

DIGITAL PROFILING
Un nuovo strumento di indagine informatica

Dott.ssa Clara Maria Colombini

INDICE

1. UN NUOVO STRUMENTO DI INDAGINE	pag. 5
2. L'INTELLIGENCE	pag. 6
3. IL PROFILING	pag. 8
4. L'APPLICAZIONE DELLE TECNICHE DI DIGITAL PROFILING ALL'ANALISI DI UN SISTEMA INFORMATICO	pag. 9
5. IL CASO DI STUDIO	pag. 10
5.1 Il reato	pag. 10
5.2 Il problema	pag. 11
6. LE AREE DI RICERCA IN UN SISTEMA INFORMATICO	pag. 13
6.1 L'analisi degli utenti	pag. 13
6.2 L'analisi dei file di testo	pag. 13
6.3 L'analisi delle cartelle di file personali	pag. 13
6.4 L'analisi dell'organizzazione delle cartelle	pag. 14
6.5 L'analisi dei nickname	pag. 14
6.6 L'analisi dei file di log e della cronologia delle connessioni	pag. 14
6.7 L'analisi delle installazioni hardware	pag. 15
6.8 L'analisi delle installazioni software	pag. 15
6.9 L'analisi dei listati di codice	pag. 15
6.10 L'analisi della timeline	pag. 15
6.11 L'analisi in macchina virtuale	pag. 16
6.12 L'analisi del "modus operandi"	pag. 16
7. I RISULTATI DELL'APPLICAZIONE AL CASO DI STUDIO	pag. 17
8. IL METODO	pag. 19

9. LA DESCRIZIONE DEL MODELLO	pag. 20
10. GLI ELEMENTI	pag. 21
11. I PROFILI	pag. 22
12. LE CARATTERISTICHE E LE FUNZIONI DEGLI ELEMENTI	pag. 23
12.1 D – Dispositivo digitale	pag. 23
12.2 f – Feature	pag. 24
12.3 A – Area di file	pag. 26
12.4 F – Insieme di feature	pag. 35
12.5 m – Feature minima	pag. 37
12.6 M – Insieme delle feature minime	pag. 39
12.7 i – Indicatore	pag. 39
12.8 I – Insieme degli indicatori	pag. 40
12.9 k – File contenente indicatori	pag. 40
12.10 K – Insieme dei file contenenti indicatori	pag. 40
13. LA SEQUENZA DELLE OPERAZIONI PER LA CREAZIONE DEL PROFILO DIGITALE	pag. 41
13.1 Ps – Profilo di sistema	pag. 43
13.2 Pc – Profilo cartella utente	pag. 45
13.3 Pd – Profilo dispositivo	pag. 50
13.4 Pu – Profilo utente	pag. 53
13.5 Puc – Profilo utente campione	pag. 57
14. IL CONFRONTO	pag. 58
15. IL CONFRONTO INCROCIATO	pag. 65
16. I DISPOSITIVI MULTI-UTENTE	pag. 66

17. LA VALUTAZIONE DEL RISULTATO	pag. 68
18. CONCLUSIONI	pag. 70
BIBLIOGRAFIA	pag. 71

1. UN NUOVO STRUMENTO DI INDAGINE

Lo sviluppo della moderna tecnologia ha portato ad una trasformazione nel ruolo dei dispositivi digitali che si vanno trasformando da meri contenitori di dati a veri e propri “diari digitali” di colui che li utilizza.

I software che vengono implementati su un numero sempre crescente e diversificato di dispositivi digitali offrono oggi un alto livello di personalizzazione: le agende di appuntamenti, la connessione a chat, blog, forum, social network, le connessioni WIFI, bluetooth, GPS, solo per fare alcuni esempi, hanno ormai trasformato il telefono cellulare, il lettore MP3, la consolle di gioco, il navigatore satellitare (e a breve il televisore) in veri e propri depositari della abitudini di vita dell’individuo che ne fa uso.

Questa profonda trasformazione ha influenzato l’analisi forense dei dispositivi che deve, di conseguenza, ampliare il proprio raggio d’azione verso l’acquisizione e l’analisi delle informazioni che è possibile reperire nelle memorie digitali.

Ad oggi l’analisi tecnica informatica, svolta sul contenuto di una memoria digitale, si limita alla ricerca di quei file che contengono dati di possibile attinenza ad un determinato reato, dietro indicazioni fornite da uno specifico quesito. Se i file vengono ritrovati, sono estrapolati e presentati al magistrato, ma se i dati non ci sono, o se sono troppo pochi, confusi, per poter divenire un’evidenza, l’analisi tecnica si ferma.

Si pensi per esempio all’applicazione delle tecniche di anti-forensics, ben conosciute ed utilizzate soprattutto dalla criminalità organizzata e dal terrorismo, che le utilizzano proprio allo scopo di occultare, e ove possibile, eliminare ogni traccia del reato compiuto e del suo autore.

Il Digital Profiling offre un ulteriore strumento di indagine: esso applica sulla memoria digitale specifiche tecniche di *intelligence* e di *profiling*, allo scopo di ottenere tutta una serie di informazioni che possono essere di utilità nella risoluzione del problema che si pone: dalla descrizione delle modalità con cui è stato compiuto il reato (modus operandi) all’identificazione dell’autore o degli autori del reato in oggetto.

Il processo si svolge attraverso la ricerca e l’analisi delle informazioni che si possono trarre dalle “tracce digitali” lasciate su di esso: il computer è una macchina, ma il suo utilizzatore è un essere umano, e come tale ha una caratteristica unica: tende inevitabilmente a personalizzare l’ambiente con cui interagisce, sia esso reale o virtuale, lasciando, anche inconsciamente, ovunque si muova, delle tracce digitali che possono essere rilevate, confrontate e riconosciute.

2. L'INTELLIGENCE

“Procedimento che attraverso la raccolta, la valutazione e l'analisi delle informazioni, consente di dare un significato all'insieme delle informazioni esaminate”.

Nasce nel lontano XIV secolo in Gran Bretagna, e si sviluppa nelle Università inglesi, a seguito della necessità di disporre di idonee metodologie che consentano di trarre, da un grande quantità di informazioni, dati concretamente utili al contrasto dell'attività criminale.

Le tecniche di Intelligence vengono applicate sia all'analisi della cosiddetta “criminalità evidente” cioè quell'insieme di attività delittuose palesemente consumate sul territorio, che alla “criminalità reale” che comprende invece quelle attività criminali non ancora individuate perché realizzate mediante tecniche innovative o di originale concezione con effetti ancora occulti.



Il processo di analisi delle informazioni dell'intelligence “applicato” si sviluppa in un ciclo (che può quindi ripetersi) suddiviso in 4 fasi:

1° Fase - Individuazione dell'obiettivo.

2° Fase - Raccolta mirata delle informazioni e loro valutazione.

3° Fase – L'analisi delle informazioni:

- a) Selezione dei dati pertinenti dalla massa dei dati raccolti (*indicatori*);
- b) Confronto delle informazioni per l'individuazione di carenze e/o discordanze;
- c) Integrazione delle informazioni, cioè la raccolta di quelle evidenziate dal confronto;
- d) Interpretazione delle informazioni:

e) sviluppo di inferenze:

- strategiche (valutazioni o previsioni);
- operative (preventive o giudiziarie).

A questo punto, se dall'analisi sorgono nuovi elementi o se giungono nuovi dati, il ciclo può ripetersi tornando alla Fase 1.

4° Fase – Utilizzazione del prodotto finale (utilizzazione concreta delle informazioni ottenute con provvedimenti di natura preventiva e/o giudiziaria).

3. IL PROFILING

Il Profiling è una branca della Criminologia che studia i modelli di comportamento criminale allo scopo di fornire un aiuto nell'identificazione dell'autore di un reato.

La sua nascita avviene ufficialmente nel 1972, quando viene istituita all'accademia di Quantico la "Behavioral Science Unit" (BSU), o unità di scienze comportamentali, reparto speciale dell'FBI dedicato all'implementazione di tecniche di analisi per l'individuazione degli autori di crimini violenti.

In Italia il Ministero dell'Interno ha costituito il SASC (Sistema Informativo per l'Analisi della Scena del Crimine) e il SACV (Sistema per l'Analisi del Crimine Violento), tramite cui è possibile, tra l'altro, effettuare ricerche specialistiche ed analisi delle informazioni direttamente sulla scena di un evento criminale, collegandosi all'Archivio Centrale.

Le tecniche che comprende si basano sull'analisi delle peculiarità del crimine commesso per arrivare all'identificazione delle principali caratteristiche di comportamento e personalità di un individuo, a partire da una prima fase che consiste nell'analisi della scena del crimine:

- Scena del delitto: la zona in cui è stato rinvenuto il cadavere, il luogo e le sue caratteristiche.
- La vittima.
- Le lesioni: caratteristiche e loro localizzazione.
- I mezzi lesivi utilizzati.
- La descrizione dei reperti: prove balistiche, fisiche, chimiche, merceologiche, biologiche, grafologiche ecc.
- Caratteristiche dei veicoli eventualmente coinvolti.
- Caratteristiche dell'aggressore.

La seconda fase di analisi è definita "case linkage" cioè il processo attraverso il quale i dati dalla ottenuti nella fase precedente si confrontano per stabilire possibili connessioni fra le prove raccolte in casi differenti.

Da questa si ottengono i modelli comportamentali che permettono di delineare infine il profilo del possibile autore.

La base delle due analisi, sebbene sviluppata su piani differenti, è la medesima: si raccolgono i dati per gruppi omogenei, se ne estraggono quelli significativi, e si confrontano per ricostruire gli eventi.

4. L'APPLICAZIONE DELLE TECNICHE DI DIGITAL PROFILING ALL'ANALISI DI UN SISTEMA INFORMATICO

Il Digital profiling applica queste tecniche all'analisi della memoria di un personal computer, come di qualsiasi altro dispositivo digitale, allo scopo di ricavarne informazioni utili al problema che si presenta.

Il processo si svolge in un ciclo che comprende 6 fasi:

- 1° Fase - Individuazione dell'obiettivo: cosa ricercare in relazione al tipo di problema in oggetto.
- 2° Fase - Raccolta mirata dei dati che contengono informazioni di possibile utilità e loro valutazione, all'interno di specifiche aree di file all'interno della memoria di massa in analisi.
- 3° Fase – Selezione delle informazioni pertinenti e di quelli caratterizzanti dalla massa dei dati raccolti (estrazione degli *indicatori*), per ogni area analizzata.
- 4° Fase - Confronto delle informazioni tratte dai dati (*indicatori*) per l'individuazione di carenze, discordanze o similitudini.
- 5° Fase - Raccolta delle informazioni evidenziati dal confronto, per costruire, laddove necessario, un "*profilo digitale*".
- 6° Fase - Interpretazione delle informazioni in funzione dell'obiettivo.

Per meglio illustrare l'applicazione delle tecniche sopracitate, si presenta qui un esempio di applicazione pratica delle tecniche di Digital Profiling all'analisi forense degli Hard Disk contenuti in personal computer nell'ambito di una indagine informatica.

Si precisa che il caso è stato opportunamente modificato allo scopo di renderlo del tutto anonimo.

5. IL CASO DI STUDIO

5.1 IL REATO

In seguito ad una denuncia relativa ad uno specifico reato, vengono effettuate delle indagini che portano all'arresto di due persone: "Pippo" e "Pluto".

Contestualmente all'arresto viene effettuata una perquisizione che porta al sequestro di n. 10 personal computer.

Il magistrato incarica il consulente tecnico di effettuare una perizia tecnica informatica sulle 10 macchine allo scopo di ricercare ed estrapolare tutti quei file il cui contenuto possa avere attinenza al reato in oggetto.

Figura 1 - I due indagati

Pippo



Pluto



Figura 2 – I dieci Personal Computer: in rosso sono evidenziati i computer contenenti i dati attinenti il reato.



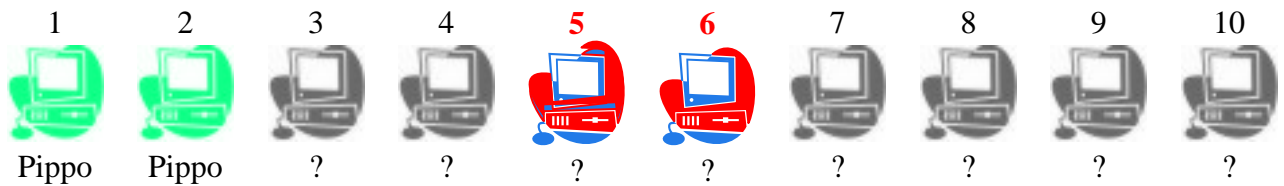
Il consulente tecnico prende in carico le 10 macchine, ed effettua l'analisi forense del contenuto delle memorie di massa in essi contenute.

L'analisi porta a rilevare, su n. 2 macchine, nello specifico la n. 5 e la n. 6, una notevole quantità di file dal contenuto di possibile attinenza al reato in oggetto. Essi vengono quindi estrapolati, descritti in una relazione e presentati al magistrato.

5.2 IL PROBLEMA

Durante la perquisizione non è stato possibile attribuire ad ogni PC il rispettivo utente¹. Solo due PC, il n. 1 ed il n. 2 sono stati riconosciuti come propri da uno dei due indagati: il sig. Pippo, che invece misconosce i due PC contenenti i dati di interesse, dichiarandoli non di sua proprietà. Ciò rende impossibile attribuire ad alcuno la responsabilità del reato.

Figura 3 - Le 10 macchine in analisi ed i rispettivi utenti: le macchine “incriminate” restano anonime.

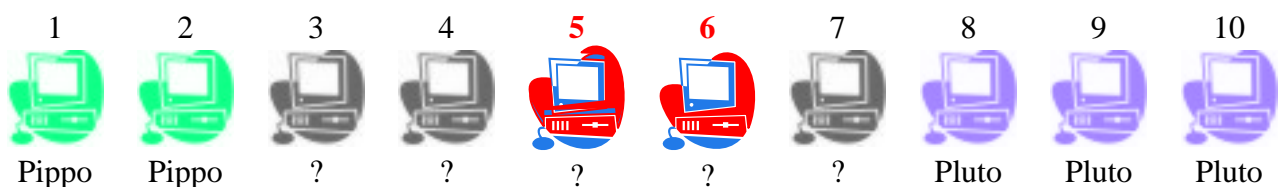


Per risolvere il problema, il magistrato incarica il consulente tecnico di una nuova analisi sulle sole due macchine contenenti i dati inerenti il reato, alla ricerca di email firmate, bollette, contratti, email, o qualsiasi altro documento che possa attribuire con certezza la “paternità” dei due PC incriminati.

L’analisi condotta sul PC n. 5 e sul PC n.6 non porta però alcun risultato: nessun documento, nota, o appunto firmato, nessuna bolletta, contratto, intestati; i messaggi email sono firmati con un nickname, gli indirizzi di posta sono formati da nomi di fantasia e registrati presso un provider estero.

Nel frattempo il secondo indagato, il sig. Pluto, durante un interrogatorio, riconosce come propri altri tre PC, il n. 8, il n. 9 e il n. 10, il che porta alla situazione illustrata nella figura 4.

Figura 3 - Le 10 macchine in analisi ed i rispettivi utenti: le macchine “incriminate” restano ancora anonime.



¹ Si intende qui per “utente” l’effettivo utilizzatore della macchina, che può essere persona diversa dal proprietario.

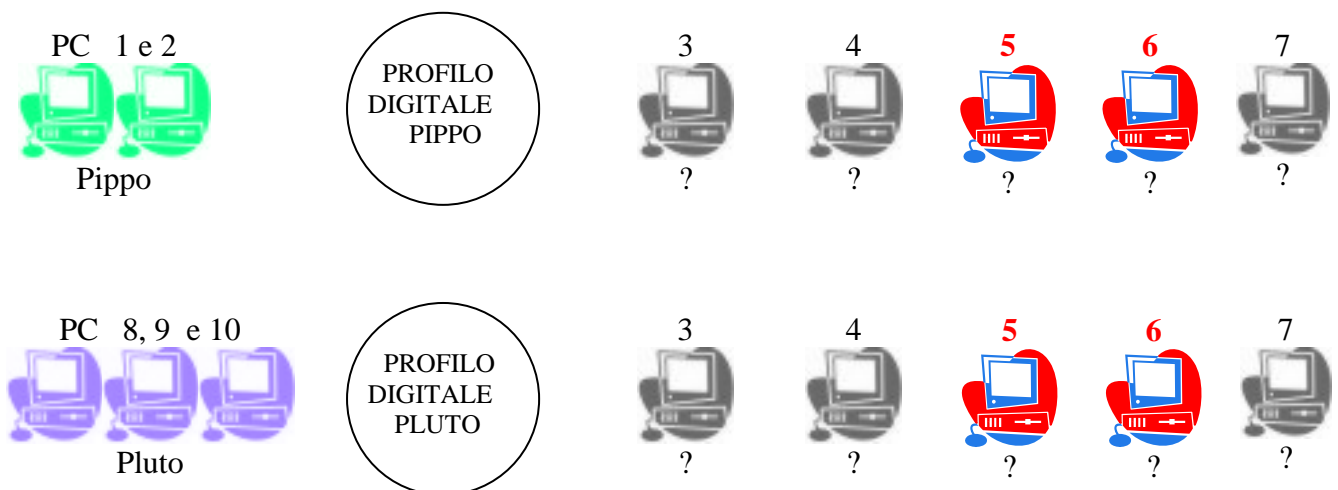
A questo punto l'analisi forense di tipo tecnico-informatico si sarebbe arrestata, ma, poichè il caso lo richiedeva, è stato possibile utilizzare il nuovo strumento di indagine offerto dal Digital Profiling, per effettuare una nuova analisi allo scopo di identificare, se possibile, gli utenti dei due computer incriminati.

