APPUNTI
Fondamenti di Sicurezza e Privacy

Università degli studi di Verona

Antonio Panfili

22 febbraio 2022

Indice

1 Introduzione

L'obiettivo della sicurezza informatica è proteggere i dispositivi che ognuno di noi utilizza ed i servizi a cui accediamo oltre a prevenire l'accesso non autorizzato alle informazioni personali che teniamo nei device ed online.
Gli **elementi della sicurezza informatica** sono:

- Confidentiality: criptazione
- Integrity: modifiche non autorizzate $->$ controlli hash
- Availability: Distributed Denial of Service (DDoS) $->$ più server
- Authenticity: furto password $->$ A2F
- Accountability: accesso non autorizzato $->$ log di sistema
- Safety

I **concetti chiave della sicurezza**:

- Assets: tutto ciò che ha valore per un'organizzazione sia software che hardware che cloud
- Vulnerability: un bug, debolezze o errori che compromettono integrità o fornitura di servizio
- Cyber Threat: ogni circostanza che può compromettere le operazioni di un'organizzazione, i suoi asset, individui, altre organizzazioni o la nazione attraverso l'utilizzo di attacchi che si servono delle vulnerabilità
- Attack: la realizzazione di un threat specifico che impatta su uno o più degli elementi della sicurezza informatica
- Threat Actor: attaccante
- Risk: livello potenziale dell'impatto del threat e la probabilità che accada
- Controlli di sicurezza: gestione e controlli tecnici prescritti per proteggere il sistema

1.1 Attaccanti - Cybercriminali

Gli attaccanti sono di norma i **Cybercriminali** tipicamente interessati al profitto illegale. La tipologia degli attacchi è:

- Malware
- Ransomware
- Data breaches
- DDoS

ed i vettori di attacco sono altri malware, email e botnet.

1.2 Attaccanti - Nation State

In alcuni casi però l'attaccante è un **Nation State** o più d'uno. Quando ad attaccare sono degli "stati", il capitale per l'attacco è molto alto e quindi di norma il codice malevolo sfrutta una o più "Zero-Day".
I Nation State sono interessati in spionaggio, sabotaggio e sovversione (es: elezioni politiche). Gli attacchi sono effettuati tramite malware molto sofisticati con tecniche

di offuscamento avanzate e colpiscono con attacchi di tipo DDoS, Data breach e malware.
Uno dei casi più famosi di attacco condotto da Nation State è **Stuxnet**, un worm creato si presume da USA e Israele per rallentare lo sviluppo del programma nucleare Iraniano.
Un altro esempio è l'attacco russo per screditare la Clinton ed influenzare le elezioni politiche degli USA.
Altri ancora sono SolarWinds [1] (vedi la sottosezione Cyber Threat Landerscape) e Kaseya.

1.3 Attaccanti - Hacktivists

Gli **Hacktivists**, attivisti, sono invece attaccanti motivati da visioni politiche, credenze religiose, attivismo, ideologie terroriste o divertimento. Agiscono sfruttando kit di exploit, email e botnet per attacchi come lo sfregio di siti web, la pubblicazione di informazioni confidenziali e i DDoS.

1.4 Attaccanti - Insider Threats

Infine gli **Insider Threats** sono ex dipendenti che essendo ancora in possesso di credenziali per l'accesso a risorse di valore, o attaccano intenzionalmente per pubblicare le informazioni sul web, installare bombe logiche o rubare e vendere informazioni, oppure postano accidentalmente contenuti classificati o visitano siti malevoli infettando conseguentemente anche la rete aziendale.

