360/versione_originale.txt",
  "is_dir":false,
  "client_mtime":"Sun, 31 Jan 2016 06:53:40
  +0000",
  "icon":"page_white_text",
  "read_only":false,
  "modifier":null,
  "bytes":30,
  "modified":"Sun, 31 Jan 2016 06:53:40 +0000",
  "size":"30 bytes",
  "root":"dropbox",
  "mime_type":"text/plain",
  "revision":165
}
```


Fig. 4 – Risposta di Dropbox alla funzione listFileRevisions

5.2. Microsoft OneDrive

OneDrive implementa uno schema di protezione più sofisticato rispetto a quello di Dropbox ed è quindi possibile impostare da programma un accesso in sola lettura a oggetti e cartelle specificando all'inizio della comunicazione con il server il parametro *wl_contacts_skydrive*. In ogni caso, l'utilizzo esclusivo di comandi di sola lettura HTTP GET costituisce un'ulteriore garanzia

ISSN 0390-0975 ISBN 978-88-495-3285-2 © Edizioni Scientifiche Italiane

di immodificabilità dei dati. Anche per il provider Microsoft la politica di ritenzione non supera i 30 giorni dalla data di eliminazione o modifica di un oggetto, con l'ulteriore penalità che le versioni storiche sono mantenute solo per i documenti nel formato Office. Una fondamentale limitazione dell'interfaccia di programmazione di OneDrive sotto il profilo della DF è la mancanza di funzioni documentate che consentano l'accesso al cestino e alle versioni precedenti di un oggetto. Ne segue che per questo CSP il metodo *listFileRevisions* non è implementabile. Al contrario, la chiamata alla funzione *listFolder* viene tradotta nel comando GET del protocollo HTTP contenente nell'indirizzo la parola chiave/files e l'identificativo univoco della cartella da sfogliare. La risposta inviata dal server in formato JSON contiene i parametri mostrati nella tab. 3.

Parametro	*Descrizione*
id	Identificatore univoco della cartella o dell'oggetto
name (from object)	Nome dell'utente che ha creato la cartella o caricato l'oggetto
id (from object)	Identificatore univoco dell'utente di cui sopra
name	Nome della cartella o dell'oggetto
description	Eventuale descrizione
count	Numero totale di oggetti nella cartella
parent_id	Identificatore univoco della cartella padre
link	Indirizzo di visualizzazione dell'oggetto
size	Dimensione in bytes dell'oggetto
upload_location	Indirizzo di caricamento dell'oggetto
comments_count	Numero di commenti associati all'oggetto
comments_enabled	Indica se i commenti sono abilitati per l'oggetto
is_embeddable	Indica se l'oggetto o la cartella possono essere *embedded*
source	L'indirizzo a cui è possibile ottenere il contenuto dell'oggetto. Informazione NON persistente
type	Indicatore di tipo: oggetto o cartella
created_time	La data/ora del primo caricamento dell'oggetto su OneDrive
updated_time	La data/ora dell'ultimo caricamento dell'oggetto su OneDrive
client_updated_time	Data/ora di modifica del file (come da sistema operativo *client*)
access	Informazioni su chi può accedere all'oggetto (es. "Just me")

Tab. 3 – Elenco dei metadati inviati da OneDrive per il comando listFolder

Come si vede dalla tab. 3 il parametro *source* non è persistente e si verifica facilmente analizzando il file di log prodotto da CDI che ripetute chiamate alla funzione *listFolder* restituiscono ogni volta un indirizzo diverso a cui scaricare l'oggetto *versione_orginale.txt*. Si è quindi portati a concludere che le acquisizioni remote sono comunque irripetibili poiché il valore dell'*inner hash* cambia ad ogni esperimento. Tuttavia, sempre alle condizioni di assen-

© Edizioni Scientifiche Italiane ISSN 0390-0975 ISBN 978-88-495-3285-2

za di variazioni esterne indicate nel par. 5, potrebbe essere accettabile anche un giudizio di sostanziale ripetibilità poiché ciò che varia ad ogni ripetizione non sono i metadati di base (come le marche temporali), ma quelli legati ad informazioni di servizio del provider che potrebbero essere valutati dal giudice come non rilevanti.

6. Conclusioni

Quando le tracce del reato sono contenute all'interno di un'area di *personal object store* potrebbe essere impossibile applicare le migliori pratiche della *digital forensics* tradizionale che consigliano di compiere una copia *bit a bit* delle memorie di massa. Un approccio alternativo consiste nella possibilità di compiere le acquisizioni di dati mediante accesso remoto all'infrastruttura *cloud* ed è pertanto fondamentale lo studio delle interfacce di programmazione pubblicate dai provider per determinare se esse siano idonee per gli scopi della DF. La maturità di tali supporti è stata dimostrata realizzando un'applicazione denominata *Cloud Data Imager* in grado di compiere accessi in sola lettura, creare un'immagine logica di un albero di directory e produrre sia codici di integrità sia registrazioni delle attività in un apposito file di log. Inoltre, l'analisi dei metadati scambiati con le piattaforme di POS ha consentito di concludere che, alle dovute condizioni, gli accertamenti compiuti su dati conservati nei server di Dropbox sono ripetibili, mentre quelli riguardanti Microsoft OneDrive sono irripetibili in senso stretto, ma sostanzialmente ripetibili in senso lato, dal momento che le alterazioni rilevate durante gli esperimenti riguardano solo alcuni metadati minori che potrebbero pertanto essere giudicati di scarsa rilevanza nel corso del dibattimento.

ISSN 0390-0975 ISBN 978-88-495-3285-2 © Edizioni Scientifiche Italiane