Sono stati ripresi in analisi tutti e dieci i PC, andando ad effettuare raccolte mirate di dati per specifiche aree, da cui estrarre informazioni rilevanti (indicatori) che potessero delineare un profilo digitale per ogni utente presente in ognuna delle macchine.

L'operazione è stata effettuata a partire dalle macchine di utente noto: con i dati caratterizzanti estrapolati dai PC 1 e 2, è stato costruito il profilo digitale del sig. Pippo, dai PC 8, 9 e 10 quello del sig. Pluto.

I due profili sono stati infine messi a confronto con quelli anonimi emersi dalle altre macchine, per evidenziare eventuali analogie che potessero farli riconoscere nei profili anonimi delle macchine incriminate per individuarne gli utenti.

Figura 4 - Confronto fra i profili digitali.



6. LE AREE DI RICERCA IN UN SISTEMA INFORMATICO

Si presentano qui di seguito le principali aree di ricerca di dati da cui è possibile trarre gli indicatori utili alla creazione del profilo digitale, all'interno di un sistema informatico.

Nello specifico si è preso in esame il caso di studio, consistente nell'analisi di personal computer con sistema operativo Windows.

Non a caso si è scelto un Personal Computer in quanto è il mezzo informatico che contiene la più grande massa di dati di cui effettuare un'analisi.

Va tenuto comunque ben presente che ogni caso è diverso dall'altro, e che le seguenti tipologie di ricerca devono essere adattate a seconda della situazione, del tipo di dispositivo in analisi, della fattispecie di reato cui ci si trova di fronte, e dell'obiettivo che ci si prefigge di raggiungere.

6.1 L'analisi degli Utenti

L'installazione di un Sistema Operativo viene registrata dai file di registro che ne contengono la data, i nomi utenti/organizzazione e numeri seriali: da essa si traggono informazioni sugli utenti che hanno accesso al PC, come la data di installazione, il numero di accessi, l'ultimo accesso, l'ultimo cambio di password ecc.

Operazione:

Estrapolazione dai file di registro "SAM" e "NTUSER" delle informazioni su quanti e quali utenti hanno accesso alla macchina, siano essi registrati come utenti dal sistema operativo, che come utenti di rete.

6.2 L'analisi dei file di testo

Ogni individuo ha un suo stile di scrittura, utilizza particolari espressioni idiomatiche, compie gli stessi errori di sintassi o di grammatica², di battitura; errori che tende a ripetere in tutto ciò che scrive.

Operazione:

Ricerca ed analisi tutti i file che contengono testi scritti dall'utente: messaggi di posta elettronica, conversazioni Chat, appunti, documenti, ecc. al fine di estrapolarne le cosiddette "firme" caratterizzanti, che possono essere confrontate e riconosciute in altri documenti, non solo digitali.

Applicazione al caso: nei PC 1 e 2 sono state rinvenute conversazioni chat e testi di messaggi email scritti dall'utente Pippo, il quale di origine straniera, utilizzava particolari espressioni idiomatiche oltre a compiere errori di grammatica tipici degli stranieri. Gli stessi sono stati riconosciuti nel PC 5 (anonimo).

6.3 L'analisi delle cartelle di file personali

L'utente che utilizza diversi computer (per esempio un fisso ed un portatile) tende a portarsi con sé le stesse raccolte di musica o di foto, ponendole spesso in una cartella con lo stesso nome, e posizionata nello stesso punto.

² Particolari errori di sintassi e di grammatica permettono di riconoscere l'origine italiana o straniera dell'utente.

Operazione:

Ricerca e analisi di tutte quelle cartelle che contengono: brani musicali, immagini, scatti fotografici, filmati, documenti, pubblicazioni ecc.

Applicazione al caso: nei PC 1 e 2 sono state rinvenute le stesse 3 cartelle di file MP3 di brani musicali del paese di origine dell'utente Pippo. Le stesse sono state ritrovate nel PC 5 (anonimo). Nel PC 6 sono state rinvenute 2 cartelle di foto scattate dall'utente Pluto che sono state ritrovate anche nel PC 6 (anonimo).

6.4 L'analisi dell'organizzazione delle cartelle

L'utente tende a ripetere il medesimo schema organizzativo di file e cartelle allo scopo di ritrovare con più rapidità i file quando si sposta da una macchina all'altra.

Operazione:

Ricerca ed estrapolazione dello schema di organizzazione delle cartelle e dei file personali.

Applicazione al caso: le cartelle di cui al punto 5.3 sono state rinvenute, nei PC anonimi, nella medesima posizione rispetto allo schema di organizzazione dei file, in cui si trovavano sui PC degli utenti Pippo e Pluto.

6.5 L'analisi dei nickname

I nickname vengono utilizzati per l'accesso a chat, blog, forum, social network, oltre a venire utilizzati per gli indirizzi di posta elettronica. Se ripetuti su differenti PC possono essere facilmente riconosciuti.

Operazione:

Ricerca ed estrapolazione degli indirizzi di posta elettronica, i nickname di chat, blog, social network.

Applicazione al caso: l'utente Pluto, utilizzava sul PC 10, (conosciuto) un particolare nickname per connettersi a due chat, nickname che è stato riconosciuto sul PC 6 (anonimo).

6.6 L'analisi dei file di log e della cronologia delle connessioni

La navigazione abituale di specifici siti Internet, come l'accesso a forum, conti online, web mail, connessioni ftp, legati ad un accesso tramite nome utente e password, è rilevabile e riconoscibile se ripetuta su PC differenti.

Operazione:

Estrapolazione dei file che contengono la cronologia dei siti internet visitati abitualmente (URL di accesso a webmail, forum, blog, accessi a conti correnti online, connessioni ftp, ecc.). Lo stesso viene fatto per i "Preferiti" del browser utilizzato per la navigazione in Internet.

Applicazione al caso: sia l'utente Pippo che l'utente Pluto si sono collegati a siti di web mail personali, a pagine web con accesso riservato e ai propri conti online dalle proprie macchine. Gli stessi URL sono stati rinvenuti sui PC 5 e 6 (anonimi).

6.7 L'analisi delle installazioni hardware

L'utente che utilizza più di un PC, vi connette spesso le stesse periferiche: dispositivi di memoria USB, telefoni cellulari, macchine fotografiche, lettori MP3 ecc. Di queste installazioni è possibile ritrovare traccia dei file di registro che riportano nome, marca e numero di serie del dispositivo collegato.

Operazione:

Ricerca ed estrapolazione dai file di registro le informazioni relative a quali periferiche (stampanti, drive esterni, memorie USB, macchine fotografiche, telefoni cellulari ecc.) siano state installate sulla macchina, al fine di rilevarne data di installazione, marca, tipo e, dove fornito, il numero di serie).

Applicazione al caso: l'utente Pippo aveva collegato al PC 2 una penna USB, una macchina fotografica ed un cellulare, i cui numeri seriali sono stati ritrovati anche nei file di installazione nel PC 5 (anonimo). L'utente Pluto aveva invece connesso al PC 10 una Penna USB, un telefono cellulare ed un drive esterno, i cui numeri seriali sono stati ritrovati anche nei file di installazione nel PC 6 (anonimo).

6.8 L'analisi delle installazioni software

Come per l'hardware, l'utente che utilizza più di un PC, vi installa spesso il medesimo software, di cui rimane traccia (nome, versione, numero di serie) nei file di registro.

Operazione:

Ricerca ed estrapolazione dai file di registro di quali applicazioni software siano state installate sulla macchina, al fine di rilevarne data di installazione, versione e numero di serie (o il medesimo crack o keygen), con particolare riguardo ai tool di formattazione, wiping, criptatura, steganografia, macchine virtuali.

Applicazione al caso: l'utente Pippo aveva scaricato da Internet attraverso il software Emule n. 5 programmi "pirata" forniti di un numero di serie che è stato ritrovato nell'installazione degli stessi programmi sia sul PC 2 che sul PC 5 (anonimo). Gli stessi software, con il medesimo numero seriale, sono stati ritrovati sul PC 10 dell'utente Pluto e sul PC6 (anonimo).

6.9 L'analisi dei listati di codice

Ogni programmatore ha un suo personale stile di programmazione: dalla scelta delle funzioni che meglio conosce all'ordine in cui le implementa, senza dimenticare il modo di commentare le righe di codice, unico per ogni programmatore.

Operazione:

Ricerca ed estrapolazione dei file contenenti listati di codice.

Dalla catalogazione di questi due punti si ricava inoltre il profilo delle conoscenze e capacità informatiche dell'utente della macchina (es: la presenza di tool di programmazione, di più sistemi operativi, di macchine virtuali ecc. denota conoscenze informatiche elevate).

6.10 L'analisi della timeline

I file di log presenti su di un computer forniscono informazioni su data e ora (e quindi degli intervalli di tempo) dell'utilizzo della macchina:

- accensione e spegnimento della macchina;

- operazioni sui file/cartelle;
- accesso ad Internet;
- accesso alla posta elettronica;
- ecc.

Operazione:

Analisi della timeline delle operazioni sui file contenuti nella memoria, della navigazione internet, dell'accesso alla posta elettronica, a chat, forum, ecc. in relazione al lasso di tempo in cui il reato in oggetto è stato compiuto (verifica di eventuali alibi).

6.11 L'analisi in macchina virtuale

“Non giudicare una persona se non hai prima camminato per 5 lune nei suoi mocassini” (detto Apache)

Operazione:

Caricamento dell'immagine della memoria in una macchina virtuale allo scopo di navigare il contenuto del computer così come vede l'utente: questa operazione permette di verificare, fra l'altro, l'esecuzione automatica di applicazioni all'avvio, l'automazione degli aggiornamenti, la disposizione del desktop, ecc.

6.12 L'analisi del “modus operandi”

Per risalire all'identità degli autori di un reato a mezzo computer (es. phishing – attacchi a server – ecc), è utile la ricostruzione e la successiva analisi del modus operandi. Allo scopo, dall'analisi dei dati possono essere ottenute le seguenti informazioni:

- obiettivo dell'attacco (frode, interruzione di servizio, attacco politico ecc.);
- tool e tecniche utilizzate per le intrusioni (rootkit, shell, worm, social engineering, ecc.);
- lasso di tempo scelto per l'attacco (giorno/notte, intra/fine settimanale ecc.);
- durata dell'attacco ed eventuale periodicità (unico o frammentato in prestabiliti intervalli temporali ecc.);
- correlazione del momento prescelto per l'attacco con eventi esterni;
- tipologia delle vittime scelte (istituzione governativa italiana o straniera, banca, organizzazione commerciale ecc.);
- tipologia delle tecniche di anti-forensic utilizzate;
- raggiungimento dell'obiettivo.

7. I RISULTATI DELL'APPLICAZIONE AL CASO DI STUDIO

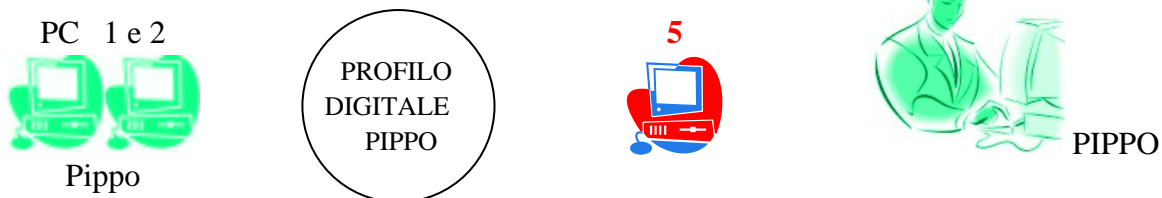
Gli indicatori estratti dalle macchine riconosciute dagli utenti e che ne hanno delineato il profilo digitale, sono stati messi a confronto con i profili digitali “anonimi” scaturiti dall'analisi dei restanti personal computer, compresi quelli non “incriminati”.

Questo ha fatto emergere i numerosi punti in comune tra i profili, che hanno contribuito all'attribuzione del reato ai due indagati:

1) il profilo digitale dell'utente Pippo (PC 1 e PC 2) ha in comune con quello del PC 5 i seguenti indicatori:

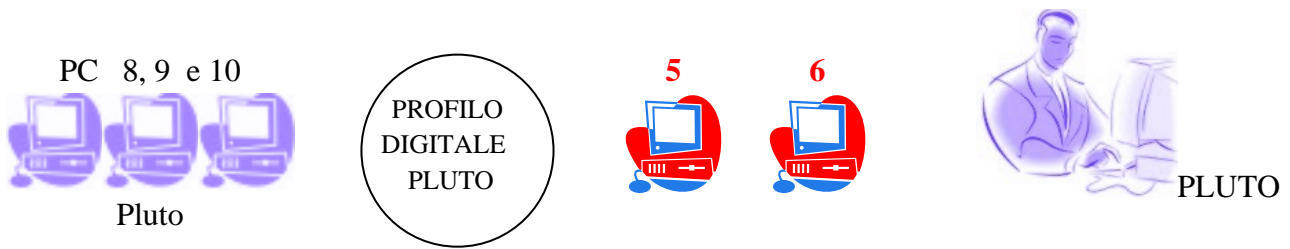
- stile conversazioni chat;
- stile messaggi email;
- cartelle di file personali;
- organizzazione delle cartelle di file personali;
- accesso a specifiche pagine web con accesso tramite identificazione;
- installazione di specifici hardware;
- installazione di specifici software.

Figura 5 - Identificazione dell'utente Pippo.



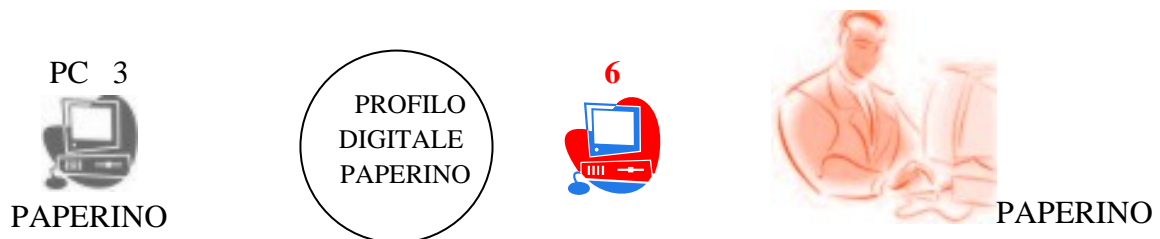
2) il profilo digitale dell'utente Pluto (PC 8 – 9 - 10) ha in comune con quello del PC 6 i seguenti indicatori:

- cartelle di file personali;
- organizzazione delle cartelle di file personali;
- nickname utilizzato;
- accesso a specifiche pagine web con accesso tramite identificazione;
- installazione di specifici hardware;
- installazione di specifici software.

Figura 6 - Identificazione dell'utente Pluto.

Infine, l'applicazione delle tecniche di Digital Profiling su tutte e 10 le macchine ha permesso di rivelare il coinvolgimento nel compimento del reato di una terza persona: dal PC 6, (una delle due macchine incriminate) è emerso un ulteriore profilo digitale, scaturito dalla presenza di un altro utente della macchina coinvolto nel reato, ma anonimo.

Esso non coincideva con i profili digitali dei due indagati, ma bensì con il profilo dell'utilizzatore, conosciuto, del PC 3, inizialmente scartato dalle indagini, in quanto la macchina non conteneva alcun dato di attinenza al quesito.

Figura 7 - Identificazione dell'utente Paperino.