2 Cyber Threat Landerscape

Per avere un'ottica degli attacchi che avvengono quotidianamente è possibile visitare il sito web `www.threatmap.checkpoint.com` che tiene traccia di tutti gli attacchi quotidiani e mostra una mappa in tempo reale degli stessi. I trend dell'anno sono:

- Attacchi legati al Covid-19
- Attacchi ransomware
- Attacchi di supply chain
- Attacchi cloud
- Attacchi IoT

2.1 Attacchi di supply chain

Supply chain si riferisce all'ecosistema del processo, le persone, organizzazioni ed i distributori coinvolti nella creazione e nell'invio del prodotto finale[2]. È scomponibile in quattro elementi chiave:

- Supplier: è il fornitore del prodotto o servizio
- Supplier assets: sono gli elementi di valore che servono al fornitore per produrre il prodotto o servizio
- Customer: è l'entità che consuma il prodotto o servizio

- Customer assets: sono gli elementi di valore posseduti dal customer

Le aziende esternalizzano sempre più l'asset (es: spostano su cloud) e per farlo si appoggiano a major vendor come Microsoft, Apple, Amazon. In alcuni casi è esternalizzato l'intero sistema. L'attaccante quindi hackerando l'attaccante e passando per tutta la catena di suppliers, raggiunge infine il customer in cui viene portato a compimento l'attacco vero e proprio.

2.2 Attaccare il venditore

Un attacco di questo tipo al supplier può essere fatto in vari modi:

- Software pre-esistente: software utilizzato dal supplier, web servers e quant'altro
- Librerie software: librerie e pacchetti di terze parti
- Codice: software prodotto dal supplier
- Configurazioni: password, chiavi API, regole del firewall, URL
- Dati: informazioni del supplier, valori dei sensori, certificati, dati personali di customer e supplier
- Processi: aggiornamento, backup, processo di firma dei certificati
- Hardware: hardware prodotto dal supplier, chip, valvole, USB
- Persone: individui con permessi di accesso ai dati, all'infrastruttura o ai dipendenti

Le tecniche utilizzate per condurre attacchi di supply chain sono:

- Infezioni di malware: spyware per rubare credenziali degli impiegati
- Ingegneria sociale: phishing, applicazioni false, convincere il venditore a compiere azioni
- Attacchi di forza bruta: indovinare una password SSH, indovinare una login
- Sfruttare una vulnerabilità software: SQL injection o buffer overflow nelle applicazioni
- Sfruttare una vulnerabilità di configurazione
- Attacco fisico o modifiche: modificare l'hardware o intrusioni fisiche
- L'intelligence open-source (OSINT): cercare online le credenziali, chiavi API o nomi utente
- Contraffazione: imitare una USB con intenzioni malevole

2.3 Attaccare il cliente

Per attaccare un cliente invece vengono sfruttate le seguenti tecniche:

- Relazioni fidate: fidarsi di un certificato, di un aggiornamento automatico o di un backup automatico
- Drive-by compromessi: script malevoli in un sito web per infettare utenti con malware
- Phishing: scrivere messaggi impersonando il venditore, false notifiche di aggiornamento

- Infezioni di malware: Remote Access Trojan (RAT), backdoor, ransomware
- Attacchi fisici o modifiche: modificare l'hardware o intrusioni fisiche
- Contraffazione: false USB, modificare motherboard, impersonare personale del venditore

e l'obiettivo è ottenere l'asset del customer che corrisponde a:

- Dati: dati di pagamento, video, documenti, emails, piani di viaggio, dati di vendita e finanziari, proprietà intellettuale
- Dati personali: dati del customer, credenziali, registro dei dipendenti
- Software: accesso al codice sorgente del cliente, modifica del codice del cliente
- Processi: documentazione dei processi interni, operazioni e configurazioni, inserimento di processi malevoli, documenti di schematiche
- Larghezza di banda: usare la banda per compiere DDoS, inviare email SPAM o compiere infezioni su larga scala
- Finanziari: rubare cryptovalute, dirottare trasferimenti bancari
- Persone: individui presi di mira per la loro posizione o conoscenza

2.4 SolarWind

Gli hacker russi, probabilmente sovvenzionati dallo stato, hanno manomesso Orion[1], una piattaforma di monitoraggio e analisi di proprietà di SolarWind, e rilasciato un aggiornamento malevolo. Essendo utilizzato da più di 300.000 aziende l'attacco ha prodotto molteplici vittime.
L'attacco ha seguito le seguenti fasi:

