

Web Forensics

LE NUOVE FRONTIERE DELLE PROVE DIGITALI



Paolo Dal Checco

- Laurea e Ph.D. in Informatica, Università di Torino
- Consulente Informatico Forense (10+ anni, 2k+ casi)
- CTP, CTU, Esperto, Perito del Giudice, CT del PM, Ausiliario di PG
- Collaborazioni con UniTO (Docente a Contratto corso Sicurezza Informatica @SUISS), UniGE (Master), UniMI e PoliMI (Master e Corsi di Perfezionamento)
- Interessi in mobile forensics, OSINT, cryptocurrency forensics, web forensics... in sostanza tutti gli aspetti della digital forensics

Andrea Lazzarotto

- Consulente informatico forense e sviluppatore
- Gli interessi includono la WhatsApp forensics e anti-forensics (manipolazione delle chat) e l'acquisizione e analisi dei metadati dei profili Instagram
- Autore di alcuni strumenti open source, inclusi **RecuperaBit** per la ricostruzione di NTFS e **Carbon14** per datare le pagine web (entrambi si trovano in CAINE)
- **Autore di Fuji**, il nuovo software open source per l'acquisizione forense dei computer con macOS



Agenda

COSTRUIRE UNA MACCHINA VIRTUALE

Ambiente per l'acquisizione di pagine web

SOFTWARE E TECNICHE AVANZATE

Programmi specifici e registrazione di API

CRISTALLIZZARE LA MACCHINA

Acquisizione forense dell'ambiente utilizzato

Costruire una macchina virtuale

AMBIENTE PER L'ACQUISIZIONE DI PAGINE WEB

Web Forensics

Proviamo a dare una definizione di “web forensics”:

Branca specializzata della digital forensics che si occupa dell'identificazione, raccolta, analisi, preservazione e presentazione di prove digitali ottenute da applicazioni web, siti web, cloud, servizi online e altre fonti accessibili tramite Internet

Abbiamo quindi due macro-categorie nella “web forensics”:

- **Acquisizione** forense
- **Analisi** forense



Cristallizzazione di risorse web

Ci occuperemo della prima categoria, acquisizione forense, che segue i primi due dei 3 principi ricavati dalla Legge 48/2008:

- **Non alterare l'originale:** facile, ma si può correre il rischio di lasciare tracce (es. click su «mi piace» o visita profilo LinkedIn) o condizionare l'acquisizione (indirizzo IP di provenienza, browser sbagliato, lingua del PC, orario di visita, etc...);
- **Copia identica all'originale:** difficile, troppi parametri variabili, necessario delineare perimetro (questioni legate a DNS, DNSCrypt, DNSSec, IP di provenienza, metadati webserver, browser, SSL, HTTP, HTML, traffico, video, audio, HTML5/AJAX, etc...)
- **Copia non modificabile e databile nel tempo:** facile, una volta salvata la copia, hash e marca temporale, volendo anche firma, doppia copia, verbale.



Cristallizzazione di risorse web

Concetti di base della web forensics:

- Rendere le acquisizioni web valide e non disconoscibili tanto quanto (!) quelle tradizionali
- Normative/guide di riferimento: ISO/IEC 27037 (Catena di Custodia), Legge 48/2008, Codice dell'Amministrazione Digitale (D.lgs. n° 82/2005) – Art. 20, ACPO Guide, SWGDE (Best Practices for Acquiring Online Content), Electronic Evidence Guide, Council of Europe, etc...

La metodologia di base

Documentare l'intera attività:

- Flusso Video
- Traffico di Rete (ricordare le chiavi SSL)
- Tracciatura dei processi, log
- DNS, Traceroute, certificati SSL, robots, sitemap, NTP, etc...
- Riferimento Temporale (inizio, durante, termine, con riferimenti oggettivi, anche blockchain)

Infine raccogliere tutto, firmare, applicare timestamp (Blockchain, CA, etc...)

Creazione di una VM

Scaricare o installare una Linux VM: una VM Windows (es. quelle “free” distribuite da MS) potrebbe dare problemi di diritti...

Installare i tool mancanti:

```
apt-get install python3-pip  
google-chrome-stable  
python3-opentimestamps ffmpeg  
  
pip3 install opentimestamps-client
```



CREAZIONE DI UNA VM

Eseguire in una prima finestra:

```
script w1.txt
```

```
export SSLKEYLOGFILE="/home/webforensics/test/sslkeylogfile.log"
```

```
strace -r -f Firefox 2>&1 | tee strace.txt
```

```
openssl s_client -connect www.msab.com:443 -servername www.msab.com  
    < /dev/null 2>/dev/null | openssl x509 -text -noout | tee cert.txt
```

(oppure salvare da Firefox)