8. IL METODO

I capitoli seguenti descrivono il metodo di applicazione delle tecniche di analisi del Digital Profiling fin qui presentato, ovvero l'estrapolazione, il confronto ed il riconoscimento del Profilo Digitale dell'utilizzatore di un dispositivo digitale.

Per meglio descrivere l'applicazione del metodo proposto, si riporta l'esempio di un caso il cui obiettivo è l'identificazione dell'autore di un determinato reato attraverso l'analisi dei dispositivi digitali che si presumono in uso al soggetto.

L'identificazione viene effettuata attraverso il confronto tra un primo profilo digitale estrapolato da un PC attribuito con certezza al soggetto e i profili estrapolati da altri dispositivi digitali con i quali è stato compiuto il reato in oggetto, non attribuibili però con certezza al soggetto stesso.

Va rilevato che il principio su cui si basa il metodo è biunivoco; è possibile cioè partire anche dal profilo digitale dell'utente "anonimo" del dispositivo, per confrontarlo con i profili di altri dispositivi (anche non coinvolti nel reato) attribuiti con certezza a determinati soggetti.

E' possibile inoltre da un profilo digitale estrapolare un "modus operandi" (es. attacco informatico) da confrontare con altri al fine di riconoscerne ed identificarne l'autore.

9. LA DESCRIZIONE DEL MODELLO

Il metodo si compone dei seguenti passi che descrivono un ciclo che si può ripetere ogniqualvolta emergono nuove informazioni:

1. costruzione del profilo digitale dell'utente (o degli utenti) di un dispositivo scelto come "profilo campione";
2. estrapolazione dei profili digitali degli utenti da eventuali altri dispositivi in analisi;
3. confronto tra i profili ottenuti ai fini di evidenziare convergenze-divergenze;
4. analisi quantitativo-qualitativa delle convergenze-divergenze tra i profili ai fini dell'identificazione del soggetto.

La creazione del seguente modello parte dallo studio delle informazioni caratterizzanti rilevabili dai file contenuti in un Personal Computer, dispositivo scelto sia per la capacità di memoria, sia per la elevata personalizzazione permessa dalle molteplici applicazioni disponibili.

Va precisato in proposito che va costruito un profilo per ognuno degli utenti rilevati per ogni sistema operativo installato sulla macchina, ivi comprese le macchine virtuali.

Il modello descrive:

1. gli elementi;
2. i profili;
3. le caratteristiche e le funzioni degli elementi;
4. la sequenza delle operazioni per la creazione del profilo digitale;
5. il confronto;
6. la valutazione del risultato ottenuto.

10. GLI ELEMENTI

Vengono qui di seguito elencati gli elementi del modello.

D	Dispositivo digitale. Di seguito gli elementi che si riferiscono ad un determinato dispositivo (indicato con D):
f(D)	Feature
A(D)	Area di file
F(D)	Insieme di feature
m(D)	Feature minima
M(D)	Insieme di feature minime
i(D)	Indicatore
I(D)	Insieme di indicatori
k	File contenente indicatori
K(D)	Insieme di file contenenti gli Indicatori

11. I PROFILI

Vengono qui di seguito elencati i profili creati dal modello.

Profilo di sistema relativo al dispositivo D

Ps(D)

Composto da:

1. **I(Ps)(D)** – insieme indicatori file di registro;
2. **K(Ps)(D)** - insieme file di registro contenenti gli indicatori.

Profilo cartella utente relativo al dispositivo D

Pc(D)

Composto da:

1. **I(Pc)(D)** – insieme indicatori cartella utente;
2. **K(Pc)(D)** – insieme file cartella utente contenenti gli indicatori.

Profilo dispositivo relativo al dispositivo D

Pd(D)

Composto da:

- 1) **I(Pd)(D)** – insieme indicatori dispositivo;
- 2) **K(Pd)(D)** - insieme file dispositivo cartella utente contenenti gli indicatori.

Profilo utente relativo al dispositivo D

Composto da:

Pu(D)

1. **I(Pu)(D)** – unione dei tre insiemi di indicatori rilevati:

$$I(Ps)(D) \cup I(Pc)(D) \cup I(Pd)(D)$$
2. **K(Pu)(D)** – unione dei tre insiemi di file:

$$K(Ps)(D) \cup K(Pc)(D) \cup K(Pd)(D)$$

Puc(D)

Profilo utente campione

Profilo di riferimento per il confronto. Coincide con il Profilo utente.

12. LE CARATTERISTICHE E LE FUNZIONI DEGLI ELEMENTI

12.1 D – dispositivo digitale

Per dispositivo digitale “D_i” si intende:

- **qualsivoglia dispositivo digitale** provvisto di memoria interna permanente in grado di memorizzare dei file.
Esempio: PC, telefono cellulare, navigatore satellitare, centralina telefonica, ecc.;
- **supporto di memoria** in grado di memorizzare dei dati.
Esempio: Hard disk, flash card, memory card, smart card, USBpen, CD, DVD, DAT, ecc.);
- **area di memoria** disponibile in remoto ove siano memorizzati file di dati creati da utenti
Esempio: Archivi in remoto;
- **macchina virtuale** contenente un sistema operativo;
- **insieme di file** di dati relativi ad accessi
Esempio: file di log.

12.2 f - feature

Per feature “ f_i ” si intende la singola caratteristica di base, hardware o software, non ulteriormente scomponibile in altre feature più elementari, analizzabile nel contesto di studio, in quanto essa contiene l’informazione che descrive il “comportamento digitale” dell’utilizzatore del dispositivo.

La feature viene ricavata dai file memorizzati all’interno del dispositivo e selezionati in base all’obiettivo dell’indagine.

Essa può essere costituita da:

- proprietà del file;
- contenuto del file.

Un file può contenere una o più feature: sono infatti considerate caratteristiche di base, secondo lo scopo dell’indagine:

- nome file: dal ritrovamento dei medesimi file di carattere personale (testi scritti, scatti fotografici, brani musicali, filmati video ecc.) in diversi dispositivi si può dedurre che vi siano stati memorizzati dallo stesso utente.
- Path: indica se il file ritrovato identico, sia stato memorizzato nella stessa posizione nell’albero delle cartelle (medesimo nome cartella o serie di cartelle).

Esempio:

PC01:

c:\Documents and Settings\Pippo\Desktop\Documenti\XXX\Branipippo\IMAG0122.jpg
 c:\Documents and Settings\Pippo\Desktop\Documenti\XXX\Branipippo\IMAG0123.jpg
 c:\Documents and Settings\Pippo\Desktop\Documenti\XXX\Branipippo\IMAG0124.jpg

PC02:

c:\Utenti\Pippo\Desktop\Documenti\XXX\Branipippo\IMAG0122.jpg
 c:\Utenti\Pippo\Desktop\Documenti\XXX\Branipippo\IMAG0123.jpg
 c:\Utenti\Pippo\Desktop\Documenti\XXX\Branipippo\IMAG0124.jpg

- MD5 (o altro algoritmo di hashing): la feature fornisce la certezza matematica della coincidenza del contenuto del file ritrovato su diversi dispositivi;

- data di creazione, di modifica, di cancellazione: queste tre feature forniscono la cronologia di:
 - memorizzazione;
 - modifica;
 - cancellazione;degli stessi file ritrovati sui diversi dispositivi;

- qualsiasi informazione ricavabile dal suo contenuto.

Esempio:

- da un file di testo scritto di pugno dall'utente (ivi compresi i messaggi email, chat, sms, ecc.) possono essere rilevate diverse feature:
 - o nickname utilizzato come firma o come appellativo;
 - o espressione idiomatica ricorrente (italiana o straniera);
 - o errore di scrittura ricorrente (di battitura, di ortografia, di coniugazione ecc.).

- Nel file di uno scatto fotografico, può essere ritenuto informazione caratterizzante, e quindi rilevato come feature il soggetto (o parte di esso) dello scatto, :
 - o una determinata persona;
 - o un luogo;
 - o un particolare (es. la targa di un'autovettura, un'insegna, un orologio, un quotidiano, ecc.).

12.3 A – Area di file

Allo scopo di individuare con maggior precisione, all'interno della memoria del dispositivo D , i file che possono contenere feature, essi sono stati suddivisi per tipologia in specifiche aree, definite come $A_i(D)$.

$$\bigcup_i A_j(D) \subset D$$

Si definisce quindi $A_i(D)$ come il sottoinsieme omogeneo di D in cui sono contenuti, suddivisi, per tipologia, tutti i file che possono contenere feature, relativamente al dispositivo D .

12.3.1 Classificazione delle aree di file A

Ogni dispositivo ha una sua specifica classificazione delle aree di file contenenti feature, secondo le sue specifiche funzionalità e le applicazioni di cui dispone.

Si presenta qui una classificazione generica delle aree di base riferita al Personal Computer.

a) A_1 – File di registro: utenti di sistema

Specifici file di registro contengono le informazioni relative agli utenti registrati dal sistema operativo come utenti di sistema o utenti di dominio, i nomi utilizzati per la registrazione, il numero di accessi, la data e l'ora dell'ultimo accesso, ecc.

Feature:

1. nome computer
2. nome utente di sistema
3. nome utente di dominio
4. numero accessi alla macchina
5. data-ora ultimo accesso
6. ultimo cambio password.

b) A_2 – File di registro: installazioni hardware

Specifici file di registro contengono le informazioni relative ai dati concernenti i dispositivi hardware collegati al PC.

Feature:

1. tipologia del dispositivo
2. nome –marca-modello del dispositivo
3. numero di serie del dispositivo
4. data ultimo accesso.

c) A₃ – File di registro: installazioni software

Specifici file di registro contengono le informazioni relative ai dati concernenti i software installati sul PC:

Feature:

1. nome, marca, versione
2. numero di serie
3. data di installazione.

d) Le aree di file personali

Vengono qui considerati “file personali” tutti quei file memorizzati dall’utente sul dispositivo al di fuori dei programmi installati, e che possono contenere informazioni caratterizzanti il “comportamento digitale” dell’utente.

Le feature possono essere ricavate sia dai metadati che descrivono il trattamento subito dai file che dal loro contenuto.

La selezione dei file di interesse in queste aree è particolarmente delicata, proprio per la natura di alcune particolari feature:

- l’operazione di selezione di un file anziché un altro utilizzando come filtro il nome, l’estensione, l’MD5 ecc. è relativamente semplice, in quanto può essere effettuata, su indicazione dell’investigatore, dal consulente tecnico con l’ausilio di tool che ne facilitino la raccolta e la selezione tramite appositi filtri;
- assai più delicata è invece l’individuazione delle feature nei file di testo, di immagini, di video, o audio (intese come registrazioni di conversazioni), poiché essa richiede la presenza, accanto al consulente tecnico che opera per la parte informatica, dell’investigatore che segue le indagini e che può più riconoscere ed indicare al consulente quali sono le feature interessanti da considerare (riferimenti a persone, immagini di soggetti o luoghi conosciuti, frasi in “codice” già rilevate durante le indagini, ecc.).

L'area dei file personali è stata suddivisa per tipologia di file nelle seguenti categorie:

A₄ – File di testi personali

File di testo scritti di pugno dall'utente (appunti, memoriali, lettere personali, ecc.) che ne evidenzino lo stile di scrittura (doc, docx, txt, rtf, odt, pdf, ma anche xls ecc.). La loro analisi può evidenziare numerose feature.

a. Feature dai metadati:

1. Nome file
2. Data creazione file
3. Data modifica file
4. Data accesso file
5. Data cancellazione file
6. Path file
7. MD5-SHA1.

Oltre alle informazioni che possono fornire i metadati, altre feature possono essere rilevate dal contenuto dei file:

b. Feature dal contenuto:

1. Firma
2. Nickname
3. Nome proprio
4. Password di accesso
5. Espressione idiomatica
6. Errore di ortografia
7. Errore di battitura
8. Riferimento ad un determinato evento
9. Riferimento ad una determinata persona
10. Riferimento ad un determinato oggetto
11. Riferimento ad un determinato luogo
12. Frase particolare
13. Indirizzo email
14. Ecc.

e) Af[5] - File personali: messaggi email personali (esclusi newsletter, pubblicità, ecc.)

Sono considerati come i file di testo, e vengono analizzati sia i messaggi inviati che quelli ricevuti.

a. Feature dai metadati:

1. Indirizzi email mittente
2. Indirizzi email destinatari
3. Data invio
4. Data ricezione
5. Data cancellazione

b. Feature dal contenuto:

1. Firma
2. Nickname
3. Nome proprio
4. Password di accesso
5. Espressione idiomatica
6. Errore di ortografia
7. Errore di battitura
8. Riferimento ad un determinato evento
9. Riferimento ad una determinata persona
10. Riferimento ad un determinato oggetto
11. Riferimento ad un determinato luogo
12. Frase particolare
13. Indirizzo email
14. Ecc.

f) A₆ – File personali: conversazioni chat

Sono considerati come i file di testo.

a. Feature dai metadati:

1. Nickname utente
2. Nickname contatto
3. Data chat
4. Data cancellazione

b. Feature dal testo:

1. Firma
2. Nickname
3. Nome proprio
4. Password di accesso
5. Espressione idiomatica
6. Errore di ortografia
7. Errore di battitura
8. Riferimento ad un determinato evento
9. Riferimento ad una determinata persona
10. Riferimento ad un determinato oggetto

11. Riferimento ad un determinato luogo
12. Frase particolare
13. Indirizzo email
14. Ecc.

g) A₇ – File personali: file di immagini fotografiche (bmp, jpg, tif, ecc.)

Si intendono qui i file di scatti fotografici memorizzati dall'utente.

a. Feature dai metadati:

1. Nome file
2. Data creazione file
3. Data modifica file
4. Data accesso file
5. Data cancellazione
6. Path file
7. MD5-SHA1

b. Feature dal contenuto:

1. Immagine fotografica di una determinata persona
2. Immagine fotografica di un determinato luogo
3. Immagine fotografica di un oggetto

h) A₈ – File personali: file di immagini grafiche (jpg, tif, dwg, ecc.)

Si intendono qui le raccolte di immagini grafiche varie, come copertine di DVD, CD, raccolte tematiche di illustrazioni, di arte, di fumetti, ecc.

a. Feature dai metadati:

1. Nome file
2. Data creazione file
3. Data modifica file
4. Data accesso file
5. Data cancellazione
6. Path file
7. MD5-SHA1

i) A₉ - File personali: file di filmati video (avi. Mpg, ecc.)

Si intendono qui le raccolte di filmati video memorizzati dall'utente.

Si pone particolare riguardo ai video realizzati da telefoni cellulari, videocamere personali, webcam, ecc.

a. Feature dai metadati:

1. Nome file
2. Data creazione file
3. Data modifica file
4. Data accesso file
5. Data cancellazione
6. Path file
7. MD5-SHA1

b. Feature dal contenuto:

1. Immagine video di una determinata persona
2. Immagine video di un determinato luogo
3. Immagine video di un determinato oggetto
4. Immagine video di un determinato evento
5. Contenuto della traccia audio: nome proprio
6. Contenuto della traccia audio: nickname
7. Contenuto della traccia audio: accento
8. Contenuto della traccia audio: espressione idiomatica
9. Contenuto della traccia audio: difetto di pronuncia
10. Contenuto della traccia audio: riferimento a una determinata persona
11. Contenuto della traccia audio: riferimento a un determinato evento
12. Contenuto della traccia audio: riferimento a un determinato luogo
13. Contenuto della traccia audio: riferimento a un determinato oggetto
14. Contenuto della traccia audio: frase particolare
15. Ecc.

j) A₁₀ - File personali: file audio (wav, mp3, ecc.)

Si intendono qui le raccolte di file audio memorizzati dall'utente.

Si pone particolare riguardo ai file realizzati da microregistratori, telefoni cellulari, ecc.

a. Feature dai metadati:

1. Nome file
2. Data creazione file
3. Data modifica file
4. Data accesso file
5. Data cancellazione
6. Path file
7. MD5-SHA1

In caso di file audio di conversazioni, vanno aggiunte le feature estraibili dal contenuto:

b. Feature dal contenuto:

- i. Contenuto della traccia audio: nome proprio
- ii. Contenuto della traccia audio: nickname
- iii. Contenuto della traccia audio: accento
- iv. Contenuto della traccia audio: espressione idiomatica
- v. Contenuto della traccia audio: difetto di pronuncia
- vi. Contenuto della traccia audio: riferimento a una determinata persona
- vii. Contenuto della traccia audio: riferimento a un determinato evento
- viii. Contenuto della traccia audio: riferimento a un determinato luogo
- ix. Contenuto della traccia audio: riferimento a un determinato oggetto
- x. Contenuto della traccia audio: frase particolare
- xi. Ecc.

k) A₁₁ – URL

Si analizzano qui i collegamenti a pagine web effettuati dall'utente.