1. Initial Infiltration: sfruttamento di una vulnerabilità nel servizio di autenticazione. Questo ha permesso agli attaccanti di accedere persistentemente nelle imprese vittime ed esaminare email e sviluppare profili degli sviluppatori da colpire
2. Reconnaissance: avviata una campagna di phishing dedicata agli sviluppatori scelti
3. Spear Phishing: infettati i le instanze locali degli sviluppatori vittima
4. Weaponization: manipolati i sistemi per inserire delle backdoor
5. Infiltration of Downstream Users: abusare della relazione di fiducia per penetrare l'infrastruttura dell'utente finale

2.5 COVID-19 Related Attacks

Con la pandemia da covid-19 gli attacchi che hanno utilizzato la situazione come pretesto sono aumentati. Si rivelano attacchi tramite phishing [3], malware, domini malevoli e fake news.
Le campagne di phishing composte da email con allegato malevolo sono state inviate sfruttando motivazioni legate alla situazione pandemica adducendo quindi a presunti salari anticipati, o al risultato di tamponi non effettuati.

[1] https://www.solarwinds.com/orion-platform

Oltre a questa tipologia persistono anche le cosiddette "frodi nigeriane" che sfruttano la debolezza umana nel voler ricevere soldi gratuitamente e quindi tombole, giveaway, eredità ecc.
Le campagne di phishing puntano sull'autorevolezza della fonte, fingendosi enti governativi, pubblici o aziende private molto famosi, e sull'urgenza in modo da spingere la vittima a scaricare l'allegato o ad inviare le credenziali senza riflettere a fondo.
Non solo le campagne di phishing hanno però sfruttato il contesto pandemico. Sono molte infatti le app malevole (Trojan Android application PacKage (APK)) spacciate per App Immuni o altre legate al Covid-19.
Alcuni malware sono sfruttati per poi scaricare altri malware sulla macchina vittima. Uno dei casi più famosi è quello di **Emotet**[2], un malware nato nel 2014, si pensa in Ucraina, per furto di dati bancari che ora viene utilizzato per scaricare altri malware.
Emotet sfrutta delle macro, ora disabilitate di default nei sistemi Windows, per far eseguire codice malevolo alla macchina vittima. Disabilitare le macro di default non è una soluzione ma solo uno step in più per l'attaccante che ora deve convincere la vittima anche ad abilitarle. Nel 2021 la collaborazione tra Germania e Ucraina ha portato al blocco dei server utilizzati per spargere l'infezione di Emotet. Prima del blocco, essendo stati sequestrati i server di comando del malware, è stato inviato dalle autorità un comando per la rimozione del malware a tutte le macchine infette.
I creatori di Emotet, essendo riusciti ad infettare molte macchine, permettevano lo sfruttamento della loro botnet per spargere altri virus o eseguire attacchi DDoS.
L'evoluzione di Emotet in loader, malware utilizzato per il download di altri malware, è stato portata alla ribalta dal propagarsi dell'infezione di **Trickbot**[3] nel 2020 (maggiori dettagli nella sottosezione TrickBot dedicata).
Trickbot è altro trojan nato nel 2016 per il furto di dati bancari e che anch'esso viene sfruttato come loader di altri malware e si pensa diventerà il nuovo Emotet.
Anche in Italia vengono sviluppati malware ed è stato riportato l'utilizzo di un ransomware spacciato come App Immuni attraverso campagne di phishing.
Da notare anche l'incremento di registrazioni di domini legati al Covid-19 utilizzati come domini malevoli [4]. Il phishing attraverso sito web malevolo funziona copiando siti web conosciuti (es: banche, e-commerce...) che registrano i dati inseriti dagli utenti. Se qualche anno fa avere il certificato Secure Sockets Layer (SSL) (quindi comunicazione garantita attraverso HTTPS e lucchetto nella barra di ricerca dei più blasonati browser web) era ritenuto un valore di sicurezza del sito web, oggi anche i siti web malevoli si adoperano per ottenere la certificazione ed aumentare la loro credibilità.
Le fake news sono state utilizzate massivamente per alimentare il panico e favorire il proliferare di phishing e virus informatici [5]. Tra i gruppi di hacker che hanno operato in questo periodo di pandemia ci sono anche quelli che si presume essere finanziati da Nation State come Russia e Nord Corea tra cui i gruppi Stronzium e NKO.
Sono stati colpiti ospedali e centri di ricerca per il furto di dati sia con attacchi di phishing che di forza bruta per ottenere le credenziali.

