LE BASI TECNICHE DELLE INVESTIGAZIONI DIGITALI

Università degli Studi di Camerino
3 Maggio 2013
Mattia Epifani
mattia.epifani@digital-forensics.it

Digital Forensics

*La **Digital Forensics** (**Informatica Forense**) è la scienza che studia*

Identificazione

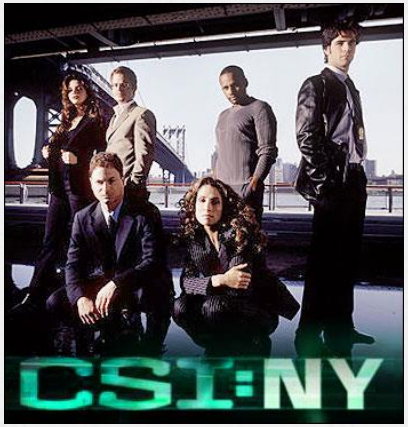
Acquisizione e conservazione

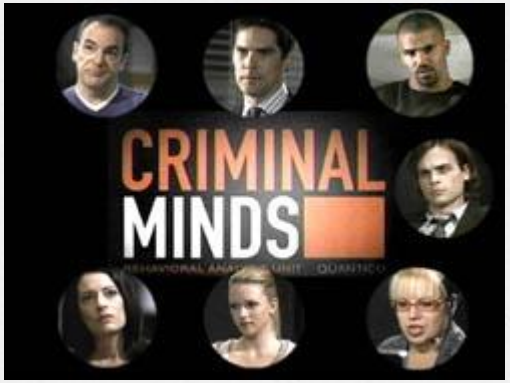
Analisi

Documentazione

*del **dato informatico** per essere valutato in un processo giuridico*

Digital Forensics NON è





Digital Evidence

- Una **digital evidence** può essere definita come **qualsiasi informazione avente valore probatorio che sia memorizzata o trasmessa in forma digitale**

- Una digital evidence può quindi essere estratta da:
  - Un **dispositivo di memorizzazione digitale**
    - Personal computer, notebook, hard disk esterno, NAS, floppy, nastro, CD/DVD, memory card, USB drive,…
    - Telefoni cellulari, SIM, SmartPhone, Tablet, Navigatori satellitari,…
  - Una **Rete Intranet/Internet**
    - Traffico di rete
    - Messaggio di posta elettronica
    - Pagina Web
    - Blog
    - Social Network
    - Chat/IM
    - Cloud Storage

Digital Evidence

- Una digital evidence è **fragile per natura**, ovvero facilmente modificabile
  - Se il dispositivo che contiene le informazioni di interesse **viene spento**, i dati che non sono stati salvati possono andare definitivamente persi
  - Se il dispositivo viene rivenuto spento, **l'accensione comporta modifiche al sistema e/o ai dati in esso contenuti**
  - Se il dispositivo è connesso ad Internet o ad una rete aziendale, **possono avvenire accessi dall'esterno con l'obiettivo di cancellare le informazioni**
  - Se la digital evidence di trova su Internet (sito web, profilo di social network, ecc.), **può essere modificata e/o rimossa dall'owner della pagina**

Digital Evidence

- I dati digitali possono essere divisi in due categorie
- **Dati volatili**: dati che sono **facilmente alterabili/persi in caso di spegnimento del dispositivo che li conserva**
  - Utenti connessi
  - File aperti
  - Software e servizi in esecuzione
  - Contenuto della RAM
  - Applicazioni aperte in uno SmartPhone
  - Contenuto di una chat attiva
- **Dati non volatili**: conservati in memorie di massa e che **non vanno persi in caso di spegnimento del dispositivo che li conserva**
  - File personali dell'utente (documenti, fogli di calcolo, archivi di posta, ecc.)
  - File di configurazione del sistema operativo
  - File dei software applicativi
  - Spazio non allocato
  - Slack space

Passi operativi

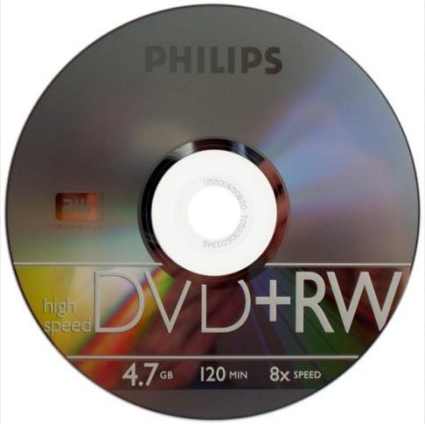
- **Identificazione e repertamento**
  - Computer spento
  - Computer acceso
- Acquisizione e verifica
- Analisi
- Valutazione e presentazione