In particolare, vengono presi in considerazione i collegamenti alle pagine di accesso via nomeutente-password a:

- posta elettronica via web;
- collegamenti FTP;
- conti online;
- pagine web con accesso autenticato;
- ecc.

a) File “Cookies”:

1. Nome file (url al sito web cui l'utente si è collegato)
2. Data creazione file
3. Data modifica file
4. Data accesso file
5. Data cancellazione

b) File di cronologia browser internet:

a. Feature dai metadati:

1. Data creazione file
2. Data modifica file
3. Data accesso file
4. Data cancellazione

b. Feature dal contenuto:

1. Nome utente che ha effettuato l'accesso
2. Indirizzo URL della pagina web visitata
3. Path del file sul PC a cui l'utente ha avuto accesso

Il numero delle aree di ricerca delle feature è flessibile: esso dipende dall'obiettivo della ricerca e dal tipo di applicazioni installate sul dispositivo.

Le ultime due aree qui descritte offrono un esempio di personalizzazione dell'analisi nel caso in cui sia ipotizzato un reato informatico (es.: attacco informatico tramite codice malevolo autoprodotta o scaricato dal web).

l) A₁₂ – LISTATI CODICE

Si prendono in considerazione i file di listati di codice di programmazione rinvenuti in cartelle create dall'utente e cioè:

- prodotti dall'utente (es. con applicazioni software presenti sulla macchina),
- scaricati dal web,
- ricevuti via email,
- ecc.

a. Feature dai metadati:

1. Nome file
2. Data creazione file
3. Data modifica file
4. Data accesso file
5. Data cancellazione

b. Feature dal contenuto:

1. commenti al codice
2. Risultato dell'esecuzione del listato (scopo dell'applicazione)

m) A₁₃ FILE ESEGUIBILI

Area che comprende file eseguibili rinvenuti in cartelle create dall'utente:

- prodotti dall'utente (es. con applicazioni software presenti sulla macchina),
- scaricati dal web,
- ricevuti via email,
- ecc.

a. Feature dai metadati:

1. Nome file
2. Data creazione file
3. Data modifica file
4. Data accesso file
5. Data cancellazione
6. Path file
7. MD5-SHA1

b. Feature dal contenuto:

1. Risultato dell'esecuzione del file (scopo dell'applicazione)

12.4 F – Insieme di feature

Una volta analizzate le diverse aree, se ne ricava un set di feature di base. Per insieme di Feature F si intende quindi l'insieme di tutte le singole caratteristiche di base analizzabili in un dispositivo digitale.

$$F = \{ f_1(A_i)(D_i), f_2(A_i)(D_i), \dots, f_n(A_i)(D_i) \}$$

ESEMPIO:

Si riporta qui un elenco di base di tipologie di feature relative ai file contenuti nella memoria di un Personal Computer.

Si tratta di un elenco generico, che va di volta in volta personalizzato secondo l'obiettivo dell'analisi stessa.

Feature rilevabili dai metadati dei file:

1. Nome file
2. Data creazione file
3. Data modifica file
4. Data accesso file
5. Data cancellazione file
6. Path file
7. MD5-SHA1 file

Feature ricavabili dai file di registro:

8. nome computer
9. nome utente di sistema
10. nome utente di dominio
11. numero accessi alla macchina
12. data-ora ultimo accesso
13. data-ora ultimo cambio password
14. Dispositivo hardware installato: tipologia del dispositivo
15. Dispositivo hardware installato: nome –marca-modello del dispositivo
16. Dispositivo hardware installato: numero di serie del dispositivo
17. Dispositivo hardware installato: data ultimo accesso
18. Software installato: nome, marca, versione
19. Software installato: numero di serie
20. Software installato: data di installazione

Feature rilevabili dai file personali (testi, email, chat, file di immagini, audio, video)

21. Firma
22. Nickname
23. Nome proprio
24. Password di accesso
25. Espressione idiomatica
26. Errore di ortografia
27. Errore di battitura
28. Riferimento ad un determinato evento
29. Riferimento ad una determinata persona
30. Riferimento ad un determinato oggetto
31. Riferimento ad un determinato luogo

32. Riferimento ad un determinato dato
33. Frase di testo particolare
34. Indirizzi email mittente
35. Indirizzi email destinatari
36. Data invio
37. Data ricezione
38. Data cancellazione
39. Data – ora chat
40. Immagine fotografica di una determinato persona
41. Immagine fotografica di un determinato luogo
42. Immagine fotografica di un determinato oggetto
43. Immagine fotografica utilizzata come sfondo del desktop
44. Immagine video di una determinato persona
45. Immagine video di un determinato luogo
46. Immagine video di un determinato oggetto
47. Contenuto della traccia audio: accento
48. Contenuto della traccia audio: difetto di pronuncia
49. Contenuto della traccia audio: frase particolare

Feature rilevabili dagli URL a pagine web:

50. URL a posta elettronica via web
51. URL a collegamenti FTP
52. URL a conti online
53. URL a webpage con accesso autenticato

Feature rilevabili dalle cronologie internet:

54. nome utente che ha effettuato l'accesso
55. indirizzo URL della pagina web visitata
56. data e ora accesso

Feature rilevabili dai file di listati di codice, file eseguibili:

57. Codice di programmazione : commento al codice

12.5 m – feature minima

Una volta fissato l'insieme di massima delle possibili feature estraibili dal dispositivo, esso va ridimensionato alle caratteristiche effettivamente presenti sul dispositivo in analisi, secondo le specifiche esigenze dell'indagine.

Allo scopo viene effettuata una prima selezione delle feature, relative ad un dato dispositivo, che ne restringe il numero, andando a formare l'insieme di feature minime.

Per m_i si intende quindi una feature consistente, appartenente all'insieme delle feature di base, selezionata in relazione alla singola indagine trattata.

$$m_i(A_i)(D_i) \in F(D_i)$$

La denominazione della feature minima è quindi:

$m_i(A_i)(D_i)$ ove:

- m_i – identifica la feature minima;
- A_i – identifica l'area di appartenenza del file di origine;
- D_i - identifica il dispositivo digitale da cui è stata estratta.

12.6 M – Insieme di feature minime

Si definisce l'insieme delle feature minime come sottoinsieme di $S(D_i)$, di numerosità minima in rapporto al singolo caso investigativo.

$$M(D_i) \in F(D_i)$$

$$M(D_i) = \{ m_1(A_i)(D_i), \{ m_2(A_i)(D_i), \dots \{ m_n(A_i)(D_i) \}$$

L'insieme di feature minime costituisce l'insieme di filtri da applicare ai file per l'estrazione delle informazioni caratterizzanti (*indicatori*) che andranno a comporre il profilo digitale.

ESEMPIO:

Feature dai metadati dei file:

1. Nome file
2. Data creazione file
3. Data modifica file
4. Data accesso file
5. Data cancellazione file
6. Path file
7. MD5-SHA1 file

Feature dai file di registro:

8. nome computer
9. nome utente di sistema
10. Dispositivo hardware installato: nome –marca-modello del dispositivo
11. Dispositivo hardware installato: numero di serie del dispositivo
12. Dispositivo hardware installato: data ultimo accesso
13. Software installato: nome, marca, versione
14. Software installato: numero di serie

Feature dai file personali:

15. Firma
16. Nickname
17. Nome proprio
18. Password di accesso
19. Espressione idiomatica
20. Errore di ortografia
21. Errore di battitura
22. Riferimento ad un determinato evento
23. Riferimento ad una determinato persona
24. Riferimento ad un determinato oggetto
25. Riferimento ad un determinato luogo
26. Riferimento ad un determinato dato
27. Frase di testo particolare
28. Indirizzi email

29. Data – ora chat
30. Immagine fotografica di una determinata persona
31. Immagine fotografica di un determinato luogo
32. Immagine fotografica di un determinato oggetto
33. Immagine fotografica utilizzata come sfondo del desktop
34. Immagine video di una determinata persona
35. Immagine video di un determinato luogo
36. Immagine video di un determinato oggetto

Feature rilevabili dagli URL a pagine web:

37. URL a posta elettronica via web
38. URL a collegamenti FTP
39. URL a conti online
40. URL a webpage con accesso autenticato

Feature rilevabili dalle cronologie internet:

41. nome utente che ha effettuato l'accesso
42. indirizzo URL della pagina web visitata
43. data e ora accesso

12.7 i - Indicatore

L'indicatore rappresenta la singola informazione caratterizzante rilevata e analizzabile nel contesto di studio a fini di profiling.

Esso si ricava dai file selezionati mediante l'applicazione delle feature minime come filtri durante l'operazione di creazione del Profilo digitale.

Si definisce quindi $i_i(I_i)(A_i)(D_i)$ l'informazione rilevata applicando come filtro, agli insiemi di file, la feature minima m_i , relativa a una determinata area (A_i) di uno specifico dispositivo (D_i). (I_i) identifica univocamente il file da cui l'indicatore è stato estratto.

L'indicatore è, a tutti gli effetti, una *traccia digitale*, che come tale può essere:

- rilevata;
- confrontata;
- riconosciuta.

12.8 I - Insieme degli indicatori

Si definisce l'insieme degli indicatori $I(D_i)$:

$$I(D_i) = \{ i_1(I_i)(A_i)(D_i) \ i_2(I_i)(A_i)(D_i) \ \dots \ i_n(I_i)(A_i)(D_i) \}$$

L'insieme degli indicatori rappresenta tutte le informazioni caratterizzanti rilevate dai file. Esso descrive il "comportamento digitale" dell'utente utilizzatore del dispositivo in analisi.

12.9 k – File contenente indicatori

$k_i(A_i)(D_i)$ identifica univocamente ogni file che contiene uno o più indicatori, ove:

- (A_i) identifica l'area in cui è stato rinvenuto il file;
- (D_i) identifica il dispositivo.

Il file che contiene uno o più indicatori costituisce la "*fonte di prova digitale*" che attesta l'origine dell'indicatore, cioè l'informazione di interesse da esso estratta.

12.10 K – Insieme dei file contenenti gli indicatori

$K(D_i)$ definisce l'insieme dei file che contengono indicatori relativi ad un determinato dispositivo (D_i) .

$$K(D_i) = \{ k_1(A_i)(D_i) \ k_2(A_i)(D_i) \ \dots \ k_n(A_i)(D_i) \}$$

13. LA SEQUENZA DELLE OPERAZIONI PER LA CREAZIONE DEL PROFILO DIGITALE

Si descrivono qui le operazioni che porteranno alla creazione del Profilo digitale, attraverso una serie di selezioni operate su insiemi di file, per mezzo dei filtri ricavati dalle feature minime.

Data la grande quantità di file che si vanno a manipolare, per lo svolgimento di tutte le operazioni si rende indispensabile l'utilizzo di un tool di analisi forense in grado di:

- Importare e gestire le bitstream image di diversi dispositivi;
- indicizzare e numerare i file in modo univoco;
- esportare i file calcolandone gli algoritmi di hash;
- creare insiemi di file distinti fra loro;
- effettuare operazioni su insiemi di file applicando filtri personalizzati;
- riconoscere file cancellati, deallocati, compressi e rinominati;
- visualizzare il contenuto della maggior parte dei file;
- visualizzare i file di registro;
- registrare i passi dell'analisi ed esportarne il risultato.

(es. FTK di Accessdata, Encase di Guidance Software, ecc.).

La sequenza di operazioni comprende la successiva estrapolazione di n.4 profili da un PC:

- il Profilo ricavato dai file di registro;
- il Profilo ricavato dai file contenuti nella cartella utente;
- il Profilo ricavato dai file nelle rimanenti aree di memoria.

Da essi si ricava:

- il Profilo utente, formato dalla loro unione;
- il Profilo campione, che coincide con il Profilo utente, ma si riferisce ad un dispositivo scelto come "campione" per il confronto con altri.

Dal Profilo campione si traggono:

- gli Indicatori, cioè le informazioni caratterizzanti da utilizzarsi per il confronto con altri profili per il riconoscimento dell'utente;
- i file che li contengono.

Tenendo ben presente che in un PC è possibile rilevare la presenza di più utenti, nella seguente esplicazione del metodo, si utilizza come esempio la realizzazione del profilo digitale di un Personal Computer attribuito ad un unico utente relativamente ad un solo sistema operativo Microsoft Windows.

Per semplicità, la numerosità degli Indicatori estratti nell'esempio è minima.

13.1 Ps - Profilo di sistema

Punto di partenza sono i file di registro (Area A_1), che forniscono tutte le informazioni (*indicatori*) relative alla configurazione della macchina da parte dell'utente.

Esse andranno a costituire il Profilo di sistema $\mathbf{Ps}_i(\mathbf{D}_i)$, ove (\mathbf{D}_i) identifica un determinato dispositivo.

$\mathbf{Ps}_i(\mathbf{D}_i)$ rappresenta il Profilo di sistema, così definito:

$$\mathbf{Ps}_i(\mathbf{D}_i) = \mathbf{I}(\mathbf{Ps}_i)(\mathbf{D}_i) \cup \mathbf{K}(\mathbf{Ps}_i)(\mathbf{D}_i)$$

ove:

- l'insieme degli indicatori rilevati dai file di registro, definito $\mathbf{I}(\mathbf{Ps}_i)(\mathbf{D}_i)$, in cui:
 - \mathbf{I} - insieme degli indicatori rilevati
 - \mathbf{Ps}_i - identifica lo specifico Profilo di sistema
 - \mathbf{D}_i - identifica lo specifico Dispositivo
- l'insieme dei file che li contiene, definito $\mathbf{K}(\mathbf{Ps}_i)(\mathbf{D}_i)$, in cui:
 - \mathbf{K} - insieme di file
 - \mathbf{Ps}_i - identifica lo specifico Profilo di sistema
 - \mathbf{D}_i - identifica lo specifico Dispositivo

in cui:

- ogni indicatore è costituita da una singola informazione non ulteriormente scomponibile;
- ogni indicatore si riferisce ad uno o più file;
- ogni file può contenere uno o più indicatori.

ESEMPIO: Personal Computer - \mathbf{D}_1

L'Area A_1 contiene i seguenti file di registro che contengono feature:

- SAM
- NTUSER.dat
- SYSTEM.dat
- SOFTWARE.dat

La loro analisi porta alla definizione dei seguenti indicatori:

- | | |
|---------------------------------|---------------|
| 1. nome computer | PC_SuperPippo |
| 2. nome utente di sistema | SuperPippo |
| 3. hardware installato-nome | USBpen Trust |
| 4. hardware installato-seriale | A01234567 |
| 5. software installato-nome | AVAST v1.34 |
| 6. software installato:-seriale | AD1234DC1234 |

Essi costituiscono i primi filtri da applicare per il confronto dei profili.

Il Profilo di sistema $\mathbf{Ps}_i(\mathbf{D}_i)$ risulta così composto:

FILE DI RIFERIMENTO	FEATURE	INDICATORE
$k_1(A_1)(D_1)$ - SAM	m8(A_1) - nome computer	$i_1(k_1)(A_1)(D_1)$ - PC_SuperPippo
	m9(A_1) - nome utente di sistema	$i_2(k_2)(A_1)(D_1)$ - SuperPippo
$k_2(A_1)(D_1)$ - SYSTEM.DAT	m10(A_1) - hardware installato-nome	$i_3(k_3)(A_1)(D_1)$ - USBpen Trust
	m14(A_1) - hardware installato-seriale	$i_4(k_4)(A_1)(D_1)$ - A01234567
$k_3(A_1)(D_1)$ - SOFTWARE.DAT	m13(A_1) - software installato: nome	$i_5(k_5)(A_1)(D_1)$ - AVAST v1.34
	m14(A_1) - software installato: seriale	$i_6(k_6)(A_1)(D_1)$ - AD1234DC1234

13.2 Pc - Profilo cartella utente

Secondo passo è l'analisi dei file memorizzati nelle cartelle utenti create dal sistema operativo. In esse infatti, è contenuta la maggior parte delle "personalizzazioni" apportate dall'utente.

Si crea quindi il profilo denominato $Pc_i(D_i)$ (Profilo cartella utente), basato sull'analisi dei file contenuti nella cartella utente creata dal sistema operativo nel dispositivo D_i .

Esiste un Pc_u per ogni cartella utente rilevata nella memoria del PC .
(es.: D_1 : PC con S.O. Windows XP : c:\Documents and Settings\ SuperPippo\...).

Se esistono più sistemi operativi (ivi compresi i SO contenuti in macchine virtuali), ognuno di essi va considerato come un diverso dispositivo.