[2]https://en.wikipedia.org/wiki/Emotet
[3]https://en.wikipedia.org/wiki/Trickbot

2.6 Ransomware

I ransomware sono incrementati del 90% negli ultimi anni. Agiscono criptando file specifici data la loro estensione e chiedendo un riscatto per ricevere la chiave di decriptazione.
In alcuni casi oltre a criptare i dati li copiano anche in modo da richiedere un doppio riscatto (chiave di decriptazione e non pubblicazione dei file).
In altri il riscatto è triplo perché viene richiesto un pagamento anche agli interessati dai dati rubati come per esempio quanto accaduto dopo il furto di dati ad una compagnia che teneva un registro delle diagnosi private ed interviste ai pazienti. Un caso simile è accaduto ad un'azienda collaboratrice di Apple alla quale hanno rubato dei blueprint di prodotti Apple. L'azienda collaboratrice non ha voluto pagare il riscatto così gli attaccanti hanno contattato direttamente Apple la quale alla fine ha preferito pagare il riscatto.
A volte i ransomware colpiscono aziende o fornitori di alto rilievo causando disagi molto importanti come accaduto alla Colonial Pipeline, fornitore di carburante in tutti gli USA. In questo caso l'attacco ha provocato un blocco delle forniture di un'infrastruttura pubblica.
Ci sono ransomware che vengono progettati per attaccare una determinata azienda come accaduto a Garmin con l'attacco WastedLocker che ha criptato tutti i file con l'estensione *.garminwasted*. Il ransomware in questione ha sfruttato il Windows Cache Manager per bypassare il software antivirus. Solitamente i ransomware vengono identificati per il comportamento "apri e cifra" ripetuto con intensità elevata nell'arco di tempo. In questo caso invece i file vengono caricati in cache e modificati all'interno della cache per poi riscaricarli senza destare sospetti.

2.7 CryptoMiner

I cryptominer sono malware che vengono installati ed eseguiti sulle macchine vittima per minare criptovalute in background.
Il mining è un'operazione che richiede una grande potenza di calcolo per ottenere un riscontro significativo quindi solitamente vengono infettate molte macchine in modo da creare una rete di bot detta "botnet" che compie le azioni di mining.
Xring è un software open source utilizzato per il mining legittimo ma che viene sfruttato dai criminali per il mining malevolo.

2.8 Android

Anche i dispositivi mobile ed Internet of Things (IoT) non sono immuni dai virus, anzi. **FluBot** è un malware che colpisce sistemi Android. Il malware viene fatto scaricare da un link inviato per SMS. Una volta scaricato l'APK e convinto l'utente ad installarlo disabilitando la protezione per l'installazione di APK di terze parti (non scaricati attraverso store ufficiali), sfrutta i permessi per rubare dati di carte di credito.
Gli smart watches sono anch'essi dispositivi oggetto di attacchi malware e possono causare gravi danni anche alle persone basti considerare app che ricordano di assumere il medicinale attraverso notifica per pazienti con deficit di memoria. Modificare

lo scheduling può far assumere i medicinali in quantità eccessive e causare fin anche la morte della vittima.
Il fatto è che non si pensa mai al rischio di un attacco ma sempre al "chi vuoi che ci attacchi?". Ciò che invece andrebbe imposto è l'utilizzo del sistema di progettazione definito "security by design"

Tra i tool utilizzati per compiere attacchi c'è anche il noto tool a pagamento (di cui si trovano le crack free online) Cobalt Strike. Questo software è venduto come tool per il penetration testing poiché permette di portare a termine varie tipologie di attacchi ma spesso viene utilizzato con scopi malevoli.