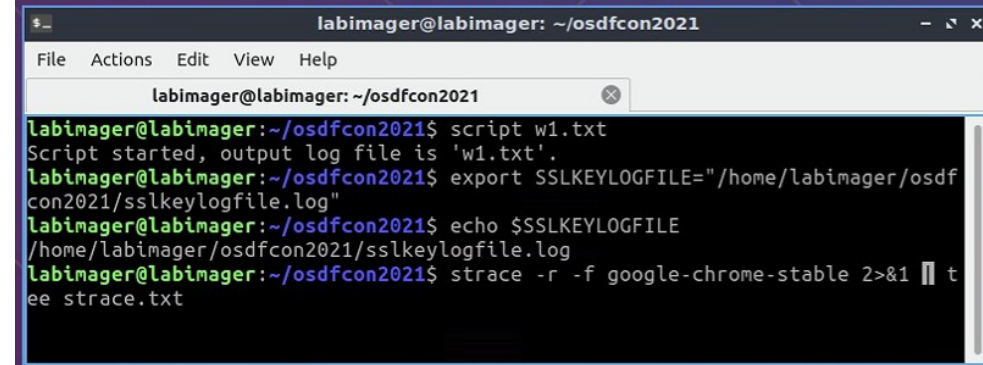
Creazione di una VM

Eeguire in una seconda finestra:

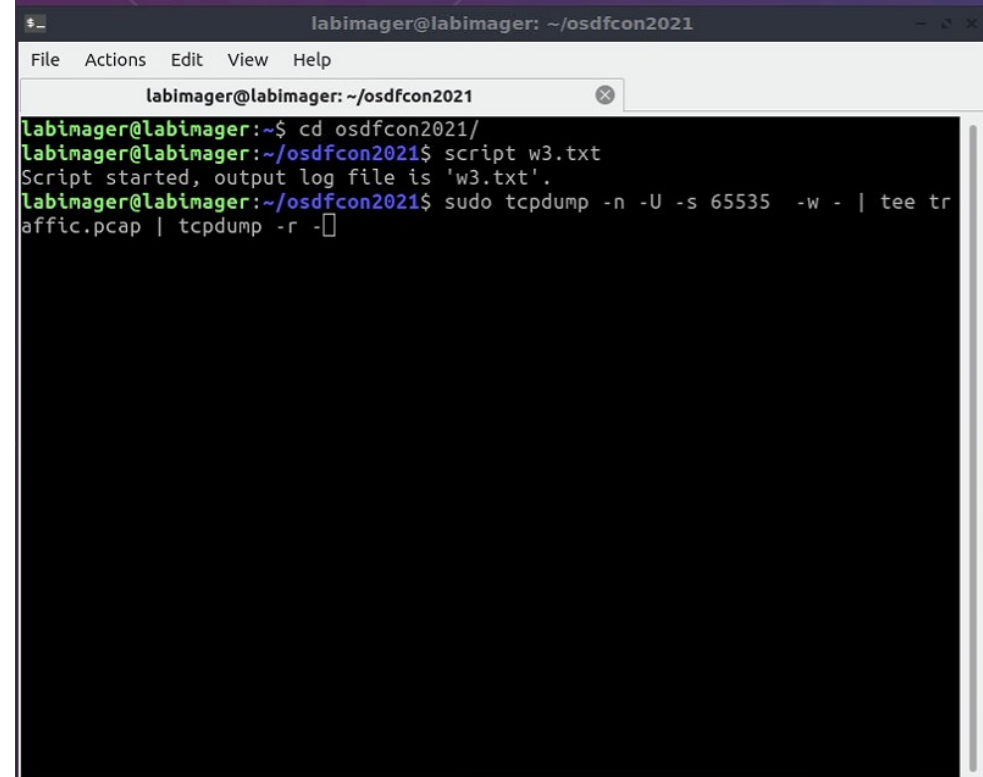
```
script w2.txt
```

```
ffmpeg -f x11grab -y -r 5  
-s 1440x900 -i :0.0  
-c:v libx264 video.avi
```

Si può anche usare OBS, più comodo per registrare l'audio

A terminal window titled 'labimager@labimager: ~/osdfcon2021' with a menu bar (File, Actions, Edit, View, Help). The terminal shows the following commands and output:

```
labimager@labimager:~/osdfcon2021$ script w1.txt  
Script started, output log file is 'w1.txt'.  
labimager@labimager:~/osdfcon2021$ export SSLKEYLOGFILE="/home/labimager/osdfcon2021/sslkeylogfile.log"  
labimager@labimager:~/osdfcon2021$ echo $SSLKEYLOGFILE  
/home/labimager/osdfcon2021/sslkeylogfile.log  
labimager@labimager:~/osdfcon2021$ strace -r -f google-chrome-stable 2>&1 | tee strace.txt
```

A terminal window titled 'labimager@labimager: ~/osdfcon2021' with a menu bar (File, Actions, Edit, View, Help). The terminal shows the following commands and output:

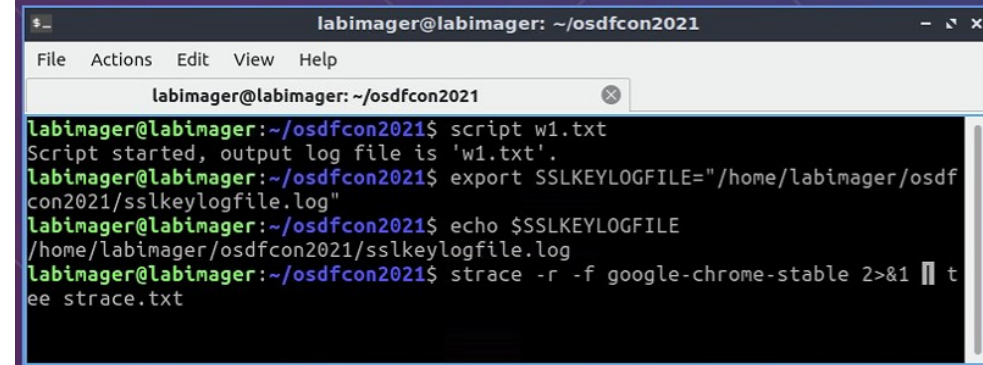
```
labimager@labimager:~$ cd osdfcon2021/  
labimager@labimager:~/osdfcon2021$ script w3.txt  
Script started, output log file is 'w3.txt'.  
labimager@labimager:~/osdfcon2021$ sudo tcpdump -n -U -s 65535 -w - | tee traffic.pcap | tcpdump -r -
```

Creazione di una VM

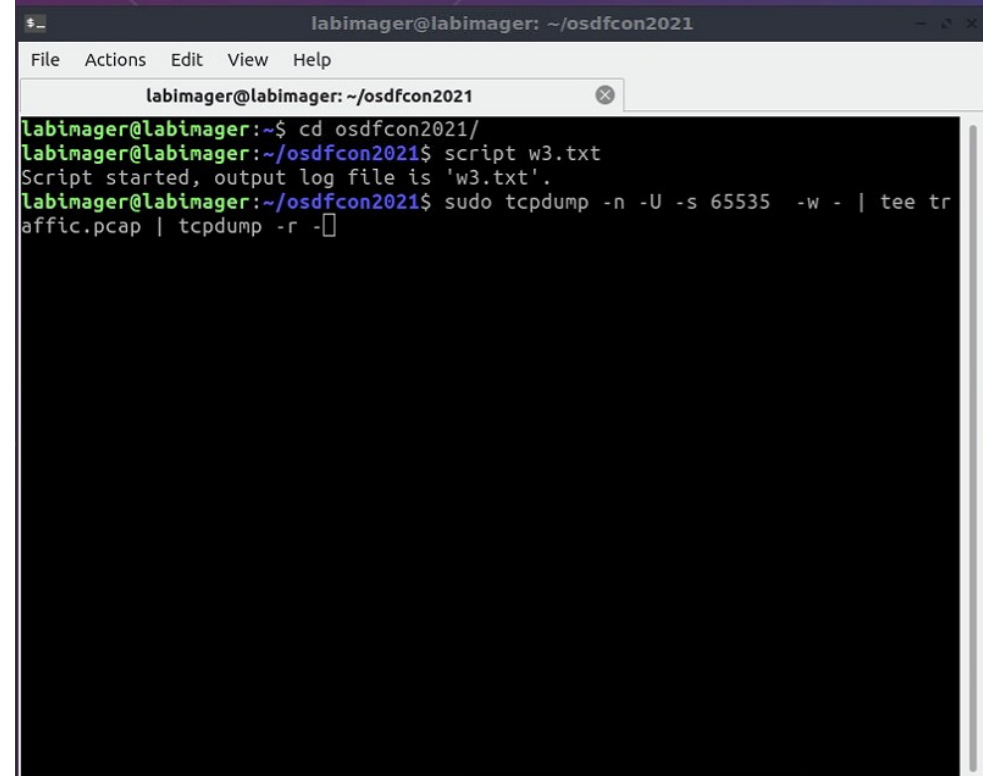
Eseguire in una terza finestra:

```
script w3.txt
```

```
sudo tcpdump -n -U -s 65535 -w -  
| tee tcpdump.pcap  
| tcpdump -r -
```

A terminal window titled 'labimager@labimager: ~/osdfcon2021' with a menu bar (File, Actions, Edit, View, Help). The terminal shows the following commands and output:

```
labimager@labimager:~/osdfcon2021$ script w1.txt  
Script started, output log file is 'w1.txt'.  
labimager@labimager:~/osdfcon2021$ export SSLKEYLOGFILE="/home/labimager/osdfcon2021/sslkeylogfile.log"  
labimager@labimager:~/osdfcon2021$ echo $SSLKEYLOGFILE  
/home/labimager/osdfcon2021/sslkeylogfile.log  
labimager@labimager:~/osdfcon2021$ strace -r -f google-chrome-stable 2>&1 | tee strace.txt
```