Il Profilo cartella utente $Pc_i(D_i)$ si definisce come:

$$Pc_i(D_i) = I(Pc_i)(D_i) \cup K(Pc_i)(D_i)$$

ove:

- $I(Pc_i)(D_i)$ è l'insieme degli indicatori rilevati dai file contenuti nella cartella utente, in cui:
 - I_i - insieme degli indicatori rilevati
 - Pc_i - identifica il profilo cartella utente
 - D_i - identifica il Dispositivo
- $K(Pc_i)(D_i)$ è l'insieme dei file che li contiene, in cui:
 - K - insieme di file
 - Pc_i - identifica il Profilo cartella utente
 - D_i - identifica il Dispositivo

in cui:

- ogni indicatore è costituita da una singola informazione non ulteriormente scomponibile;
- ogni indicatore si riferisce ad uno o più file;
- ogni file può contenere uno o più indicatori.

ESEMPIO: D₁: PC con S.O. Windows XP - cartella utente:
c:\Documents and Settings\Pippo\.

Da esso si ricavano tanti sottoinsiemi di file (compresi i file cancellati o deallocati) quante sono le aree di interesse A_i che si rilevano:

A₂ – file di testo personali (txt ,doc , docx , rtf , odt , ecc)

A₃ - file di messaggi email (inviati, ricevuti, eliminati)

A₄ - file di conversazioni chat (dat, ecc.)

A₅ - file di scatti fotografici (bmp, jpg ecc.)

A₆ - file grafici (bmp, jpg, png, dwg, cdr. ecc.)

A₇ – file video (mpg, avi, vob, ecc.)

A₈ - file audio (wav, mp3, ecc.)

A₉ - URL (cookies, file di cronologia browser).

L' analisi dei file ne seleziona quelli di interesse, quelli cioè che contengono feature:

A₂ – n. 2 file di testo personali

- o *Notamia.txt*

Da cui vengono estratti i seguenti indicatori:

- o nomefile: Notamia.txt
- o path: c:\Documents and Settings\SuperPippo\Desktop\XXX\
- o nickname: superpippo

- o *xxx.doc*

Da cui vengono estratte estratti i seguenti indicatori:

- o nomefile: xxx.doc
- o path: c:\Documents and Settings\SuperPippo\Desktop\XXX\
- o riferimento a persona: ilgiaguaro

A₃ - n. 3 file di messaggi email

1. *message01.eml*

2. *message02.eml*

3. *message03.eml*

Da cui vengono estratti i seguenti indicatori:

- o indirizzo email: ilgiaguaro@jahoo.com

A₄ - n. 1 file di conversazione chat Skype

- *0261f112b3f57021.dat*

Da cui vengono estratti i seguenti indicatori:

- o nickname: ilgiaguaro
- o nickname: superpippo
- o espressione idiomatica: ola hombre
- o frase particolare: non mi hai lasciato le sigarette nel solito posto ieri
- o riferimento a oggetto: sigarette

- riferimento a luogo: solito posto
- riferimento a data: 24/07/2010 (*nel testo: "ieri"*)
- riferimento a evento: mancata consegna
- riferimento a persona: giaguaro

A₅ - n. 1 file di scatto fotografico

- *DSC_0001.jpg*

Da cui si ricavano i seguenti indicatori:

- Nomefile: DSC_0001.jpg
- Path: c:\Documents and Settings\SuperPippo\101ND040\
- Immagine di determinato oggetto: particolare di auto gialla con targa MI01234567
- Riferimento a oggetto: targa MI01234567
- Riferimento a dato: MI01234567

A₆ - n. 3 file grafici

- *Dvd01.tif*
- *Dvd02.tif*
- *Dvd03.tif*

Da cui si ricavano i seguenti indicatori:

- Nomefile: Dvd01.tif
- Nomefile: Dvd02.tif
- Nomefile: Dvd03.tif
- Path: c:\Documents and Settings\SuperPippo\Desktop\XXX\copertine dvd\

A₈ - file audio

- *La cumparsita.mp3*
- *El dindondero.mp3*

Da cui si ricavano i seguenti indicatori:

- Nomefile: La cumparsita.mp3
- Nomefile: El dindondero.mp3
- Path: c:\Documents and Settings\SuperPippo\Desktop\XXX\miomp3\

A₉ - URL

- *History.dat*

Da cui si ricavano i seguenti indicatori:

- <http://www.facebook.com/superpippo2345cdk0945.php>
- <http://www.ilmiosito.com/superpippo234sdfgoap43.php>
- <http://www.lamiaposta.com/superpippo3456asdf567.php>

ORGANIZZAZIONE CARTELLE

Organizzazione di file e cartelle personali:

c:\Documents and Settings\SuperPippo\Desktop\XXX\
c:\Documents and Settings\SuperPippo\Desktop\XXX\copertine dvd\
c:\Documents and Settings\SuperPippo\Desktop\XXX\miomp3\
c:\Documents and Settings\SuperPippo\101ND040\

Il Profilo cartella utente $Pc_1(D_1)$ risulta così composto:

FILE DI RIFERIMENTO	FEATURE	INDICATORE
$k_4(A_2)(D_1)$ - XXX.DOC	$m_1(A_2)$ - Nome file	$i_7(k_4)(A_2)(D_1)$ - xxx.doc
	$m_6(A_2)$ - Path	$i_8(k_4)(A_2)(D_1)$ - c:\Documents and Settings\SuperPippo\Desktop\XXX\
	$m_{16}(A_2)$ - nickname	$i_9(k_4)(A_2)(D_1)$ - ilgiaguaro
	$m_7(A_2)$ - MD5	B1E5CBE1E019E12E5B73EB4AFB619B5A
$k_5(A_2)(D_1)$ - NOTAMIA.TXT	$m_1(A_2)$ - Nome file	$i_{10}(k_5)(A_2)(D_1)$ - Notamia.txt
	$m_{16}(A_2)$ - nickname	$i_{11}(k_5)(A_2)(D_1)$ - superpippo
	$m_6(A_2)$ - Path	$i_{12}(k_5)(A_2)(D_1)$ - c:\Documents and Settings\SuperPippo\Desktop\XXX\
	$m_7(A_2)$ - MD5	$i_{13}(k_5)(A_2)(D_1)$ - C1E5CBE1E019E12E5B73EB4AFB619B5A
$k_6(A_3)(D_1)$ - message01.eml	$m_{28}(A_3)$ - Indirizzo email	$i_{14}(k_6)(A_3)(D_1)$ - superpippo@lamiaposta.com
	$m_{28}(A_3)$ - Indirizzo email	$i_{15}(k_6)(A_3)(D_1)$ - ilgiaguaro@jahoo.com
$k_7(A_3)(D_1)$ - message02.eml	$m_{28}(A_3)$ - Indirizzo email	$i_{16}(k_7)(A_3)(D_1)$ - superpippo@lamiaposta.com
	$m_{28}(A_3)$ - Indirizzo email	$i_{17}(k_7)(A_3)(D_1)$ - ilgiaguaro@jahoo.com
$k_8(A_3)(D_1)$ - message03.eml	$m_{28}(A_3)$ - Indirizzo email	$i_{18}(k_8)(A_3)(D_1)$ - superpippo@lamiaposta.com
	$m_{28}(A_3)$ - Indirizzo email	$i_{19}(k_8)(A_3)(D_1)$ - ilgiaguaro@jahoo.com
$k_9(A_4)(D_1)$ - 0261f112b3f57021.dat	$m_{19}(A_4)$ - espressione idiomatica	$i_{20}(k_9)(A_4)(D_1)$ - ola hombre
	$m_{16}(A_4)$ - Nickname	$i_{21}(k_9)(A_4)(D_1)$ - ilgiaguaro
	$m_{16}(A_4)$ - Nickname	$i_{22}(k_9)(A_4)(D_1)$ - superpippo
	$m_{27}(A_4)$ - frase particolare	$i_{23}(k_9)(A_4)(D_1)$ - non mi hai lasciato le sigarette nel solito posto ieri
	$m_{24}(A_4)$ - riferimento a oggetto	$i_{24}(k_9)(A_4)(D_1)$ - sigarette
	$m_{25}(A_4)$ - riferimento a luogo	$i_{25}(k_9)(A_4)(D_1)$ - solito posto
	$m_{26}(A_4)$ - riferimento a dato	$i_{26}(k_9)(A_4)(D_1)$ - 24/12/2009
	$m_{22}(A_4)$ - riferimento a evento	$i_{27}(k_9)(A_4)(D_1)$ - mancata consegna
$m_{23}(A_4)$ - riferimento a	$i_{28}(k_9)(A_4)(D_1)$ - giaguaro	

	persona	
k ₁₀ (A ₅)(D ₁) - DSC_0001.jpg	m ₁ (A ₅) - nome file	i ₂₉ (k ₁₀)(A ₅)(D ₁) - DSC_0001.jpg
	m ₆ (A ₅) - path	i ₂₃ (k ₁₀)(A ₅)(D ₁) - c:\Documents and Settings\SuperPippo\101ND040\
	m ₃₂ (A ₅) - immagine di determinato oggetto	i ₃₀ (k ₁₀)(A ₅)(D ₁) - particolare di auto gialla con targa MI01234567
	m ₂₄ (A ₅) - riferimento a oggetto	i ₃₁ (k ₁₀)(A ₅)(D ₁) - Automobile gialla
	m ₂₄ (A ₅) - riferimento a oggetto	i ₃₂ (k ₁₀)(A ₅)(D ₁) - targa MI01234567
	m ₇ (A ₅) - MD5	i ₃₃ (k ₁₀)(A ₅)(D ₁) - D1E5CBE1E019E12E5B73EB4AFB619B5A
k ₁₁ (A ₆)(D ₁) - Dvd01.tif	m ₁ (A ₆) - nome file	i ₃₄ (k ₁₁)(A ₆)(D ₁) - Dvd01.tif
	m ₆ (A ₆) - path	i ₃₅ (k ₁₁)(A ₆)(D ₁) - Dvd01.tif c:\Documents and Settings\SuperPippo\Desktop\XXX\copertine dvd\
	m ₇ (A ₆) - MD5	i ₃₆ (k ₁₁)(A ₆)(D ₁) - A2E5CBE1E019E12E5B73EB4AFB619B5A
k ₁₂ (A ₆)(D ₁) - Dvd02.tif	m ₁ (A ₆) - nome file	i ₃₇ (k ₁₂)(A ₆)(D ₁) - Dvd02.tif
	m ₆ (A ₆) - path	i ₃₈ (k ₁₂)(A ₆)(D ₁) - c:\Documents and Settings\SuperPippo\Desktop\XXX\copertine dvd\
	m ₇ (A ₆) - MD5	i ₃₉ (k ₁₂)(A ₆)(D ₁) - A3E5CBE1E019E12E5B73EB4AFB619B5A
k ₁₃ (A ₆)(D ₁) - Dvd03.tif	m ₁ (A ₆) - nome file	i ₄₀ (k ₁₃)(A ₆)(D ₁) - Dvd03.tif
	m ₆ (A ₆) - path	i ₄₁ (k ₁₃)(A ₆)(D ₁) - c:\Documents and Settings\SuperPippo\Desktop\XXX\copertine dvd\
	m ₇ (A ₆) - MD5	i ₄₂ (k ₁₃)(A ₆)(D ₁) - B6E5CBE1E019E12E5B73EB4AFB619B5A
k ₁₄ (A ₈)(D ₁) - La cumparsita.mp3	m ₁ (A ₈) - nome file	i ₄₃ (k ₁₄)(A ₈)(D ₁) - La cumparsita.mp3
	m ₆ (A ₈) - path	i ₄₃ (k ₁₄)(A ₈)(D ₁) - c:\Documents and Settings\SuperPippo\Desktop\XXX\miomp3\
	m ₇ (A ₈) - MD5	i ₄₄ (k ₁₄)(A ₈)(D ₁) - C3E5CBE1E019E12E5B73EB4AFB619B5A
k ₁₅ (A ₈)(D ₁) - El dindondero.mp3	m ₁ (A ₈) - nome file	i ₄₅ (k ₁₄)(A ₈)(D ₁) - El dindondero.mp3
	m ₆ (A ₈) - path	c:\Documents and Settings\SuperPippo\Desktop\XXX\miomp3\
	m ₇ (A ₈) - MD5	i ₄₆ (k ₁₄)(A ₈)(D ₁) - E6E5CBE1E019E12E5B73EB4AFB619B5A
k ₁₆ (A ₉)(D ₁) - History.dat	m ₄₀ (A ₉) - URL	i ₄₇ (k ₁₆)(A ₉)(D ₁) - http://www.facebook.com/superpippo2345cdk0945.php
	m ₄₀ (A ₉) - URL	i ₄₈ (k ₁₆)(A ₉)(D ₁) - http://www.ilmiosito.com/superpippo234sdfgoap43.php
	m ₃₇ (A ₉) - URL	i ₄₉ (k ₁₆)(A ₉)(D ₁) - http://www.lamiaposta.com/superpippo3456asdf567.php

13.3 Pd - Profilo dispositivo

La creazione del Profilo cartella utente non è sufficiente a delineare l'intero profilo dell'utente della macchina, poiché altre feature possono essere rilevate da file memorizzati in aree non comprese nelle cartelle utenti canoniche.

Il Profilo dispositivo comprende quei file, ad esempio, contenuti in cartelle di archivio su altre partizioni, su ulteriori dischi fissi, comprendendo anche i file deallocati, ecc..

Viene effettuata perciò una seconda selezione, utilizzando le feature dell'insieme M (feature minime), volta a evidenziare tutti quei file contenenti feature memorizzati al di fuori delle cartelle utenti.

Il Profilo dispositivo $\mathbf{Pd}_i(\mathbf{D}_i)$, si definisce come:

$$\mathbf{Pd}_i(\mathbf{D}_i) = \mathbf{I}(\mathbf{Pd}_i)(\mathbf{D}_i) \cup \mathbf{K}(\mathbf{Pd}_i)(\mathbf{D}_i)$$

ove:

- l'insieme degli indicatori estratti dai file non contenuti nella cartella utente, definito $\mathbf{I}_i(\mathbf{Pd}_i)(\mathbf{D}_i)$, in cui:
 - \mathbf{I} - insieme degli indicatori rilevati
 - \mathbf{Pd}_i - identifica il profilo dispositivo
 - \mathbf{D}_i , - identifica il Dispositivo
- l'insieme dei file che li contiene, definito $\mathbf{K}_i(\mathbf{Pd}_i)(\mathbf{D}_i)$, in cui:
 - \mathbf{K} - insieme di file
 - \mathbf{Pd}_i - identifica il Profilo dispositivo
 - \mathbf{D}_i , - identifica il Dispositivo

in cui:

- ogni indicatore è costituita da una singola informazione non ulteriormente scomponibile;
- ogni indicatore si riferisce ad uno o più file;
- ogni file può contenere uno o più indicatori.

ESEMPIO: D₁: - PC con S.O. Windows XP.

Da una prima selezione si ricavano tanti sottoinsiemi di file (compresi i file cancellati o deallocati) quante sono le aree di interesse A_i che si rilevano:

A₂ – file di testo personali (txt ,doc , docx , rtf , odt , ecc)

A₅ - file di scatti fotografici (bmp, jpg ecc.)

A₇ – file video (mpg, avi, vob, 3gp, ecc.)