2.9 IoT

L'IoT in questo contesto è molto vulnerabile agli attacchi informatici. Il 57% dei dispositivi IoT sono vulnerabili ad attacchi di media o alta severità.
Gli strumenti IoT comunicano verso una app di riferimento oppure verso un cloud centrale. Spesso queste comunicazioni non sono criptate e dunque è possibile leggerle e modificarle.
Consideriamo il caso di un paziente diabetico con un iniettore di insulina IoT al quale vengono inviati dei dati modificati per far si che somministri una dose mortale di insulina. Nel Maggio 2021 una falla di questo tipo ha costretto al ritiro di un dispositivo di questo genere[4].
Molto spesso viene utilizzata una password di default oppure il sistema in esecuzione non è aggiornato (XP, 7 non più supportati) o non è nemmeno aggiornabile.
Gli IoT inoltre, per limitare il consumo di energia (pensiamo a prodotti wearable) tende ad avere una memoria interna bassa con una potenza computazionale minima e questo rende difficile implementare sistemi di autenticazione per rendere più difficile l'attacco da parte di criminali.
Ci sono vari esempi di come falle in librerie o errori di scrittura del codice possono portare a compormettere sistemi compresi quelli IoT.
Il C è un linguaggio in cui è facile generare errori di memoria. Per esempio se ad una `malloc` viene passato un numero troppo grande si ottiene un "Integer Overflow" che rende possibile l'esecuzione di istruzioni malevole.
Nel 2020 è sono state scoperte 19 vulnerabilità nella libreria **Ripple20** per la comunicazione Transmission Control Protocol/Internet Protocol (TCP/IP). Alcune di queste vulnerabilità sono state valutate con il livello massimo di gravità perché permettono l'accesso a dati sensibili o l'accesso da remoto a macchine vittima.
Il caso singolare dei comandi inviati attraverso ultrasuoni (Siri), laser o clonando la voce della vittima ma interpretati comunque dagli assistenti vocali rende l'idea delle potenziali vulnerabilità legate a questa tecnologia.
Non mancano poi attacchi effettivi come quelli compiuti ai danni di videocamere, elettrodomestici di vario genere ed anche baby monitor. Il problema molto spesso risiede nella mancanza di una password o nella debolezza della stessa.
Il più grande portale/motore di ricerca di dispositivi connessi in rete è **SHODAN**[5].

[4]https://www.fda.gov/medical-devices/medical-device-recalls/medtronic-recalls-remote-controllers-used-paradigm-and-508-minimed-insulin-pumps-potential

[5]https://www.shodan.io/

Fornisce molte informazioni su ogni dispositivo dalla tipologia di prodotto alle porte aperte.
Il problema delle password deboli è stato ben sfruttato dalla botnet **Mirai** che ha sfruttato le macchine vittime per attacchi DDoS. Il sistema di infezione di Mirai scansiona costantemente il web alla ricerca di devices IoT accessibili via internet per poi inserire le login di default del prodotto o tentare login molto comuni. Una volta ottenuto l'accesso alla macchina vittima Mirai installa il malware che la forza a fare riferimento ad un server centrale di controllo che la rende un bot da utilizzare in attacchi DDoS.
I dispositivi più vulnerabili nel mondo medicale sono soprattutto dispositivi di diagnostica (Imaging System, Patient Monitoring, Medical Device Gateway).

2.10 Cloud Attacks

Gli attacchi a sistemi cloud sono per la maggior parte conseguenza di errate configurazioni dei sistemi stessi.
È emblematico il caso dell'esposizione dei dati degli utenti relativo al servizio Amazon S3 Bucket, servizio di Amazon Web Services (AWS) nel quale è molto difficile non fare errori di configurazione dei bucket. Alcuni penetration tester hanno infatti scoperto alcuni bucket aziendali con tutte le informazioni erroneamente accessibili al pubblico. Questi tipi di vulnerabilità sono stati scoperti anche in cloud universitari e di altre aziende.
Microsoft non ne è immune; il sistema di permessi di default di Power Apps ha causato errori di configurazione che hanno rivelato milioni di dati privati. Il sistema di Application Programming Interface (API) per accedere ai dati, costringeva infatti a riconfigurare la privatizzazione dei permessi di lettura dei dati manualmente.
Un altro caso è quello dell'attacco di **DroppelPaymer** nei confronti di **Bretagne Télécom** nel quale è stato sfruttato una falla di Citrix ADC per installare un ransomware. L'azienda si è salvata solo grazie al backup scollegato.
Ed ancora le due vulnerabilità di Microsoft Azure una su Azure Stack e Data service per i quali attraverso determinati percorsi non era richiesta l'autenticazione per accedere al servizio.