Identificazione

- La fase di identificazione avviene in corrispondenza dell'**analisi della scena del crimine**
- Il processo di identificazione deve seguire le cosiddette "**best practises**"
- Può sembrare la fase più semplice, perché si tratta **«unicamente» di individuare e catalogare il potenziale contenitore delle informazioni ricercate**
- Tuttavia, vista l'enorme quantità di strumenti atti a conservare dati che si possiedono, è fondamentale individuare tutto quello che può essere utile
- Per esempio….
  - Quanto ci vuole per occultare una scheda microSD?
  - E' sempre facile individuare un Pen Drive?

Hard disk

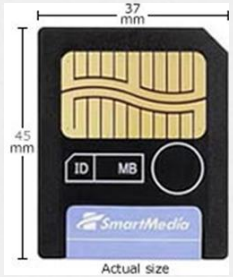
Floppy, ZIP, CD, DVD, BluRay



Memory Card


SanDisk
4.0GB
CompactFlash®


CompactFlash
SanDisk
ultra II
8.0GB
SanDisk


37 mm
45 mm
ID
MB
SmartMedia
Actual size


SanDisk
4.0
MEMORY STICK PRO
MAGICGATE


24 mm
32 mm
pqi
1GB
MultiMediaCard
Actual size


SanDisk
Ultra II
MEMORY STICK PRO DUO
8GB
MAGICGATE


SanDisk
M2
16GB


SanDisk
SD Card
2GB


RS-MMC
SanDisk
2.0GB


SanDisk
SDHC Card
8GB


SD
XC
64GB


1GB
miniSD
SanDisk


4.0GB
miniSD
SanDisk


SanDisk
Premier
2GB
microSD


SanDisk
16GB
microSDHC


FUJIFILM
xD
xD-Picture Card
M2GB


iPod
Music
Videos
Photos
Podcasts
Extras
Settings
Shuffle Songs
Now Playing
MENU


Logitech


Video
PHILIPS


SONY
Cyber-shot
8.1
MEGA PIXELS


Full HD
1080
HD


AIPTEK
AIPTEK
PenCam Trio HD

Smartphone, GPS, Tablet

E se la “digital evidence” e su Internet?

Repertamento

- A seconda della tipologia di dispositivo e/o localizzazione, si possono identificare delle "**best practises**" per il repertamento
- Analizziamo 2 casi:
  - Computer spento (**Post Mortem Forensics**)
  - Computer acceso (**Live Forensics**)

Best Practices Internazionali

- Esistono linee guida dettagliate con le corrette metodologie di acquisizione:
  - RFC3227 - Guidelines for Evidence Collection and Archiving (2002)
  - USA – Department of Justice - Searching and Seizing Computers (2002)
  - USA – IACP - Best Practices for Seizing Electronic Evidence (2006)
  - USA – DoJ – Electronic Crime Scene Investigation v. 2 (2008)
  - UK – ACPO – Computer Based Evidence Guidelines v.4 (2008)
  - **ISO 27037 - Guidelines for identification, collection, acquisition and preservation of digital evidence**

Passi operativi

- **Identificazione e repertamento**
  - **Computer spento**
  - Computer acceso
- Acquisizione e verifica
- Analisi
- Valutazione e presentazione

Repertamento di un computer spento

- **Mettere in sicurezza** la scena
- **Allontanare** le persone presenti dai dispositivi digitali
- **Fotografare** o fare una ripresa video della scena del crimine
- **Assicurarsi** che il computer sia **effettivamente spento**
- **NON ACCENDERE IL COMPUTER PER NESSUN MOTIVO**

Repertamento di un computer spento

- **Rimuovere** la batteria
- **Scollegare** l'alimentazione
- **Etichettare** le porte e i cavi
- Assicurarsi che tutte gli oggetti siano stati **sigillati e siglati**
- Identificare eventuali indicazioni del **modello** e del **numero di serie** presenti
- **Compilare un report** di sequestro per ogni oggetto
- Ricercare diari, appunti o pezzi di carta con password
- **Prendere nota dettagliata di tutte le operazioni compiute** in relazione ai dispositivi informatici