A terminal window titled 'labimager@labimager: ~/osdfcon2021' with a menu bar (File, Actions, Edit, View, Help). The terminal shows the following commands and output:

```
labimager@labimager:~$ cd osdfcon2021/  
labimager@labimager:~/osdfcon2021$ script w3.txt  
Script started, output log file is 'w3.txt'.  
labimager@labimager:~/osdfcon2021$ sudo tcpdump -n -U -s 65535 -w - | tee traffic.pcap | tcpdump -r -
```

Creazione di una VM

Eseguire in una quarta finestra:

```
sudo ntpdate 1.ro.pool.ntp.org | tee ntpdate.txt
```

```
cp /etc/hosts ./
```

```
traceroute www.msab.com | dig traceroute.txt
```

```
dig +trace www.msab.com | tee dig.txt
```

```
whois msab.com | tee whois.txt
```

IMPORTANZA DELLE CHIAVI SSL

TeamSpeak2
TELNET
Teredo
TETRA
TFP
TFTP
Thread
Thrift
Tibia
TIME
TIPC
TiVoConnect
TLS
TNS
Token-Ring
TCP
TPKT
TPM2.0
TPNCP
TRANSUM
TSDNS
TSP

Transport Layer Security

RSA keys list

TLS debug file

☒ Reassemble TLS records spanning multiple TCP segments
☒ Reassemble TLS Application Data spanning multiple TLS records
☐ Message Authentication Code (MAC), ignore "mac failed"

Pre-Shared-Key

(Pre)-Master-Secret log filename



Perma.cc ∞

archive.today
webpage capture

Incrementare la “forensicità”

- Volendo si può registrare anche dump di rete e video della VM dall'esterno (quindi dell'host)
- Durante la registrazione, fare cristallizzazioni saltuarie con servizi esterni (Web Archive, Perma.cc, Archive.is, etc...)
- Esportare le pagine strategiche anche in formato HAR, WEBP, Warc direttamente dal browser
- Filtrare ulteriormente tramite mitmproxy per mostrare i comandi e le risposte HTTP

Software e tecniche avanzate

PROGRAMMI SPECIFICI E REGISTRAZIONE DI API

STRUMENTI DI ACQUISIZIONE REMOTI



Minore controllo

Non richiedono configurazione ma c'è meno flessibilità, spesso non sono interattivi



Maggiore autorevolezza

Risulta difficile contestarne la credibilità accusandoli di non essere "terzi"

Alcuni esempi

SITI DI ARCHIVIAZIONE

Nascono essenzialmente per supportare il nuovo lavoro archivistico digitale, naturale evoluzione della preservazione della conoscenza attuata dai bibliotecari.

I principali sono **Archive.org** e **Archive Today**.

ALTRI SERVIZI SPECIFICI

Esistono alcune soluzioni remote specifiche per l'acquisizione forense interattiva, per esempio Kopjra, LegalEye o Eviquire.

PROGRAMMI IN LOCALE



Controllo totale

Generalmente sono flessibili, consentono di usare proxy e accedere a pagine nella LAN



Possibili contestazioni

Qualcuno potrebbe obiettare che l'acquisizione sia potenzialmente modificabile in locale



FIT: Freezing Internet Tool

Il progetto è stato creato da Fabio Zito, come tesi di master.

Ciò è sfociato in un progetto **open-source**, **multi-piattaforma**, **modulare ed estendibile** per l'acquisizione di contenuti web, chiamato "Freezing Internet Tool".

FIT è **sviluppato da informatici forensi per informatici forensi**, quindi perfettamente in linea con le esigenze della professione.

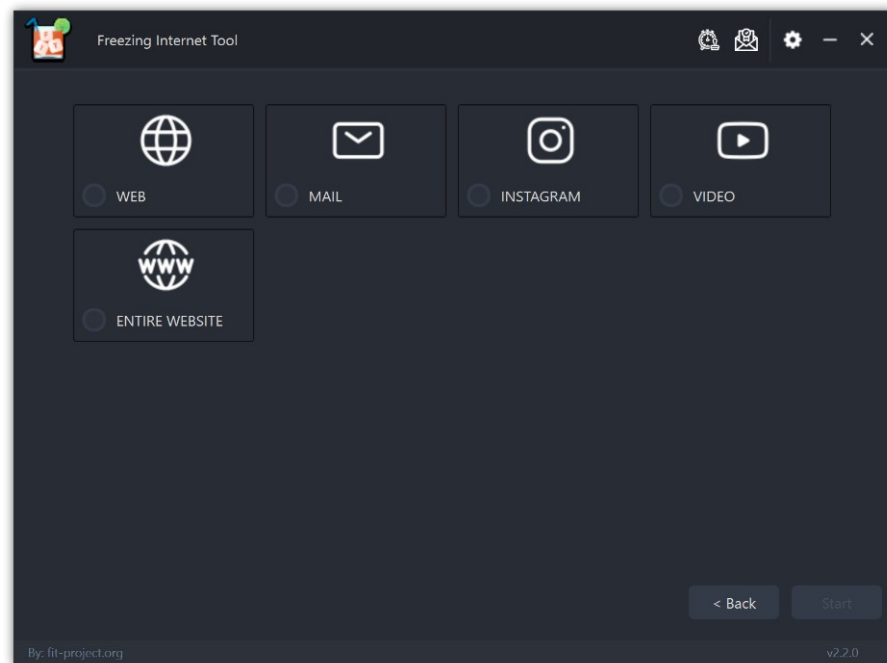
[HTTPS://GITHUB.COM/FIT-PROJECT/FIT](https://github.com/FIT-PROJECT/FIT)