3 Cyber Kill Chain

La cyber kill chain è la catena di controllo del sistema che permette di capire se si è sotto attacco e di che tipo di attacco.
Si divide in fasi:

- Reconnaissance
- Weaponization
- Delivery
- Exploitation
- Installation
- Command and Control
- Actions on Objectives

3.1 Reconnaissance

Obiettivo: selezionare ed ottenere informazioni sull'individuo vittima (cliente dell'azienda oppure dell'IT...) Fase passiva: ottenere informazioni senza interagire con la vittima (whois, shodan, google, social media, mantego) Fase attiva: ottenere informazioni attraverso l'interazione con la vittima (nmap, port scanning, vulnerability scanners)

3.2 Weaponization

Obiettivo: trovare o creare l'attacco per sfruttare la vulnerabilità attraverso l'utilizzo di

- Metasploit
- Exploit DB
- Veil Framework
- Social Engineering Toolkit
- Cain and Abel
- Aircrack
- SQL Map
- Malware ad hoc

3.3 Delivery

Obiettivo: scegliere come inviare l'attacco

- Web site: compromettendo siti web molto utilizzati in modo che la visita inneschi il download del file malevolo
- Social Media: utilizzando profili fake per attacchi di ingegneria sociale
- User Input: avendo accesso a tastiere ecc... della vittima
- Email: inviando il file malevolo per email
- USB: lasciando pendrive incustodite

3.4 Exploitation

Obiettivo: sfruttare una vulnerabilità già nota come ad esempio:

- SQL Injection
- Buffer overflow su software già presente nella macchina
- Malware
- Javascript hijacking cioè redirect attraverso js
- User exploitation

3.5 Installation

Obiettivo: mantenere l'accesso nel sistema vittima attraverso l'utilizzo delle seguenti tecniche

- DLL Hijacking cioè cambiare la libreria legittima di un programma con quella malevola che avvia il malware all'avvio del programma legittimo
- Meterpreter
- Remote Access Trojan
- Registry Changes cioè la modifica dello scheduling d'avvio in modo da avviare il malware in automatico all'accensione della macchina.
- PowerShell commands

3.6 Command and Control (C2)

Obiettivo: stabilire un canale di comunicazione tra attaccante e macchina vittima in modo da manipolare la stessa da remoto ad esempio aprendo canali di comunicazione a due vie attraverso HTTP/HTTPS o cloud, agendo su DNS e protocolli email e collegandosi direttamente a server di controllo in possesso degli attaccanti oppure ad altre macchine vittima utilizzate per questo scopo

3.7 Actions on Objectives

Obiettivo: azioni intraprese per conseguire l'obiettivo dell'attacco tra cui

- Ottenere i dati dell'utente
- Ottenere maggiori privilegi
- Ricognizione interna
- Muoversi all'interno del sistema
- Rubare dati
- Distruggere il sistema
- Sovrascrivere o corrompere dati
- Modificare dati di nascosto

3.8 TrickBot

È un trojan avanzato che i criminali spargono principalmente attraverso campagne di email phishing in cui viene inviato un allegato malevolo o un link ad un server malevolo che fa scaricare il malware contenente il payload che una volta eseguito si collega al server C2 e scarica TrickBot.
La struttura del malware è modulare tale per cui i moduli del malware lo spargono in giro abusando del protocollo Server Message Block (SMB).
Gli attaccanti possono usare TrickBot per: È capace di rubare dati attraverso un server C2, minare criptovalute e riconoscere altri host nella rete tracciandone i dettagli. In quest'ultimo caso a TrickBot è aggiunto un file di configurazione dedicato che ne specifica i passaggi da compiere.