Catena di custodia

- La digital evidence deve essere **trattata e conservata molto attentamente per evitare contaminazioni, danni e qualsiasi azione che potrebbe renderla inutilizzabile**
- Si deve predisporre una **catena di custodia** che identifichi tutte **le persone che hanno avuto accesso al supporto originale**
- La catena di custodia deve contenere alcune informazioni fondamentali, come:
  - Dati identificativi del caso (numero, investigatore, ecc.)
  - Dati identificativi del supporto (produttore, modello, numero di serie)
  - Dati identificati del sequestro (data e ora di inizio custodia, luogo)
- Ogni volta che i supporti oggetto di indagini sono affidati a un nuovo investigatore, nella catena di custodia dovrà essere aggiunta un'informazione contenente:
  - Nome della persona che ha preso in carico il supporto
  - Data e ora di consegna e data e ora di restituzione

Hard Drive/Computer Details

Description:		
Manufacturer:	Model #:	Serial #:

Chain of Custody

Date/Time:	From:	To:	Reason:
Date:	Name/Organization:	Name/Organization:	
Time:	Signature:	Signature:	
Date:	Name/Organization:	Name/Organization:	
Time:	Signature:	Signature:	
Date:	Name/Organization:	Name/Organization:	
Time:	Signature:	Signature:	
Date:	Name/Organization:	Name/Organization:	
Time:	Signature:	Signature:	
Date:	Name/Organization:	Name/Organization:	
Time:	Signature:	Signature:	
Date:	Name/Organization:	Name/Organization:	
Time:	Signature:	Signature:	

Passi operativi

- Identificazione e repertamento
  - Computer spento
  - Computer acceso
- **Acquisizione e verifica**
- Analisi
- Valutazione e presentazione

REGOLA FONDAMENTALE!

PRESERVARE L'ORIGINALE!

Copia Forense

- **L'originale non deve mai essere utilizzato per l'analisi dei dati, per garantirne l'integrità**
- Per effettuare l'acquisizione di tutti i dati presenti sul dispositivo è necessario (ove possibile) effettuare una **copia bit-a-bit** (**bit-stream image**) del supporto originale
- Questa operazione è **differente da un semplice backup** dei dati, che consiste nella copia di **file noti** e tralascia file cancellati, slack space, spazio non allocato, ecc.
- **L'acquisizione viene solitamente effettuata leggendo ogni bit del supporto originale (prevenendo qualsiasi possibile scrittura) e scrivendo un file immagine su un supporto di destinazione**

Copia Forense
Duplicatori e Write Blocker

Copia Forense
DEFT

Copia Forense
CAINE

Verifica dell'integrità della copia

- Come posso verificare la conformità e la successiva integrità della copia?
- **Verifica bit a bit**: Richiede tempi molto lunghi ed e possibile solo disponendo dell'originale
- Usando **funzioni di hash**