Approccio modulare

Il software è dotato di vari moduli e può essere facilmente ampliato aggiungendone degli altri.

Non tutti i contenuti si acquisiscono usando lo stesso metodo, quindi FIT fornisce vari strumenti:

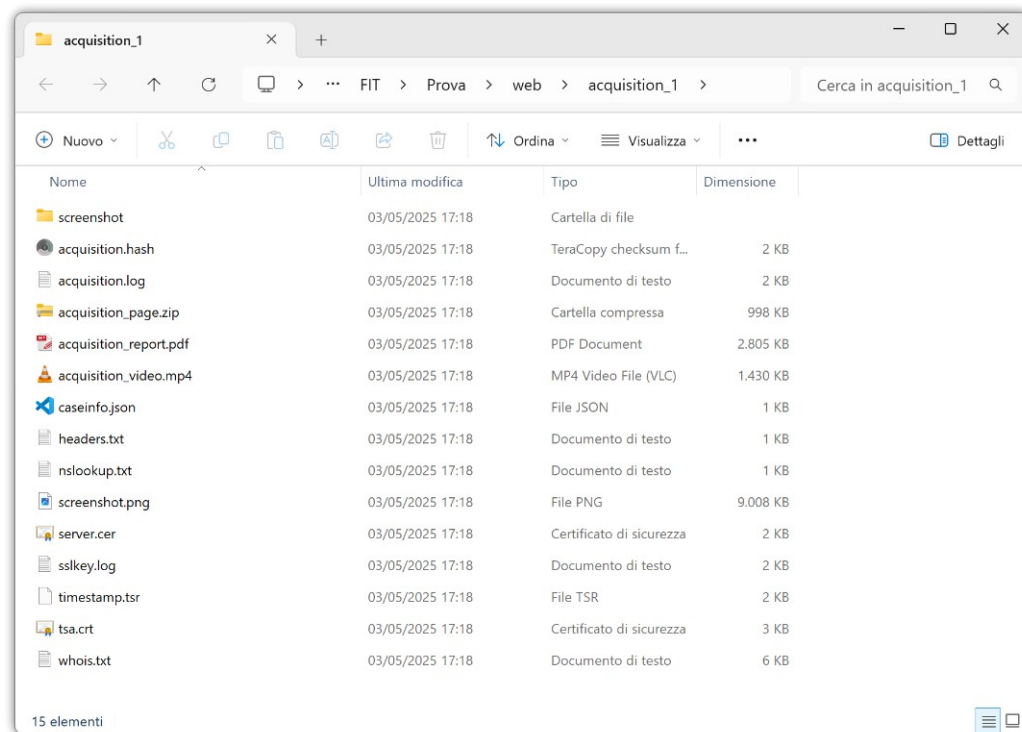
- Web
- Email
- Instagram
- Video



ACQUISIZIONE WEB



RISULTATI PRODOTTI



Il software genera una cartella con tutti i file:

- Registrazione video
- Traffico di rete
- Codice HTML
- Report in PDF
- Hash
- ...

A green padlock is attached to a blue door. The door has water droplets on it, suggesting it might be raining or has been wet. The padlock is a standard combination padlock with a silver-colored shackle. The door is a vibrant blue color.