- Scaricare altri malware come **Ryuk**[6] e **Conti**[7] ransomware
- Essere utilizzato come un downloader stile Emotet

[6]https://en.wikipedia.org/wiki/Ryuk_(ransomware)
[7]https://en.wikipedia.org/wiki/Conti_(ransomware)

Figura 1: fonte `https://blog.gigamon.com/2019/03/02/revisiting-prolific-crimeware-to-improve-network-detection-trickbot/`

3.9 Ransomware Cyber Kill Chain

1. La vittima riceve un'email di phishing contenente un link ad un sito web infetto e lo visita
2. Il web server malevolo trova una vulnerabilità nel sistema vittima
3. Il ransomware viene automaticamente scaricato e si esegue
4. L'eseguibile cancella eventuali copie di sé stesso e si propaga attraverso il file system
5. Il ransomware trova i file con un'estensione specifica e li cripta
6. Il ransomware contatta il server C2 e invia la chiave di criptazione assieme ai dettagli della macchina vittima
7. Il ransomware riceve le informazioni di pagamento dal server C2
8. Il ransomware imposta un conto alla rovescia prima di cancellare ogni file della vittima
9. Il ransomware mostra le informazioni di pagamento
10. Se la vittima paga gli attaccanti, il ransomware potrebbe contattare il server C2 per ricevere la chiave di decriptazione
11. Se la vittima non paga gli attaccanti in tempo, la chiave di decrittazione viene distrutta

La fase di weaponization si differenzia a seconda della scelta fatta dagli attaccanti. Il ransomware è basato su script e quindi scaricato da uno script e caricato direttamente in memoria; è nascosto all'interno di un file di formato differente e tutte le corrispondenze del file nel Master File Table (MFT) vengono rimosse.
Per evitare di essere scovati dagli anti-virus i ransomware cifrano ad intervalli di tempo oppure solo quando la vittima esegue un backup, attendono che il codice sia in memoria pronto all'esecuzione per eliminare i file del malware stesso, cifrano le comunicazioni attraverso rete **Tor**, per prevenire il reverse engineering o il codice viene cifrato oppure riempito di funzioni inutili.
La fase di delivery è piuttosto standard attraverso phishing, spear phishing, malvertisement o sistemi di distribuzione del traffico.
La exploitation si avvale di **Anger EK** per sfruttare le vulnerabilità di Adobe Flash e Microsoft Silverlight, **Neutrino** e **Magnitude EK** che sfruttano le vulnerabilità di Adobe Flash e **Blackhole EK** che invece si focalizza sulle vulnerabilità di Adobe Reader e Adobe Flash Java.
Oltre agli exploit kits c'è anche la possibilità che sfruttino vulnerabilità ancora non note "Zero-Day" ed altre vulnerabilità scoperte durante la fase di ricognizione.
L'installazione prevede di rendere i file della vittima inutilizzabili criptandoli, comprimendoli in zip con password, criptando l'MBR e distruggendo i backup ma anche la diffusione del ransomware attraverso la rete. La propagazione dell'infezione può essere ottenuta attraverso varie modalità.
L'utilizzo di dispositivi rimovibili è piuttosto semplice considerando che sono facili da trovare nel sistema, facili da scollegare, possono contenere e condividere informazioni e si può fare leva sulla funzionalità dell'"AutoPlay".
Per eseguire un l'auto-play/auto-run è necessaria la presenza di un file denominato "autorun.inf" che è utilizzato per facilitare l'avvio dell'interfaccia grafica per i CD.
Un'altra modalità di propagazione dell'infezione è attraverso la condivisione di file