L' analisi dei file ne seleziona quelli di interesse, quelli cioè che contengono feature:

A₂ – n. 1 file deallocato di testo personale:

○ *Carved[123456789].doc*

Da cui vengono estratti i seguenti indicatori:

- | | |
|-------------------------|----------------------------------|
| ○ MD5 | D1E9ABE1E009E12E5B23EB4DFB689B5E |
| ○ nickname | superpippo |
| ○ password | piùvelocedellaluce |
| ○ indirizzo email | superpippo@lamiaposta.com |
| ○ riferimento a oggetto | 339123456 |

A₅ - n. 1 file deallocato di scatto fotografico

- *Carved[123456749].jpg*

Da cui si ricavano i seguenti indicatori:

- | | |
|-----------------------|---|
| ○ MD5 | A1E5CBE1E019E12E5B73EB4AFB619B5A |
| ○ Immagine di oggetto | particolare di carta di credito Bankamericard |
| ○ Riferimento a dato | Bankamericard |
| ○ Riferimento a dato | 4935 1500 4556 5784 |

A₇ - n. 1 file video

- *Carved[123451049].3gp*

Da cui si ricavano i seguenti indicatori:

- | | |
|-----------------------------------|----------------------------------|
| ○ MD5 | B1E5CBE1E019E13E5B73EB4AFB619B5D |
| ○ Immagine di determinata persona | Rossi Mario |
| ○ Riferimento a persona | Rossi Mario |

Il Profilo dispositivo **Pd₁(D₁)** risulta perciò così composto:

FILE DI RIFERIMENTO	FEATURE	INDICATORE
K ₁₇ (A ₂)(D ₁) - carved[123456789].doc	m ₁₆ (A ₂) - nickname	i ₅₀ (k ₁₇)(A ₂)(D ₁) - superpippo
	m ₁₈ (A ₂) - password	i ₅₁ (k ₁₇)(A ₂)(D ₁) - piùvelocedellaluce
	m ₂₈ (A ₂) - indirizzo email	i ₅₂ (k ₁₇)(A ₂)(D ₁) - superpippo@lamiaposta.com
	m ₂₆ (A ₂) - riferimento a dato	i ₅₃ (k ₁₇)(A ₂)(D ₁) - 339123456
	m ₇ (A ₂) - MD5	i ₅₄ (k ₁₇)(A ₂)(D ₁) - D1E9ABE1E009E12E5B23EB4DFB689B5E
K ₁₈ (A ₅)(D ₁) - carved[123456749].jpg	m ₃₂ (A ₅) - immagine di oggetto	i ₅₅ (k ₁₇)(A ₅)(D ₁) - particolare di carta di credito Bankamericard
	m ₂₆ (A ₂) - riferimento a dato	i ₅₆ (k ₁₇)(A ₅)(D ₁) - Bankamericard
	m ₂₆ (A ₂) - riferimento a dato	i ₅₇ (k ₁₇)(A ₅)(D ₁) - 4935 1500 4556 5784
	m ₇ (A ₂) - MD5	i ₅₈ (k ₁₇)(A ₅)(D ₁) - A1E5CBE1E019E12E5B73EB4AFB619B5A
K ₁₈ (A ₉)(D ₁) - carved[123451049].3gp	m ₃₀ (A ₂) - immagine di determinata persona	i ₅₉ (k ₁₇)(A ₉)(D ₁) - Rossi Mario
	m ₂₃ (A ₂) - riferimento a persona	i ₆₀ (k ₁₇)(A ₂)(D ₁) - Rossi Mario
	m ₇ (A ₂) - MD5	i ₆₁ (k ₁₇)(A ₂)(D ₁) - B1E5CBE1E019E13E5B73EB4AFB619B5D

13.4 Pu - Profilo utente

I profili estrapolati fin qui rappresentano tutti gli elementi necessari per la creazione del profilo utente $Pu(D_i)$.

Esso rappresenta il modello comportamentale digitale che descrive l'interazione dell'utente con il dispositivo digitale analizzato.

Esso è perciò composto da:

- tutte le informazioni caratterizzanti (*indicatori*) rilevate sull'intera macchina durante l'analisi;
- tutti i file che le contengono.

Il Profilo utente $Pu(D_i)$ è quindi definito da:

$$Pu(D_i) = I(Pu)(D_i) \cup K(Pu)(D_i)$$

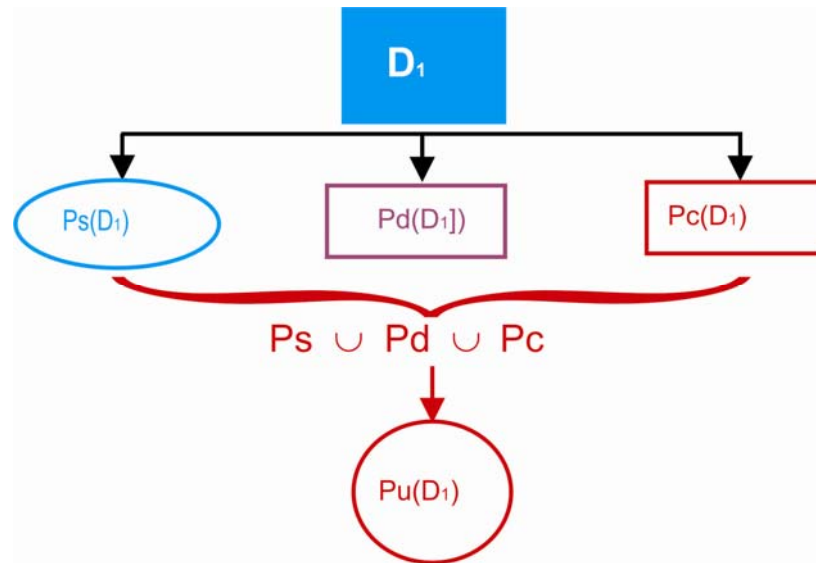
ove:

- $I(Pu)(D_i)$ – insieme nato dall'unione dei tre insiemi di indicatori rilevati:
 - $I(Ps)(D_i) \cup I(Pc)(D_i) \cup I(Pd)(D_i)$
- $K(Pu)(D_i)$ – insieme nato dall'unione dei tre insiemi di file:
 - $K(Ps)(D_i) \cup K(Pc)(D_i) \cup K(Pd)(D_i)$

in cui:

- ogni indicatore è costituito da una singola informazione non ulteriormente scomponibile;
- ogni indicatore si riferisce ad uno o più file;
- ogni file può contenere uno o più indicatori.

Figura 1 – Il profilo utente Pu.



ESEMPIO: D_1 : - PC con S.O. Windows XP: Profilo utente.

FILE DI RIFERIMENTO	FEATURE	INDICATORE
$k_1(A_1)(D_1)$ - SAM	$m_8(A_1)$ - nome computer	$i_1(k_1)(A_1)(D_1)$ - PC_SuperPippo
	$m_9(A_1)$ - nome utente di sistema	$i_2(k_2)(A_1)(D_1)$ - SuperPippo
$k_2(A_1)(D_1)$ - SYSTEM.DAT	$m_{10}(A_1)$ - hardware installato-nome	$i_3(k_3)(A_1)(D_1)$ - USBpen Trust
	$m_{14}(A_1)$ - hardware installato-seriale	$i_4(k_4)(A_1)(D_1)$ - A01234567
$k_3(A_1)(D_1)$ - SOFTWARE.DAT	$m_{13}(A_1)$ - software installato: nome	$i_5(k_5)(A_1)(D_1)$ - AVAST v1.34
	$m_{14}(A_1)$ - software installato: seriale	$i_6(k_6)(A_1)(D_1)$ - AD1234DC1234
$k_4(A_2)(D_1)$ - XXX.DOC	$m_1(A_2)$ - Nome file	$i_7(k_4)(A_2)(D_1)$ - xxx.doc
	$m_6(A_2)$ - Path	$i_8(k_4)(A_2)(D_1)$ - c:\Documents and Settings\SuperPippo\Desktop\XXX\
	$m_{16}(A_2)$ - nickname	$i_9(k_4)(A_2)(D_1)$ - ilgiaguaro
	$m_7(A_2)$ - MD5	B1E5CBE1E019E12E5B73EB4AFB619B5A
$k_5(A_2)(D_1)$ - NOTAMIA.TXT	$m_1(A_2)$ - Nome file	$i_{10}(k_5)(A_2)(D_1)$ - Notamia.txt
	$m_{16}(A_2)$ - nickname	$i_{11}(k_5)(A_2)(D_1)$ - superpippo
	$m_6(A_2)$ - Path	$i_{12}(k_5)(A_2)(D_1)$ - c:\Documents and Settings\SuperPippo\Desktop\XXX\
	$m_7(A_2)$ - MD5	$i_{13}(k_5)(A_2)(D_1)$ - C1E5CBE1E019E12E5B73EB4AFB619B5A
$k_6(A_3)(D_1)$ - message01.eml	$m_{28}(A_3)$ - Indirizzo email	$i_{14}(k_6)(A_3)(D_1)$ - superpippo@lamiaposta.com
	$m_{28}(A_3)$ - Indirizzo email	$i_{15}(k_6)(A_3)(D_1)$ - ilgiaguaro@jahoo.com
$k_7(A_3)(D_1)$ - message02.eml	$m_{28}(A_3)$ - Indirizzo email	$i_{16}(k_7)(A_3)(D_1)$ - superpippo@lamiaposta.com
	$m_{28}(A_3)$ - Indirizzo email	$i_{17}(k_7)(A_3)(D_1)$ - ilgiaguaro@jahoo.com
$k_8(A_3)(D_1)$ - message03.eml	$m_{28}(A_3)$ - Indirizzo email	$i_{18}(k_8)(A_3)(D_1)$ - superpippo@lamiaposta.com
	$m_{28}(A_3)$ - Indirizzo email	$i_{19}(k_8)(A_3)(D_1)$ - ilgiaguaro@jahoo.com
$k_9(A_4)(D_1)$ -	$m_{19}(A_4)$ - espressione	$i_{20}(k_9)(A_4)(D_1)$ - ola hombre

0261f112b3f57021.dat	idiomatica	
	m ₁₆ (A ₄) - Nickname	i ₂₁ (k ₉)(A ₄)(D ₁) - ilgiaguaro
	m ₁₆ (A ₄) - Nickname	i ₂₂ (k ₉)(A ₄)(D ₁) - superpippo
	m ₂₇ (A ₄) - frase particolare	i ₂₃ (k ₉)(A ₄)(D ₁) - non mi hai lasciato le sigarette nel solito posto ieri
	m ₂₄ (A ₄) - riferimento a oggetto	i ₂₄ (k ₉)(A ₄)(D ₁) - sigarette
	m ₂₅ (A ₄) - riferimento a luogo	i ₂₅ (k ₉)(A ₄)(D ₁) - solito posto
	m ₂₆ (A ₄) - riferimento a dato	i ₂₆ (k ₉)(A ₄)(D ₁) - 24/12/2009
	m ₂₂ (A ₄) - riferimento a evento	i ₂₇ (k ₉)(A ₄)(D ₁) - mancata consegna
	m ₂₃ (A ₄) - riferimento a persona	i ₂₈ (k ₉)(A ₄)(D ₁) - giaguaro
k ₁₀ (A ₅)(D ₁) - DSC_0001.jpg	m ₁ (A ₅) - nome file	i ₂₉ (k ₁₀)(A ₅)(D ₁) - DSC_0001.jpg
	m ₆ (A ₅) - path	i ₂₃ (k ₁₀)(A ₅)(D ₁) - c:\Documents and Settings\SuperPippo\101ND040\
	m ₃₂ (A ₅) - immagine di determinato oggetto	i ₃₀ (k ₁₀)(A ₅)(D ₁) - particolare di auto gialla con targa MIO1234567
	m ₂₄ (A ₅) - riferimento a oggetto	i ₃₁ (k ₁₀)(A ₅)(D ₁) - Automobile gialla
	m ₂₄ (A ₅) - riferimento a oggetto	i ₃₂ (k ₁₀)(A ₅)(D ₁) - targa MIO1234567
	m ₇ (A ₅) - MD5	i ₃₃ (k ₁₀)(A ₅)(D ₁) - D1E5CBE1E019E12E5B73EB4AFB619B5A
k ₁₁ (A ₆)(D ₁) - Dvd01.tif	m ₁ (A ₆) - nome file	i ₃₄ (k ₁₁)(A ₆)(D ₁) - Dvd01.tif
	m ₆ (A ₆) - path	i ₃₅ (k ₁₁)(A ₆)(D ₁) - Dvd01.tif c:\Documents and Settings\SuperPippo\Desktop\XXX\copertine dvd\
	m ₇ (A ₆) - MD5	i ₃₆ (k ₁₁)(A ₆)(D ₁) - A2E5CBE1E019E12E5B73EB4AFB619B5A
k ₁₂ (A ₆)(D ₁) - Dvd02.tif	m ₁ (A ₆) - nome file	i ₃₇ (k ₁₂)(A ₆)(D ₁) - Dvd02.tif
	m ₆ (A ₆) - path	i ₃₈ (k ₁₂)(A ₆)(D ₁) - c:\Documents and Settings\SuperPippo\Desktop\XXX\copertine dvd\
	m ₇ (A ₆) - MD5	i ₃₉ (k ₁₂)(A ₆)(D ₁) - A3E5CBE1E019E12E5B73EB4AFB619B5A
k ₁₃ (A ₆)(D ₁) - Dvd03.tif	m ₁ (A ₆) - nome file	i ₄₀ (k ₁₃)(A ₆)(D ₁) - Dvd03.tif
	m ₆ (A ₆) - path	i ₄₁ (k ₁₃)(A ₆)(D ₁) - c:\Documents and Settings\SuperPippo\Desktop\XXX\copertine dvd\
	m ₇ (A ₆) - MD5	i ₄₂ (k ₁₃)(A ₆)(D ₁) - B6E5CBE1E019E12E5B73EB4AFB619B5A
k ₁₄ (A ₈)(D ₁) - La cumparsita.mp3	m ₁ (A ₈) - nome file	i ₄₃ (k ₁₄)(A ₈)(D ₁) - La cumparsita.mp3
	m ₆ (A ₈) - path	i ₄₃ (k ₁₄)(A ₈)(D ₁) - c:\Documents and Settings\SuperPippo\Desktop\XXX\miomp3\
	m ₇ (A ₈) - MD5	i ₄₄ (k ₁₄)(A ₈)(D ₁) - C3E5CBE1E019E12E5B73EB4AFB619B5A
k ₁₅ (A ₈)(D ₁) - El dindondero.mp3	m ₁ (A ₈) - nome file	i ₄₅ (k ₁₄)(A ₈)(D ₁) - El dindondero.mp3
	m ₆ (A ₈) - path	c:\Documents and Settings\SuperPippo\Desktop\XXX\miomp3\
	m ₇ (A ₈) - MD5	i ₄₆ (k ₁₄)(A ₈)(D ₁) - E6E5CBE1E019E12E5B73EB4AFB619B5A
k ₁₆ (A ₉)(D ₁) - History.dat	m ₄₀ (A ₉) - URL	i ₄₇ (k ₁₆)(A ₉)(D ₁) - http://www.facebook.com/superpippo2345cdk0945.php

	m ₄₀ (A ₉) - URL	i ₄₈ (k ₁₆)(A ₉)(D ₁) - http://www.ilmiosito.com/superpippo234sdfgoap43.php
	m ₃₇ (A ₉) - URL	i ₄₉ (k ₁₆)(A ₉)(D ₁) - http://www.lamiaposta.com/superpippo3456asdf567.php
K ₁₇ (A ₂)(D ₁) - carved[123456789].doc	m ₁₆ (A ₂) - nickname	i ₅₀ (k ₁₇)(A ₂)(D ₁) - superpippo
	m ₁₈ (A ₂) - password	i ₅₁ (k ₁₇)(A ₂)(D ₁) - piùvelocedellaluce
	m ₂₈ (A ₂) - indirizzo email	i ₅₂ (k ₁₇)(A ₂)(D ₁) - superpippo@lamiaposta.com
	m ₂₆ (A ₂) - riferimento a dato	i ₅₃ (k ₁₇)(A ₂)(D ₁) - 339123456
	m ₇ (A ₂) - MD5	i ₅₄ (k ₁₇)(A ₂)(D ₁) - D1E9ABE1E009E12E5B23EB4DFB689B5E
K ₁₈ (A ₅)(D ₁) - carved[123456749].jpg	m ₃₂ (A ₅) - immagine di oggetto	i ₅₅ (k ₁₇)(A ₅)(D ₁) - particolare di carta di credito Bankamericard
	m ₂₆ (A ₂) - riferimento a dato	i ₅₆ (k ₁₇)(A ₅)(D ₁) - Bankamericard
	m ₂₆ (A ₂) - riferimento a dato	i ₅₇ (k ₁₇)(A ₅)(D ₁) - 4935 1500 4556 5784
	m ₇ (A ₂) - MD5	i ₅₈ (k ₁₇)(A ₅)(D ₁) - A1E5CBE1E019E12E5B73EB4AFB619B5A
K ₁₈ (A ₉)(D ₁) - carved[123451049].3gp	m ₃₀ (A ₂) - immagine di determinata persona	i ₅₉ (k ₁₇)(A ₉)(D ₁) - Rossi Mario
	m ₂₃ (A ₂) - riferimento a persona	i ₆₀ (k ₁₇)(A ₂)(D ₁) - Rossi Mario
	m ₇ (A ₂) - MD5	i ₆₁ (k ₁₇)(A ₂)(D ₁) - B1E5CBE1E019E13E5B73EB4AFB619B5D

ORGANIZZAZIONE CARTELLE

Organizzazione di file e cartelle personali utente "Superpippo" in D₁ .

c:\Documents and Settings\SuperPippo\Desktop\XXX\
c:\Documents and Settings\SuperPippo\Desktop\XXX\copertine dvd\
c:\Documents and Settings\SuperPippo\Desktop\XXX\miomp3\
c:\Documents and Settings\SuperPippo\101ND040\

13.5 Puc - Profilo utente campione

Il profilo utente campione $Puc(D_i)$ coincide con il Profilo utente $Pu(D_i)$, da cui si differenzia solo per definizione in quanto viene fissato come punto di riferimento per il confronto con altri dispositivi.