Misure anti-manomissione

FIT utilizza diverse tecniche per garantire l'integrità dell'acquisizione e rendere possibile successive verifiche:

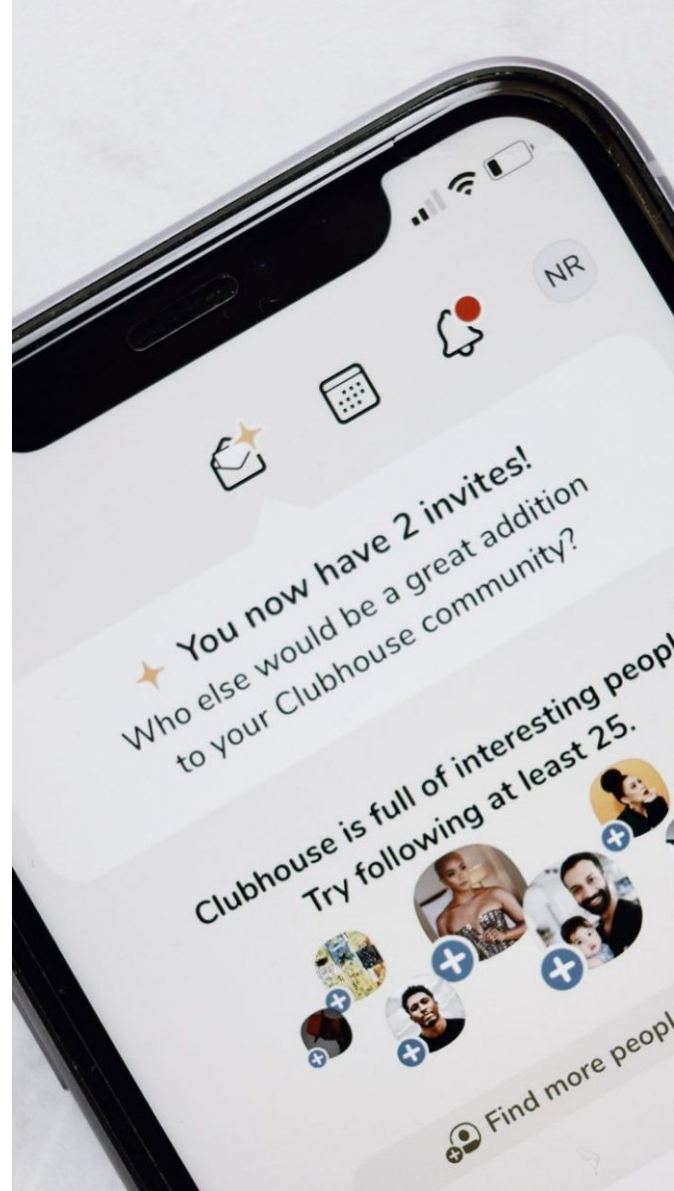
- Registra il traffico di rete e le chiavi dei certificati SSL/TLS
- Calcola l'hash di tutti gli elementi con tre algoritmi
- Usa un servizio esterno di marcatura temporale
- Permette l'invio tramite PEC
- Include strumenti interni per verificare un'acquisizione

Contenuti più “complessi”

In alcuni casi potremmo trovarci di fronte alla necessità di dover acquisire dei contenuti complessi, come i video pubblicati sulle piattaforme di streaming.

Alcuni contenuti distribuiti tramite tecnologie web di fatto **non sono neppure fruibili del tutto tramite browser**, ma solo con modalità specifiche.

Si pensi alle applicazioni mobili sviluppate con tecnologie “ibride” o quelle che fanno uso di API di tipo REST. In questi casi potrebbe essere necessario “intercettare” uno smartphone.





L'essenziale è invisibile agli occhi.

LA VOLPE AL PICCOLO PRINCIPE





Mitmproxy

È un programma che permette di registrare tutto il traffico HTTP, anche quello cifrato (tramite un certificato TLS).

Possiamo usarlo per leggere e acquisire i flussi di comunicazione dei browser ma anche di **molte app per dispositivi mobili**.

L'interfaccia web si può avviare col comando:

```
mitmweb --set http2=false --set http3=false
```

REGISTRAZIONE DEI FLUSSI

The screenshot displays the Wireshark interface with a network traffic capture. The 'Flow List' pane on the left shows a list of HTTP requests. The 'Flow' pane on the right shows the details of the selected request.

Path	Method	Status	Size	Time
https://prod-api.hingeaws.net/metric/v2	POST	200	3.6kb	158ms
https://prod-api.hingeaws.net/standouts/v2	GET	200	17b	142ms
https://prod-api.hingeaws.net/user/export/status	GET	200	68b	155ms
https://prod-api.hingeaws.net/consent/user/blocking	GET	200	2.7kb	134ms
https://prod-api.hingeaws.net/config/v3?country=IT	GET	200	448b	280ms
https://prod-api.hingeaws.net/constable/recordcheck	POST	200	15b	138ms
https://prod-api.hingeaws.net/store/boost	GET	200	106b	138ms
https://prod-api.hingeaws.net/store/v4/product	POST	200	2.6kb	289ms
https://prod-api.hingeaws.net/boost	GET	200	2b	128ms
https://prod-api.hingeaws.net/connection/v2	GET	200	18b	147ms
https://prod-api.hingeaws.net/inappnotifications/v1/views	GET	200	189b	140ms
https://prod-api.hingeaws.net/like/v2	GET	200	255b	137ms
https://prod-api.hingeaws.net/user/v2/public?ids=	GET	200	2b	130ms
https://prod-api.hingeaws.net/content/v1/public?ids=	GET	200	4b	136ms
https://api-3cd...sendbird.com/...	GET	200	44b	270ms
https://media.hingenexus.com/video/upload/dd2...	GET	206	2b	36ms
https://media.hingenexus.com/video/upload/dd2...	GET	206	67.1kb	70ms
https://outlook.office365.com/Microsoft-Server-ActiveSync?User=a.la...	POST	...	141b	...
https://events.split.io/api/testImpressions/bulk	POST	200	313b	290ms
https://prod-api.hingeaws.net/metric/v2	POST	202	6.8kb	127ms