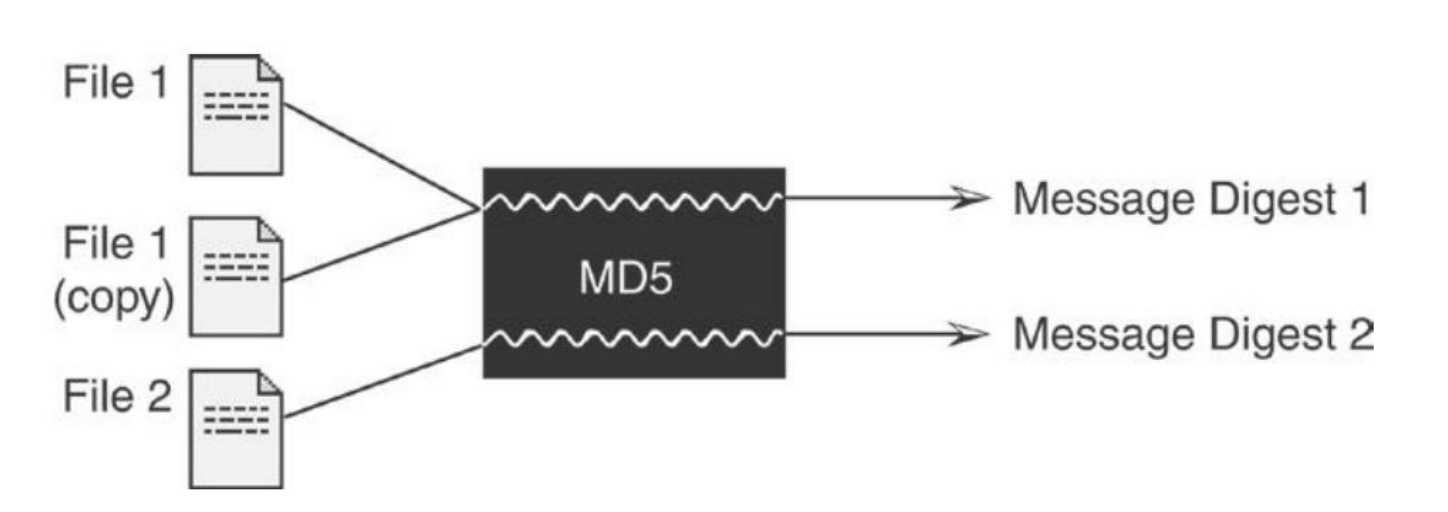
Funzioni di hash

- Una funzione crittografica di hash **trasforma** dei **dati di lunghezza arbitraria** (un sequenza di bit) in una **stringa di dimensione fissa** chiamata **valore di hash** o **cheksum**
- Una funzione di hash deve essere tale che:
  - **Non sia possibile** (o infinitamente oneroso in termini di calcolo) **risalire al dato originario**
  - **Una piccola variazione nel dato originario si traduce in una grande variazione del risultato**
  - E raro che due dati presi a caso diano il medesimo risultato (**collisioni**)