Di fatto, gli indicatori rilevati saranno utilizzati come filtri per la ricerca di informazioni coincidenti all'interno delle memorie di altri dispositivi.

ESEMPIO:

Dispositivi: n. 3 dispositivi digitali in analisi perché connessi ad un reato:

- D_1 - n. 1 PC (attribuito con certezza ad un determinato soggetto, definito S_1)
- D_2 - n. 1 Netbook (non attribuito)
- D_3 - n. 1 penna USB (non attribuito)

Obiettivo:

Verifica della possibile appartenenza dei dispositivi D_2 e D_3 al soggetto S_1

Operazioni:

- Costruzione del Set di feature minime del D_1 relativamente al soggetto S_1
- Creazione del Profilo utente del D_1 , fissato come $Puc(D_1)$
- Ricerca degli indicatori rilevati dal $Puc(D_1)$ nei file contenuti nelle memorie di D_2 e D_3 .

14 IL CONFRONTO

Una volta ottenuto il Profilo Campione Puc(D₁) da un dispositivo, gli indicatori rilevati vengono utilizzati come filtri per la ricerca degli stessi su altri dispositivi, al fine di rilevare coincidenze e/o divergenze.

ESEMPIO:

DISPOSITIVI:

a) D₁ - n. 1 PC - Profilo Campione Puc(D₁): FILTRI PER CONFRONTO

N.	Tipo	INDICATORE
1	hardware	USBpen Trust sn A01234567
2	software	AVAST v1.34 sn AD1234DC1234
3	File	xxx.doc
4	File	Notamia.txt
5	File	DSC_0001.jpg
6	File	Dvd02.tif
7	File	Dvd01.tif
8	File	La cumparsita.mp3
9	File	El dindondero.mp3
10	Email	superpippo@lamiaposta.com
11	Email	ilgiaguaro@jahoo.com
12	testo	superpippo
14	testo	piùvelocedellaluce
15	testo	ola hombre
16	testo	L'amico del giaguaro
17	testo	ilgiaguaro
18	testo	giaguaro
19	testo	sigarette
20	testo	solito posto
21	testo	Bankamericard
22	testo	Rossi Mario
23	Data/ora	24/12/2009
24	Dati	4935 1500 4556 5784
25	Dati	339123456
26	Dati	MI01234567
27	MD5	D1E5CBE1E019E12E5B73EB4AFB619B5A
28	MD5	C1E5CBE1E019E12E5B73EB4AFB619B5A

29	MD5	A2E5CBE1E019E12E5B73EB4AFB619B5A
30	MD5	A3E5CBE1E019E12E5B73EB4AFB619B5A
31	MD5	B1E5CBE1E019E13E5B73EB4AFB619B5D
32	MD5	D1E9ABE1E009E12E5B23EB4DFB689B5E
33	MD5	A1E5CBE1E019E12E5B73EB4AFB619B5A
34	MD5	B6E5CBE1E019E12E5B73EB4AFB619B5A
35	MD5	C3E5CBE1E019E12E5B73EB4AFB619B5A
36	MD5	B1E5CBE1E019E12E5B73EB4AFB619B5A
37	MD5	E6E5CBE1E019E12E5B73EB4AFB619B5A
38	URL	http://www.facebook.com/superpippo2345cdk0945.php
39	URL	http://www.ilmiosito.com/ superpippo234sdfgoap43.php
40	URL	http://www.lamiaposta.com/superpippo3456asdf567.php

b) D₂ - n. 1 Netbook - esito applicazione filtri alla ricerca: elenco file e indicatori rilevati

N.	Tipo	FILTRO APPLICATO	FILE RILEVATO su D[2]
1	hardware	USBpen Trust sn A01234567	File[1] SYSTEM.DAT
2	software	AVAST v1.34 sn AD1234DC1234	File[2] SOFTWARE.DAT
3	File	Dvd02.tif	c:\Utenti\SuperPippo\Desktop\XXX\copertine dvd\ Dvd02.tif
4	File	Dvd01.tif	c:\Utenti\SuperPippo\Desktop\XXX\copertine dvd\ Dvd01.tif
5	File	La cumparsita.mp3	c:\Utenti\SuperPippo\Desktop\XXX\miomp3\ La cumparsita.mp3
6	File	El dindondero.mp3	c:\Utenti\SuperPippo\Desktop\XXX\miomp3\dindon dero.mp3
7	Path	...\SuperPippo\Desktop\ copertine dvd\ Dvd01.tif	...\SuperPippo\Desktop\ copertine dvd\ Dvd01.tif
8	Path	...\SuperPippo\Desktop\ copertine dvd\ Dvd02.tif	...\SuperPippo\Desktop\ copertine dvd\ Dvd02.tif
9	Path	...\SuperPippo\Desktop\XXX\miomp3\ La cumparsita.mp3	...\SuperPippo\Desktop\XXX\miomp3\ La cumparsita.mp3
10	Path	...\SuperPippo\Desktop\XXX\miomp3\ dindondero.mp3	...\SuperPippo\Desktop\XXX\miomp3\ dindondero.mp3
11	Email	superpippo@lamiaposta.com	Carved[323456749].eml
12	Email	ilgiaguaro@jahoo.com	

13	testo	SuperPippo	0241f112b3f570231.dat
14	testo	piùvelocedellaluce	
15	testo	ola hombre	
16	testo	L'amico del giaguaro	
17	testo	ilgiaguaro	
18	testo	giaguaro	
19	testo	sigarette	
20	testo	solito posto	
21	testo	Bankamericard	0241f112b3f570231.dat
22	testo	4935 1500 4556 5784	
23	testo	339123456	
24	MD5	B1E5CBE1E019E13E5B73EB4AFB619B5D	Ilgiaguaro.3gp
25	URL	http://www.facebook.com/superpippo2345cdk0945.php	History.dat
26	URL	http://www.ilmiosito.com/superpippo234sdfgoap43.php	
27	URL	http://www.lamiaposta.com/superpippo3456asdf567.php	

ORGANIZZAZIONE CARTELLE

Dall'applicazione dei filtri "Path" si è rilevata la medesima organizzazione di file e cartelle personali:

28	c:\Utenti\SuperPippo\Desktop\XXX\
29	c:\Utenti\SuperPippo\Desktop\XXX\copertine dvd\
30	c:\Utenti\SuperPippo\Desktop\XXX\miomp3\

RISULTATO

Dall'applicazione su D₂ di n. 40 filtri ricavati dagli indicatori tratti dal profilo Puc(D₁) sono stati evidenziati n. 30 indicatori coincidenti che hanno rilevato:

1. La medesima organizzazione di 3 cartelle di file personali .

D ₁ (Windows XP)	D ₂ (Windows 7)
c:\Documents and Settings\SuperPippo\Desktop\XXX\	c:\Utenti\SuperPippo\Desktop\XXX\
c:\Documents and Settings\SuperPippo\Desktop\XXX\copertine dvd\	c:\Utenti\SuperPippo\Desktop\XXX\copertine dvd\
c:\Documents and Settings\SuperPippo\Desktop\XXX\miomp3\	c:\Utenti\SuperPippo\Desktop\XXX\miomp3\

2. Il ritrovamento degli stessi 4 file nella medesima posizione:

D ₁ (Windows XP)	D ₂ (Windows 7)
c:\Documents and Settings\SuperPippo\Desktop\copertine dvd\ Dvd02.tif	c:\Utenti\SuperPippo\Desktop\XXX\copertine dvd\ Dvd02.tif
c:\Documents and Settings\SuperPippo\Desktop\copertine dvd\ Dvd01.tif	c:\Utenti\SuperPippo\Desktop\XXX\copertine dvd\ Dvd01.tif
c:\Documents and Settings\SuperPippo\Desktop\XXX\miomp3\ La cumparsita.mp3	c:\Utenti\SuperPippo\Desktop\XXX\miomp3\ La cumparsita.mp3
c:\Documents and Settings\SuperPippo\Desktop\XXX\miomp3\ dindondero.mp3	c:\Utenti\SuperPippo\Desktop\XXX\miomp3\ dindondero.mp3

3. Un messaggio email (file deallocato) **Carved[323456749].eml** in cui sono stati ritrovati:
 - a. lo stesso mittente: superpippo@lamiaposta.com
 - b. lo stesso destinatario: ilgiaguaro@jahoo.com

4. Un file di conversazioni chat: **0241f112b3f570231.dat** nei cui testi sono stati rinvenuti:
- a. Tre nickname coincidenti:
 - i. SuperPippo (mittente)
 - ii. L'amico del giaguaro (destinatario)
 - b. Una password coincidente:
 - i. piùvelocedellaluce (accesso a Skype)
 - c. La medesima espressione idiomatica ripetuta:
 - i. ola hombre
 - d. Il riferimento a persona:
 - i. Ilgiaguaro
 - ii. Giaguaro
 - e. Il riferimento a oggetto:
 - i. Sigarette
 - f. Il riferimento ad un luogo:
 - i. Solito posto
5. Un file di conversazioni chat: **0241f112b3f570231.dat** nei cui testi sono stati rinvenuti:
- g. La parola Bankamericard
 - h. Il numero della carta Bankamericard: 4935 1500 4556 5784
 - i. Il numero di telefono cellulare: 339123456
6. Un file video: **ilgiaguaro.3gp** il cui MD5:
- a. B1E5CBE1E019E13E5B73EB4AFB619B5D corrisponde al file:
 - i. carved[123451049].3gp del D[1] e che mostra:
 1. ROSSI MARIO *alla guida di una*
 2. automobile gialla
 3. *con targa* MI01234567
7. Un file di cronologia internet **History.dat** che contiene i medesimi URL a pagine personali:
- a. <http://www.facebook.com/superpippo2345cdk0945.php>
 - b. <http://www.ilmiosito.com/superpippo234sdfgoap43.php>
 - c. <http://www.lamiaposta.com/superpippo3456asdf567.php>

N.	Feature	Indicatore	D ₁	D ₂
1	Organizzazione cartelle	...\SuperPippo\Desktop\XXX\	●	●
2	Organizzazione cartelle	...\SuperPippo\Desktop\XXX\copertine dvd\	●	●
3	Organizzazione cartelle	...\SuperPippo\Desktop\XXX\miomp3\	●	●
4	Path file	...\SuperPippo\Desktop\ copertine dvd\ Dvd01.tif	●	●
5	Path file	...\SuperPippo\Desktop\ copertine dvd\ Dvd02.tif	●	●
6	Path file	...\SuperPippo\Desktop\XXX\miomp3\ La cumparsita.mp3	●	●
7	Path file	...\SuperPippo\Desktop\XXX\miomp3\ dindondero.mp3	●	●
8	File personali	Dvd01.tif	●	●
9	File personali	Dvd02.tif	●	●
10	File personali	La cumparsita.mp3	●	●
11	File personali	dindondero.mp3	●	●
12	Mittente email	superpippo@lamiaposta.com	●	●
13	Destinatario email	ilgiaguaro@jahoo.com	●	●
14	Nickname mittente Skype	SuperPippo	●	●
15	Nickname destinatario Skype	L'amico del giaguaro	●	●
16	Password Skype	piùvelocedellaluce	●	●
17	Espressione idiomatica	ola hombre	●	●
18	nickname	ilgiaguaro	●	●
19	Frase particolare	non mi hai lasciato le sigarette nel solito posto ieri	●	●
20	Riferimento a oggetto	sigarette	●	●
21	Riferimento a luogo	Solito posto	●	●
22	Riferimento a data	24/07/2010	●	●
23	Riferimento a oggetto	Bankamericard n. s. 4935 1500 4556 5784	●	●
24	Riferimento a n. di telefono	339123456	●	●
25	Riferimento a veicolo	Automobile gialla con targa MI01234567	●	●
26	URL	http://www.facebook.com/superpippo2345cdk0945.php	●	●
27	URL	http://www.ilmiosito.com/superpippo234sdfgoap43.php	●	●
28	URL	http://www.lamiaposta.com/superpippo3456asdf567.php	●	●
29	Hardware	USBpen Trust sn A01234567	●	●
30	Software	AVAST v1.34 sn AD1234DC1234	●	●

Tabella 1 – Prospetto riepilogativo degli indicatori coincidenti.

Passo finale, se il caso lo richiede, è il confronto fra le date di creazione/modifica/cancellazione dei file estrapolati dai due dispositivi, allo scopo di ricostruire la cronologia delle azioni dell'utente sui dispositivi nel tempo.

L'esempio ha illustrato come la ricerca degli indicatori, estrapolati dal dispositivo D_1 , nei file contenuti nella memoria del dispositivo D_2 , ha 30 informazioni caratterizzanti l'utente, in comune (il 75% dei filtri applicati).

Esse dimostrano che entrambi i dispositivi sono stati utilizzati dallo stesso utente.

Tuttavia questo tipo di confronto è unidirezionale: la ricerca di informazioni caratterizzanti viene effettuata sulla base degli indicatori riscontrati in un solo dispositivo, detto "campione", tralasciando l'analisi e quindi la ricerca di eventuali indicatori sugli altri dispositivi.

Per ovviare al problema è possibile eseguire un'ulteriore operazione di affinamento dei profili attraverso il confronto incrociato, che si basa sull'analisi del contenuto di memoria di tutti i dispositivi.

15 IL CONFRONTO INCROCIATO

Questo tipo di confronto consiste nell'analisi incrociata di tutte le informazioni raccolte di ogni dispositivo in analisi. La sua realizzazione comporta i seguenti passi:

- 1) estrapolazione dei profili utente campione P_{uc} di tutti i dispositivi in analisi, ognuno dei quali sarà composto da:
 - a. unione dei tre insiemi di indicatori rilevati:

$$I(P_s)(D_i) \cup I(P_c)(D_i) \cup I(P_d)(D_i)$$
 - b. unione dei tre insiemi di file:

$$K(P_s)(D_i) \cup K(P_c)(D_i) \cup K(P_d)(D_i)$$
- 2) estrazione degli indicatori (e relativi file) da tutti i profili:
 - a. $I(P_u)(D_i)$
 - b. $K(P_u)(D_i)$
 - c. applicazione di ogni set dei filtri tratti dagli indicatori $I(P_u)(D_i)$ ad ognuno dei dispositivi;
- 3) aggiornamento dei singoli profili ai nuovi indicatori rilevati.

La procedura, pur presentando lo svantaggio di allungare i tempi di realizzazione, può rivelarsi di grande utilità nei caso in cui le informazioni rilevate dall'analisi di un solo dispositivo siano poco consistenti poiché consente di:

- analizzare i dati presenti in tutti i dispositivi;
- incrementare il numero degli indicatori ottenuti;
- rendere più consistenti i profili degli utenti.

Essa inoltre permette di rilevare profili di eventuali ulteriori utenti.

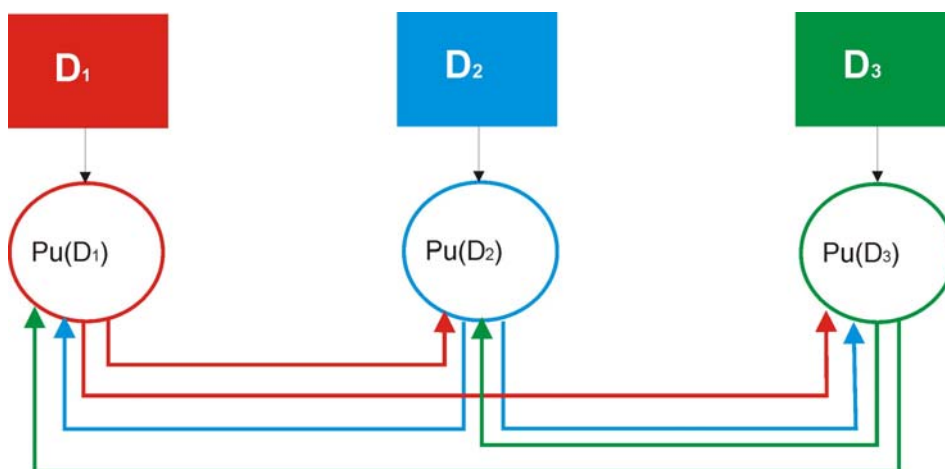


Figura 2 – Confronto incrociato: ognuno dei tre profili rilevati viene confrontato con gli altri due.