The detailed view of the selected GET request shows the following information:

- Request: GET https://media.hingenexus.com/video/upload/dd2... .aac HTTP/1.1
- Host: media.hingenexus.com
- X-Playback-Session-Id: DD483612
- Range: bytes=0-1
- Accept: */*
- User-Agent: AppleCoreMedia/1.0.0.22E252 (iPhone; U; CPU OS 18_4_1 like Mac OS X; it_it)
- Accept-Language: it-IT,it;q=0.9
- Accept-Encoding: identity
- Connection: keep-alive
- No content

Questo test mostra le chiamate HTTP effettuate dall'app Hinge per iOS.

È possibile notare, in chiaro, l'indirizzo di un file audio aggiunto da una utente al proprio profilo pubblico.

Cristallizzare la macchina

ACQUISIZIONE FORENSE DELL'AMBIENTE UTILIZZATO

Una copia “immodificabile”

- Chiudere il browser, interrompere tcpdump e ffmpeg/obs
- Salvare la history tramite “script” con CTRL+D
- Archiviare la cartella: `tar -czvf acquisition.tar.gz`
- Applicare marca temporale:
 - `ots stamp acquisition.tar.gz`
 - Tool di timestamp
 - Invio hash via PEC
- Eventualmente copiare all'esterno l'archivio compresso





Una copia “immodificabile”

- Chiudere la VM
- Comprimere l'intero folder della VM
- Applicare marca temporale:
 - `ots stamp acquisition.tar.gz`
 - Tool di timestamp
 - Invio hash via PEC
- Applicare firma digitale

Nel caso dei Mac...

Apple ha introdotto la crittografia hardware con il chip T2 nel 2017 e l'ha perfezionata con Apple Silicon alla fine del 2020.

I modelli M1, M2 e M3 utilizzano un'architettura ARM, non x64.

Questi Mac non possono avviare distribuzioni Linux forensi, anzi non possono avviare del tutto sistemi operativi esterni:

“ Yes, you can create a bootable installer [...], **but your Mac won't actually start up from it.** Instead, it will start up from an internal copy of macOS Recovery, and only leverage your bootable installer when you choose to reinstall macOS.

[HTTPS://DISCUSSIONS.APPLE.COM/THREAD/254091163](https://discussions.apple.com/thread/254091163)





Un nuovo paradigma

Non possiamo ottenere un'immagine fisica (decifrabile).

È utile pensare all'acquisizione forense dei Mac con Apple Silicon **nello stesso modo in cui si opera sui moderni smartphone.**

Quando non è possibile ottenere un'immagine fisica, ci sforziamo di ottenere un'estrazione Full File System (FFS) mentre il dispositivo è acceso.

Fuji: Forensic Unattended Juicy Imaging

Fuji è un'applicazione per l'acquisizione forense dei Mac, che fornisce al consulente un'immagine **Full File System**.

Offre un'interfaccia grafica modulare, estensibile e facile da usare, che sfrutta vari strumenti di macOS. **È gratis e open source.**

Fuji è anche una tipologia di mela.



[HTTPS://GITHUB.COM/LAZZA/FUJI](https://github.com/lazza/fuji)

Interfaccia

The screenshot shows the Fuji - Forensic Unattended Juicy Imaging application window. The title bar reads "Fuji - Forensic Unattended Juicy Imaging". The main title is "Fuji" in a large, bold font, followed by "Forensic Unattended Juicy Imaging" in a smaller font. Below this, it says "Version 1.1.0 by Andrea Lazzarotto" and provides a URL: <https://andrealazzarotto.com>.

The interface contains several input fields and buttons:

- Case name:** A text field containing "Fuji test".
- Examiner:** A text field containing "Andrea Lazzarotto".
- Notes:** An empty text field.
- Image name:** A text field containing "0123456789_Acquisition".
- Source:** A text field containing "/" and a "Browse" button. Below it is a button labeled "List of drives and partitions".
- Temp image location:** A text field containing "/Volumes/Fuji" and a "Browse" button.
- DMG destination:** A text field containing "/Volumes/Fuji" and a "Browse" button.
- Acquisition method:** A dropdown menu showing "ASR" with a blue arrow icon.