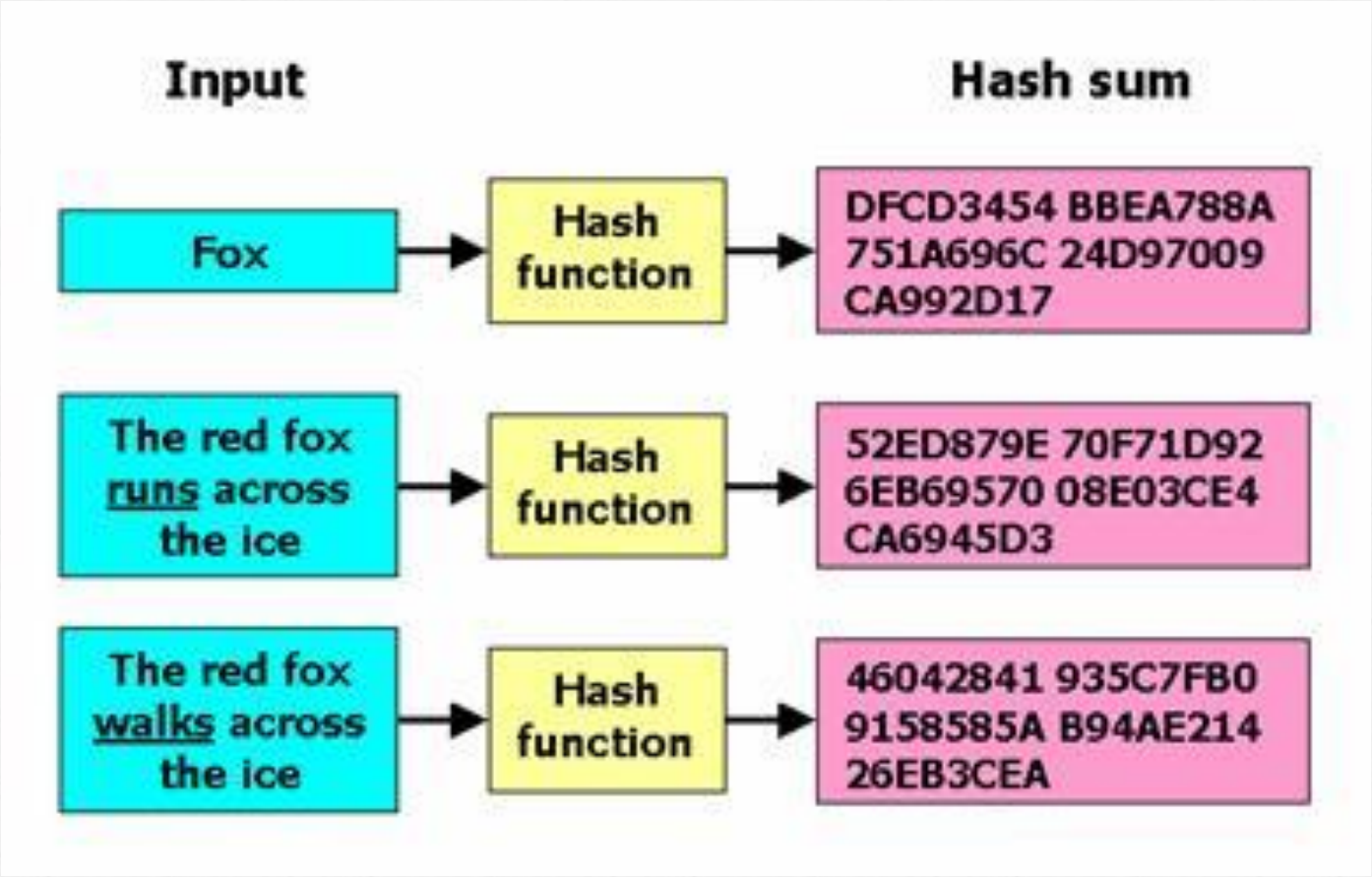
Funzioni di hash

- Le funzioni di hash più utilizzate sono:
  - **MD5 (128 bit)**
  - **SHA-1 (160 bit)**
  - **SHA-256 (256 bit)**
  - **SHA-512 (512 bit)**
- Gli strumenti di acquisizione (hardware o software) **calcolano l'hash del supporto originale e dell'immagine per verificare il processo di copia**

Funzioni di hash



Funzioni di hash



Passi operativi

- **Identificazione e repertamento**
  - Computer spento
  - **Computer acceso**
- Acquisizione e verifica
- Analisi
- Valutazione e presentazione

Live Forensics

- Quando ci si trova davanti a un computer acceso si deve effettuare una scelta:
  - **Spegnerlo subito** per effettuare una copia forense
  - **Esaminarlo** mentre è in esecuzione
- La scelta dipende da diversi fattori
  - Competenza e/o conoscenza dello specifico sistema
  - Strumenti disponibili
  - Rilevanza dei dati rispetto all'indagine