16 I DISPOSITIVI MULTI-UTENTE

Un caso più complesso si può verificare nel caso in cui lo stesso dispositivo D_i sia utilizzato da più di un soggetto (es. Personal Computer).

Viene allora estrapolato un profilo per ognuno degli utenti, con la seguente modalità:

- 1) si costruisce un profilo P_c per ogni utente (es. P_{c1} , P_{c2} , ecc.), relativo ad ognuna delle cartelle utenti presenti sulla macchina;
- 2) si costruisce un Profilo di Sistema P_s (es. P_{s1} , P_{s2} , ecc.), per ogni utente;
- 3) si costruisce un unico profilo P_d ;
- 4) si effettua un confronto incrociato tra ogni P_c e il P_d da cui si ottengono tanti Profili dispositivi utente P_u quanti sono i Profili cartella utente P_c ;
- 5) Ogni Profilo utente $P_{u_i}(D_i)$ sarà quindi definito come:

$$P_{u_i}(D_i) = P_{c_i}(D_i) \cup P_d(P_{c_i})(D_i) \cup P_s(P_{c_i})(D_i)$$

Le operazioni di confronto tra i diversi Profili cartella utente P_c e il profilo dispositivo P_d hanno lo scopo di:

- riconoscere i propri indicatori nelle aree di memoria comprese nel Profilo dispositivo;
- estrapolare i file che li contengono ed aggiungerli al proprio **P_{du}**, ove per **P_{du}** si intende il Profilo dispositivo utente, sottoinsieme di **P_d**, costituito da:
 - o l'insieme degli indicatori in comune con il P_c ;
 - o l'insieme dei file che li contengono;
- formare tanti Profili utente **P_{u_i}(D_i)** quante sono le cartelle utenti (non vuote) costituiti da:
 - o $P_{c_i}(D_i) \cup P_{du_i}(P_{c_i})(D_i) \cup P_{s_i}(D_i)$
- diminuire la consistenza del profilo P_d che sarà in ultima istanza composto da quegli indicatori (e relativi file) non compresi nei diversi Profili utente.

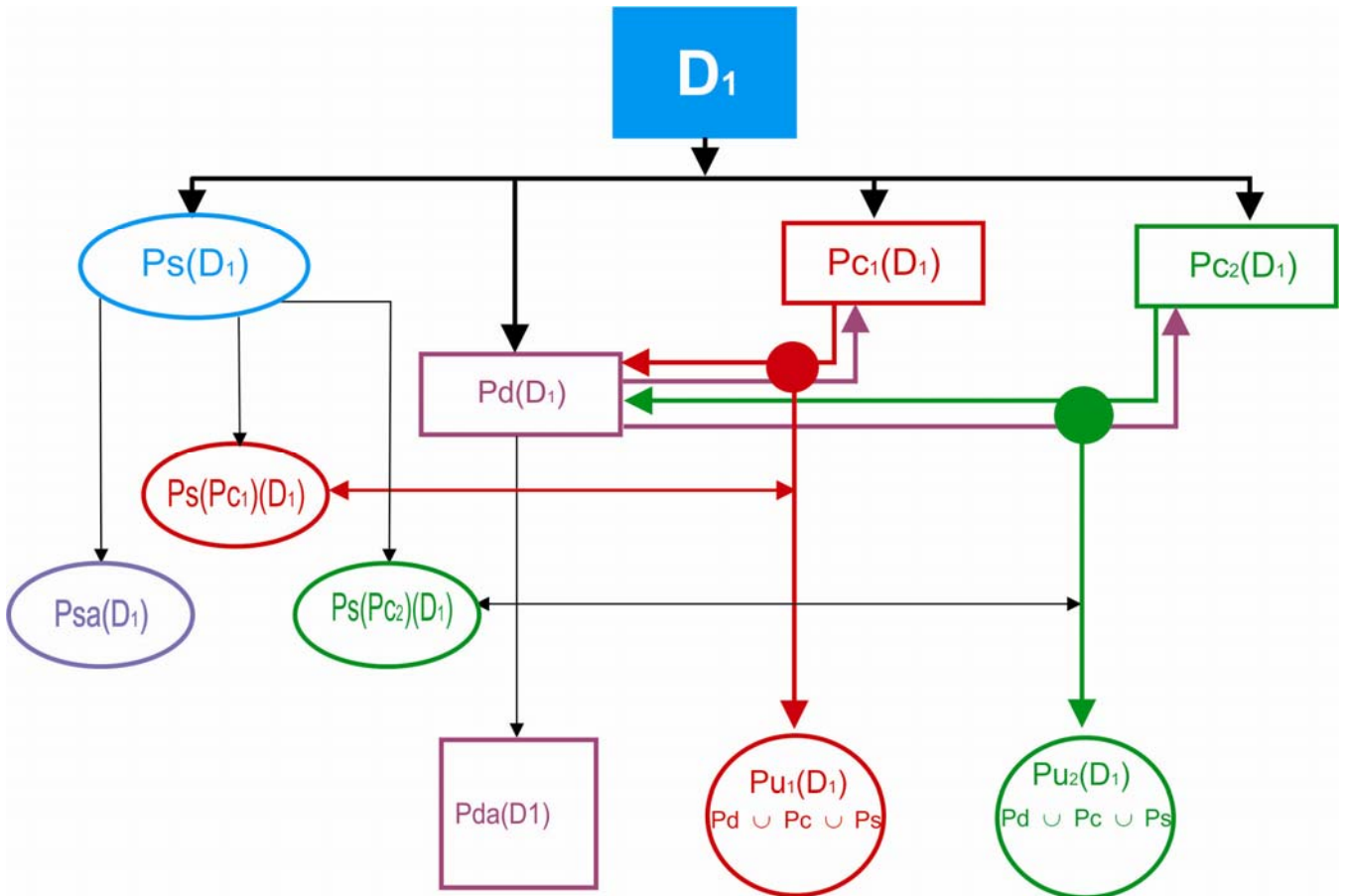


Figura XX – Il processo di rilevazione di n. 2 profili da un dispositivo multiutente.

Il risultato finale consiste in:

- n Profili utente - cioè l'insieme delle informazioni caratterizzanti che descrivono il comportamento digitale degli utenti rilevati sulla macchina;
- n. 1 Profilo dispositivo anonimo Pda (se esiste) – cioè un insieme di informazioni caratterizzanti non associabili ai suddetti utenti, che comprenderà anche quelle informazioni sulla configurazione del sistema non attribuite agli utenti rilevati.

Questo ultimo profilo non viene eliminato, ma viene classificato come Profilo anonimo poiché le informazioni che eventualmente contiene, proprio perché finora non attribuite ad alcuno, potrebbero rivelarsi utili per l'identificazione di altri soggetti, mediante confronto con altri dispositivi in analisi successive.

17 LA VALUTAZIONE DEL RISULTATO

La valutazione del risultato ottenuto (operazione di stretta pertinenza dell'investigatore), si effettua sia in senso quantitativo (cioè valutando il numero degli indicatori coincidenti rilevati), che in senso qualitativo (cioè la veridicità dell'informazione), in quanto anche un'unica informazione rinvenuta può rivelarsi come la soluzione del problema posto dall'obiettivo.

La valutazione quantitativa

Essa si effettua in modo statistico calcolando la percentuale degli indicatori coincidenti rinvenuti tramite il confronto rispetto al totale di quelli utilizzati come filtro.

ESEMPIO:

VALUTAZIONE QUANTITATIVA DEL RISULTATO OTTENUTO TRAMITE CONFRONTO SEMPLICE
(relativa al caso presentato in esempio)

OPERAZIONE 1 - Creazione Profilo utente campione $P_{cu}(D_1)$

N. filtri applicati (tratti dal set di feature minime)	44
RISULTATO: n. indicatori estratti	40

OPERAZIONE 2 - Ricerca degli indicatori tramite filtri sul dispositivo D_2 :

N. filtri applicati (tratti dal set di indicatori)	40
RISULTATO: n. indicatori coincidenti rilevati	30

RISULTATO – su un totale di n. 40 indicatori/filtro applicati tramite confronto semplice, sono stati rilevati n. 30 indicatori coincidenti, cioè il 75%.

La valutazione qualitativa

Essa assegna alle informazioni ottenute (gli indicatori) un valore di “rilevanza” ai singoli indicatori, in relazione al loro grado di utilità per l’obiettivo.

Data la natura digitale dell’analisi, non vengono valutate le *fonti*, in quanto costituite dai file presenti all’interno del dispositivo: se correttamente estrapolati e verificati tramite algoritmi di hashing, si ritengono “*completamente affidabili*”.

Per quanto riguarda le *informazioni* ottenute nel caso qui presentato, non si illustra alcuna valutazione qualitativa (comunque di esclusiva competenza dell’investigatore, nell’ambito dell’indagine), in quanto la specifica ricerca descritta dall’esempio fin qui riportato ha mirato esclusivamente alla rilevazione di dati coincidenti (in possesso cioè di soli due valori : *coincide/non coincide*), che potessero riportare con certezza alla medesima identità del o dei soggetti in questione.

18 CONCLUSIONI

Il Digital Profiling rappresenta un nuovo strumento di indagine informatica mirato all'estrazione di informazioni utili all'identificazione di soggetti, dall'analisi dei dati contenuti nella memoria di un qualsiasi dispositivo digitale.

A questo tipo di analisi si prestano tutti quei dispositivi che, per le numerose applicazioni che offrono e la discreta quantità di memoria che contengono, consentono l'immagazzinamento di una notevole quantità di dati: primi fra tutti i personal computer, seguiti dai telefoni cellulari di tipo Smartphone.

Non vanno dimenticati i dispositivi embedded: un navigatore satellitare, per fare un solo esempio, anche se di primo acchito può sembrare non possa contenere dati attinenti il reato, può invece fornire informazioni preziose sugli spostamenti di un soggetto, come i luoghi in cui si è recato, gli itinerari abituali che, se confrontati con la posizione della sua abitazione, possono aiutare a delineare il raggio d'azione delle sue attività.

Le tecniche di Digital Profiling possono inoltre essere applicate anche al contenuto di aree di archiviazione messe a disposizione in remoto da provider, e a flussi di dati, selezionati per esempio in un determinato intervallo di tempo relativo ad un attacco informatico.

I campi in cui può essere di utilità questo tipo di analisi sono numerosi, e comprendono tutti quei reati che coinvolgono un dispositivo digitale (chi al giorno d'oggi non possiede almeno un telefono cellulare?), in cui sorge la necessità di analisi delle memorie digitali mirata all'identificazione degli autori.

Essa si rivela di particolare utilità nei reati compiuti a mezzo PC, come le frodi informatiche, il cyberstalking, la pedopornografia e l'hacking in genere, specie laddove vengono adottate tecniche di anti-forensics allo scopo di occultare o cancellare le prove del reato commesso.

Ha quindi particolare utilità in operazioni contro la criminalità organizzata, operazioni di anti-terrorismo, operazioni di intelligence, ove può interfacciarsi con lo studio statistico nella previsione e prevenzione degli eventi criminali.

BIBLIOGRAFIA

1. ALFIERO C., *Il ruolo strategico della D.I.A. nel nuovo rapporto tra criminalità organizzata e territorio*, in "Rivista di Polizia", n° 2, 1999.
2. ANCONELLI M., *Introduzione al Digital Profiling*, www.cybercrimes.it.
3. CARRIER B., *File system forensic analysis*, Pearson Education, 2005.
4. CASEY E., *Digital Evidence & Computer Crime*, Second Edition, Elsevier Academic Press, 2004.
5. CENTRO INTERNAZIONALE DI RICERCHE E STUDI MESSINA, *Metodologia di ricerca in tema di alta criminalità*, Giuffrè, Milano 1987.
6. CORNELI A., *Intelligence diffusa e cultura dell'intelligente*, in "Per aspera ad veritatem".
7. COSSIGA F., *Intelligence: istruzioni per l'uso*, in "liMes", n° 3, 1997.
8. D'AMBROSIO L. e VIGNA P.L., *La pratica di polizia giudiziaria*, Cedam, Padova 1996.
9. DI PAOLO A.M., *Elementi di Intelligence e tecniche di analisi investigativa*, Robuffo, 2009.
10. FBI –Federal Bureau of Investigation - *Best practices for forensic image analysis*, USA, 2008.
11. HOSMER C., *Time-lining Computer Evidence*, WetStone Technologies Inc., 1998.
12. IANNIZZO V. A., *Elementi metodologici di intelligente e di analisi induttiva delle informazioni*. *Rivista di intelligence e di cultura professionale*", n° 1, 1995.
13. K. MANDI K., PROSISE C., PEPE M., *Incident response and computer forensics*, McGraw-Hill, 2nd edition, 2005.
14. LOIA V., MATTIUCCI M., SENATORE S., VENIERO, *Computer Crime Investigation by Means of Fuzzy Semantic Maps*, M. Web Intelligence and Intelligent Agent Technologies, 2009. WI-IAT apos;09. IEEE/WIC/ACM International Joint Conferences on Volume 3, Issue , 15-18 Sept. 2009 Page(s):183 - 186 Digital Object Identifier 10.1109/WIIAT. 2009.258.
15. MARCELLA A.J., GREENFIELD R.S., *Cyber Forensics, a field manual for collecting, examining and preserving evidence of computer crimes*, Auerbach, 2007.
16. MATTIUCCI M., DELFINIS G., *Forensic Computing*, Rassegna dell'Arma dei Carabinieri - Roma, agosto 2006.

17. MATTIUCCI M., *Forensic Computing*, Rivista Scientifica dell'Arma dei Carabinieri - Roma - Anno 2005.
18. MC KEMMISH R., *What is Forensic Computing, Trends and Issues in Crime and Criminal Justice* (118), Australian Institute of Criminology, 1999.
19. PETERSON M.B., *Applications in Criminal Analysis*, Greenwood publishing group, Westport (Connecticut) 1994.
20. PICOZZI M., ZAPPALÀ A., *Criminal Profiling: Dall'analisi della scena del delitto al profilo psicologico del criminale*, McGraw-Hill 2002.
21. PONTI G., *Compendio di Criminologia*, Cortina, Milano 1999.
22. RABON D., *Investigative discourse analysis*, Carolina Accademic Press, Durham (North Carolina) 1994.
23. SCHULTZ E.E., SHUMWAY R., *Incident Response: A Strategic Guide to Handling System and Network Security Breaches*, Sams 2002.
24. SHINDER D.L., TITTEL E., *Scene of Cyercrime*, Syngress, 2006.
25. SIDOTI E., *Morale e metodo nell'intelligente*, Cacucci Editore, Bari 1998.
26. STRANO M., E ALTRI... MATTIUCCI M., *Abusi sui minori: manuale investigativo*, Nuovo Studio Tecna, Roma, febbraio 2006 (sponsored by Symantec).
27. STRANO M., E ALTRI... MATTIUCCI M., *Manuale di investigazione criminale - Accertamenti tecnici su cellulari e smartpone*, Nuovo Studio Tecna, Roma, febbraio 2008.
28. STRANO M., *Nuove frontiere delle tecniche di criminal profiling*, ANFP Forze Civili Anno 3 N.1 2005.
29. TURVEY B., *Criminal Profiling: an introduction to behavioural evidence analysis*, Academic Press 1999.
30. TURVEY B., *Deductive Criminal Profiling: Comparing Applied Methodologies Between Inductive and Deductive Criminal Profiling Techniques*, Knowledge Solutions Library, January 1998.
31. UNICRI - AUSTRALIAN INSTITUTE OF CRIMINOLOGY, *Environmental crime, sanctioning strategies and sustainable development*, United Nations Publication, Roma 1993.

Un doveroso ringraziamento:

Magg. Ing. Marco Mattiucci, Comandante del Reparto Tecnologie Informatiche – RTI, RACIS Roma, Arma dei Carabinieri, che mi ha fornito gli spunti che hanno permesso l’avvio e lo sviluppo di questa ricerca;

Ten. Col. Dott. Antonio Colella, Criminologo e Informatico Forense, Stato Maggiore dell’Esercito, Roma, che ha integrato e completato la ricerca nella versione inglese (Colombini – Colella, Digital Profiling: a computer Forensic Approach, Springer ed., 2011, <http://www.springerlink.com/content/p0v637g733757621/>);

App. CC. Dott. Vincenzo Scognamiglio, Digital Forensic Expert - Consulente Tecnico Informatico per l’Autorità Giudiziaria - Responsabile CIT per la Procura della Repubblica di Monza, per i preziosi consigli tecnici;

Dott. Cesare Gallotti, matematico illuminato, paziente e meticoloso “correttore di bozze ” delle formule matematiche.