Below the dropdown menu, there are three sections of text:

- ASR:** Apple Software Restore logical acquisition. This is the recommended option, but it works only for volumes.
- Rsync:** Files and directories are copied using Rsync. This is slower but it can be used on any source directory. Errors are ignored.
- Sysdiagnose:** System logs and configuration. This only acquires system data and unified logs (converted to SQLite).

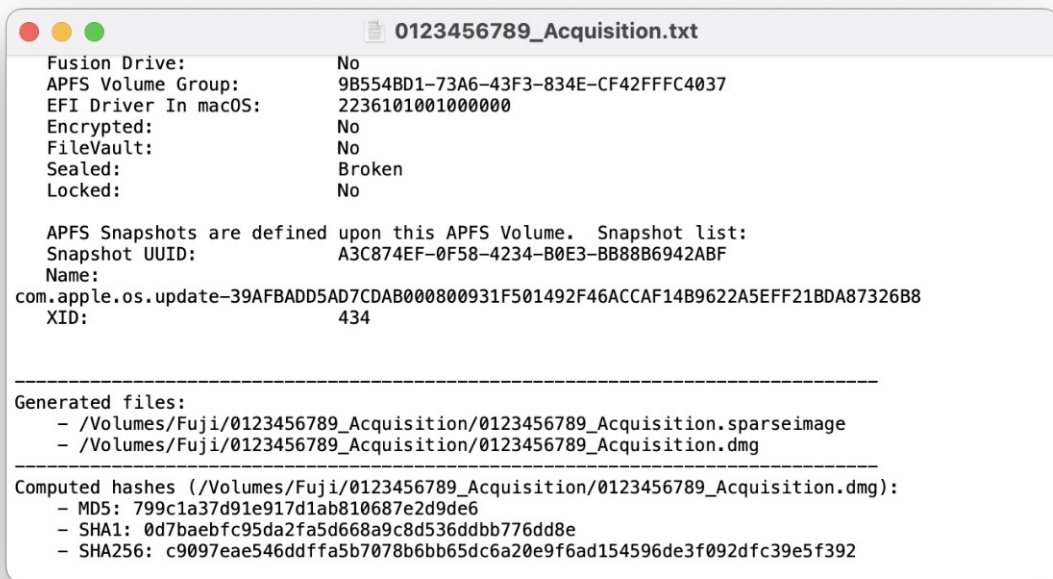
At the bottom, there is a checkbox labeled "Play loud sound when acquisition is completed" which is checked, and a "Continue" button.

DATI DEL CASO

SORGENTE E DESTINAZIONE

METODO DI ACQUISIZIONE

RISULTATO



```
0123456789_Acquisition.txt

Fusion Drive:                No
APFS Volume Group:           9B554BD1-73A6-43F3-834E-CF42FFFC4037
EFI Driver In macOS:         2236101001000000
Encrypted:                   No
FileVault:                   No
Sealed:                      Broken
Locked:                      No

APFS Snapshots are defined upon this APFS Volume.  Snapshot list:
Snapshot UUID:               A3C874EF-0F58-4234-B0E3-BB88B6942ABF
Name:
com.apple.os.update-39AFBADD5AD7CDAB000800931F501492F46ACCAF14B9622A5EFF21BDA87326B8
XID:                          434

-----
Generated files:
- /Volumes/Fuji/0123456789_Acquisition/0123456789_Acquisition.sparseimage
- /Volumes/Fuji/0123456789_Acquisition/0123456789_Acquisition.dmg
-----
Computed hashes (/Volumes/Fuji/0123456789_Acquisition/0123456789_Acquisition.dmg):
- MD5: 799c1a37d91e917d1ab810687e2d9de6
- SHA1: 0d7baebfc95da2fa5d668a9c8d536ddbb776dd8e
- SHA256: c9097eae546ddffa5b7078b6bb65dc6a20e9f6ad154596de3f092dfc39e5f392
```

Fuji genera un report e un file DMG in sola lettura contenente tutti i dati acquisiti.

Può essere aperto con le principali suite di analisi forense.

Alla fine si può eliminare la *sparse image* temporanea.

COMPATIBILITÀ CON I SISTEMI OPERATIVI

10.10+

Rsync

L'opzione più compatibile: funziona con qualsiasi Mac rilasciato negli ultimi dieci anni.

11+

ASR e Sysdiagnose

Entrambi i metodi sono particolarmente adatti ai nuovi Mac, Apple Silicon e Intel.

RIFERIMENTI E CONTATTI

Web

www.dalchecco.it

Company

www.forensier.it

X (Twitter)

[@forensico](https://twitter.com/forensico)

LinkedIn

[dalchecco](https://www.linkedin.com/company/dalchecco)

Web

andrealazzarotto.com

GitHub

[Lazza](#)

X / Twitter

[@thelazza](https://twitter.com/thelazza)

Mastodon

[@lazza@mastodon.social](https://mastodon.social/@lazza)