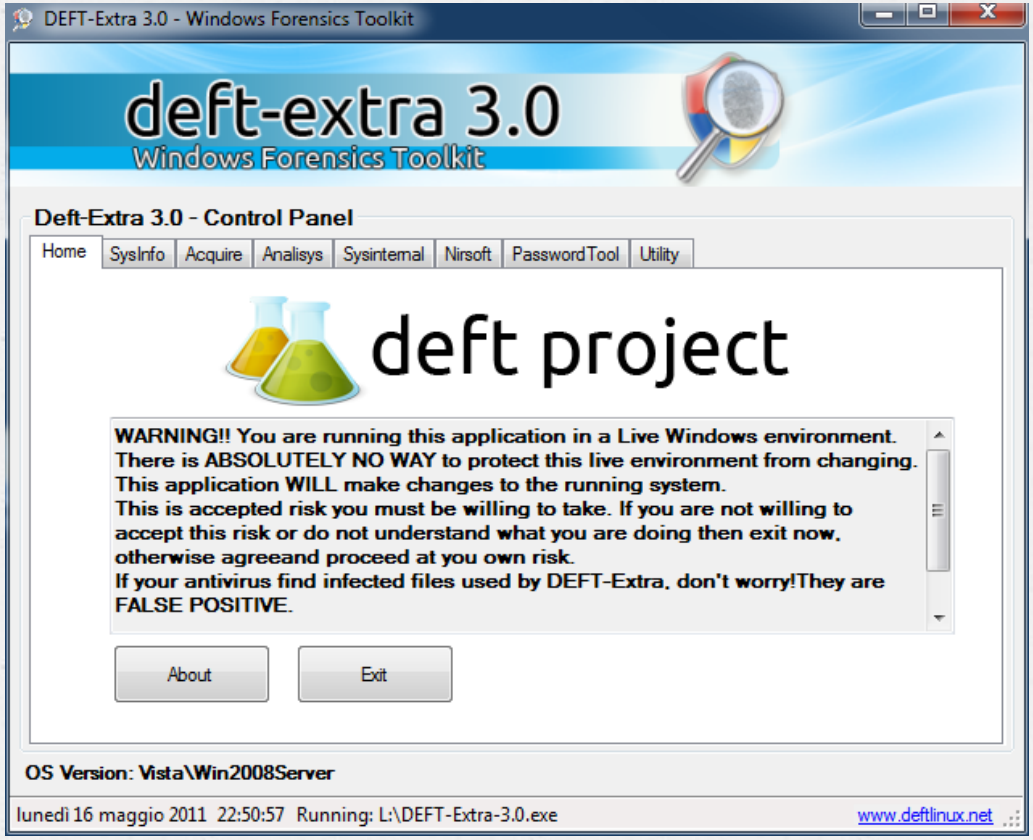
Necessità vs. Invasività

- Un intervento di *live forensics* si rende **necessario** (o molto utile) quando:
  - Il sistema **non è fisicamente rimovibile**
  - Il sistema **non può essere spento**
    - Militari
    - Videosorveglianza
    - Strumenti medicali
    - Database server condivisi
    - Server in hosting/housing
  - Il sistema **non può essere acquisito nella sua interezza**
  - **Le informazioni "volatili" sono rilevanti** rispetto alle indagini (es. stato della rete, chat/download in corso, memorie volatili, ecc.)
  - Siamo in presenza di **volumi cifrati** (BitLocker, FileVault, TrueCrypt, PGP, ecc.)

Necessità vs. Invasività

- Per contro utilizzando tecniche di *live forensics*:
  - Il sistema viene **sicuramente perturbato**:
    - Le modifiche apportate sono note?
    - Le modifiche apportate sono documentabili?
    - Le modifiche apportate intaccano significativamente il risultato dell'analisi?
    - Ogni modifica apportata può distruggere un altro dato
  - **Gli accertamenti svolti su sistemi accesi non saranno ripetibili**

DEFT Extra (ora DART)



Passi operativi

- Identificazione e repertamento
  - Computer spento
  - Computer acceso
- Acquisizione e verifica
- **Analisi**
- Valutazione e presentazione

Analisi

- Le modalità di analisi variano a seconda del **tipo di caso** investigato
- Questa fase comprende (elenco non esaustivo…):
  - Identificazione della **struttura logica del supporto**
  - Recupero di **file e informazioni cancellate**
  - Analisi del **contenuto dei file**
    - Documenti
    - Immagini
    - Video
    - Audio
  - Estrazione delle informazioni sul **sistema operativo**
  - Analisi dei principali **software applicativi**
    - Navigazione su Internet
    - Posta Elettronica
    - Chat
    - Social Network
    - Webmail
  - Generazione della **timeline** di utilizzo del computer
  - Ricerca per **keywords**

Pedopornografia

- Utilizzo di sistemi di **file sharing**
  - File scaricati
  - Parole chiave utilizzate
  - Condivisione dei file
- **Navigazione su Internet**
  - Siti acceduti
  - Parole chiave ricercate
- **Posta elettronica**
- Accesso a **Social Network** (Facebook, Badoo, Netlog)
- Ricerca per **keywords**
- Utilizzo di sistemi di **cloud storage** (Dropbox, Google Drive, SkyDrive)
- Utilizzo di **visualizzatori di immagini** e **player video**
- Utilizzo di **software di masterizzazione**

Stalking

- **Posta elettronica**
- **Registro chiamate su cellulare e/o centralino VoIP**
- **SMS/MMS**
- **App messaggistica** (Whatsapp, Viber, BlackBerry Mess)
- **Chat** (Skype, Messenger, ICQ)
- **Social Network** (Facebook, Linkedin, Badoo)
- Analisi dei **documenti**
- Ricerca per **keyword**

Spionaggio Industriale

- Utilizzo di **periferiche USB**
- Utilizzo di **software di masterizzazione**
- Utilizzo di sistemi di **cloud storage** o **server FTP**
- Utilizzo delle **stampanti**
- Analisi della presenza di **malware** (anche su **mobile**!)
- **Posta elettronica** e accessi a sistemi di **WebMail**
- **Timeline degli eventi**
- Utilizzo di **sistemi di cifratura** (file ZIP cifrati, TrueCrypt)

Furto d'identità

- Analisi della presenza di **malware**
- Attività su **Social Network**
- **Password memorizzate** sui dispositivi
- Registro chiamate
- **Navigazione su Internet** (utilizzo di carte di credito online)

Accesso abusivo

- Log di **accesso al computer in locale** (registro eventi, log sul server, ecc.)
- **Log di accesso al computer da remoto** (VPN, Software di gestione remota es. TeamViewer o LogMeIn)
- Log dei **dispositivi di rete** (Router, Switch)
- Log dei **dispositivi di sicurezza** (Firewall, IDS/IPS)
- Configurazioni e log di accesso dei **dispositivi WiFi**
- Analisi della presenza di **malware** (anche su **mobile**!)

Diffamazione su Internet

- Acquisizione di **pagine Web**
- Analisi delle informazioni disponibili sul **blog/sito**
  - **Nome di dominio**
  - **Contatti telefonici e email**
  - **Indirizzo IP del server**
  - **Informative per la rimozione dei messaggi**
- Attività su **Social Network**

Strumenti di analisi

- Forte contrapposizione

opensource vs proprietario

- Come in ogni altro settore dell'informatica il compromesso sta nel mezzo, ovvero utilizzare opensource finché si può, ma anche tool proprietari per non rovinarsi la vita!

Whatsapp Xtract

- Fabio Sangiacomo (fabio.sangiacomo@digital-forensics.it) ha realizzato un tool per il parsing delle chat di Whattsapp su dispositivi Android e iOS

- Il tool è liberamente scaricabile dal sito
  http://blog.digital-forensics.it/2012/05/whatsapp-forensics.html

- Nei dispositivi iOS Il database delle chat viene automaticamente inserito da iTunes all'interno di un backup
  Net.whatsapp.WhatsApp/Documents/ChatStorage.sqlite

- All'interno del database sono presenti 4 tabelle di interesse
  - ZWASTATUS contatti
  - ZWACHATSESSION sessioni di chat
  - ZWAMESSAGE messaggi
  - ZWAMEDIAITEM contenuti multimediali allegati

Come si fa…

- Scaricare il tool dal sito blog.digital-forensics.it
- Estrarre il file compresso all'interno di una cartella del PC (es: C:\WhatsApp)
- Installare un interprete Python sul computer (es. ActivePython)
- Collegare iPhone al computer
- Effettuare un backup non cifrato utilizzanto iTunes
- Individuare nella cartella di backup il file 1b6b187a1b60b9ae8b720c79e2c67f472bab09c0
- Copiare il file nella cartella dove è stato decompresso il tool
- Rinominare il file come ChatStorage.sqlite
- Eseguire whatsapp_xtract_iphone.bat
- Attendere l'esecuzione e l'apertura del report automatico in HTML

- Stiamo lavorando per semplificare la procedura, eliminando la necessità di estrarre e copiare manualmente il file, ovvero recuperandolo direttamente dalla cartella di backup di iTunes

Standardizzazione

- Il NIST ha creato il **CFTT (Computer Forensics Tool Testing)** per la validazione degli strumenti (hardware e software) di computer forensic

http://www.cftt.nist.gov

Passi operativi

- Identificazione e repertamento
  - Computer spento
  - Computer acceso
- Acquisizione e verifica
- Analisi
- **Valutazione e presentazione**

Valutazione e presentazione

- I risultati e le conclusioni dedotte devono essere presentate in **forma facilmente comprensibile**
- I destinatari (giudici, avvocati, amministratori di aziende) non hanno sempre competenze informatiche approfondite
- Tuttavia è probabile che la relazione venga esaminata da un tecnico della controparte
- **Semplicità e chiarezza, non superficialità e approssimazione**

Materiali di riferimento

- “Investigazione penale e tecnologia informatica” – Giuffrè – 2007
- “Computer Forensics”– Apogeo – 2009
- “Sistema penale e criminalità informatica” – Giuffrè – 2009
- “Computer Forensics e Indagini Digitali” – Experta - 2011

Riferimenti online

- http://www.digital-forensics.it
- http://www.iisfa.it/
- http://www.cftt.nist.gov/
- http://www.marcomattiucci.it/
- http://www.ictlex.net/
- http://www.computerforensics.unimi.it/
- http://www.deftlinux.net/
- http://www.caine-live.net/
- http://www.e-evidence.info/
- http://www.forensicswiki.org/
- http://www.forensicfocus.com/
- http://www.opensourceforensics.org/
- http://www.nirsoft.net
- http://www.tzworks.net
- http://redwolfcomputerforensics.com
- http://www.mitec.cz
- http://www.woanware.co.uk
- http://www.sysinternals.com
- http://www.passwordforensics.com

Grazie per l'attenzione!

Mattia Epifani

Mail: **mattia.epifani@digital-forensics.it**
Web: **http://www.digital-forensics.it - http://blog.digital-forensics.it**
Linkedin: **http://www.linkedin.com/in/mattiaepifani**