molto comune nel mondo IT nata prima della condivisione via cloud e della tecnologia di sincronizzazione.
In pratica è un server di condivisione dei documenti. Basta che uno solo dei dipendenti carichi un file infetto perché anche gli altri lo ricevano e vengano infettati.
Nel caso di cartelle condivise, per evitare di far copiare l'intero malware col rischio che venga bloccato dall'anti-virus del ricevente, si usa creare un collegamento al desktop della macchina infetta e copiare solo il collegamento. Il collegamento contiene il comando di esecuzione del malware così una volta eseguito dall'utente viene eseguito il malware della macchina infetta nella nuova macchina vittima.
La fase di command and control viene prima di iniziare la criptazione così da ricevere subito la chiave di criptaggio e dopo così che il malware possa ricevere i dettagli per il pagamento. Per contattare il server C2 a volte viene inserito l'indirizzo IP nel codice, altre volte si utilizza un algoritmo di generazione dei domini in modo che venga aggiornato il dominio ogni volta così da non poterlo inserire nella lista nera dei domini malevoli, altre volte ancora attraverso una botnet.
L'action and objectives si ottiene ricevendo il pagamento. Nei primi ransomware si chiedeva un pagamento via Paypal, oggi è più comune la richiesta di Bitcoin o attraverso portale o attraverso portali anonimi che gestiscono direttamente il pagamento e l'invio del software di decriptazione.
Per prevenire un attacco ransomware non cliccare mai su link non verificati, non aprire allegati email non fidati, non scaricare mai software da siti web non fidati, non pubblicare informazioni personali, mantieni aggiornati i software e l'OS e fai il backup dei dati.
Per rispondere ad un attacco ransomware la prima cosa da fare è scollegarsi dalla rete per non far propagare il malware, eseguire una scansione con l'internet security software, utilizzare un software per la decriptazione o, nel caso si abbia un backup, ripristinare il backup.
Esistono anche dei siti che permettono di risolvere alcuni attacchi ransomware come `www.nomoreransom.org`

4 MITRE Attack

La cyber kill chain non è ottimale per definire attacchi più complessi ed inoltre è una struttura ad alto livello.
Esiste quindi **MITRE**[8], un'associazione che ha creato un database con molti attacchi noti e non e la loro descrizione.
Il sistema di definizione MITRE si basa su alcune matrici "PRE-ATT&CK" e "ATT&CK" oltre ad altre più specifiche per piattaforme enterprise e mobile. La prima matrice copre le fasi di Reconnaissance e Weaponization, la seconda dalla Deliver alla Maintain. Cercando TrickBot si può vedere dettagliatamente ogni tecnica utilizzata in ogni fase dell'attacco.

[8]https://attack.mitre.org

5 Social Engineering

L'ingegneria sociale, è lo studio del comportamento di una persona al fine di carpire informazioni utili.
In un'infrastruttura ben protetta il punto debole è spesso l'uomo; il 22% degli attacchi di phishing utilizza l'ingegneria sociale.
Alcune delle figure che utilizzano in modo criminale l'ingegneria sociale sono:

- Hacker: l'obiettivo è far installare alla vittima un software per poi accedere a conti correnti ecc...
- Identity Thieve: l'obiettivo è rubare i dati della vittima per impersonarla oppure per venderli nel dark web
- Scam Artist: l'obiettivo è convincere la vittima ad inviare soldi al truffatore

Un caso fortunato è quello della **Bangladesh Bank** che ha sventato una truffa da un miliardo di dollari grazie ad un errore di scrittura dei truffatori.
Meno fortunato invece il presidente francese **Macron**. Attraverso l'ingegneria sociale gli attaccanti sono riusciti ad avere accesso al suo account social e hanno invitato i suoi stessi follower a non votarlo.
Parecchi danni li ha fatti anche la **Facebook Lottery Scam** cioè lotterie truffa che si fa forte del nome Facebook per attirare vittime. I malintenzionati chiedono soldi alle vittime promettendo un regalo di molto più alto valore per ingannarle.

5.1 Attack Life Cycle

Il primo passo del ciclo di vita di un attacco è ottenere informazioni.
Un sistema per farlo è quello di spiare la vittima (es: spiare persone che lavorano al pc in treno o in luoghi pubblici), un altro è frugare nella spazzatura alla ricerca di post-it e documenti.
Il secondo è stabilire relazioni impersonando qualcuno per poi passare al terzo passo che è quello esplorativo in cui si rubano le informazioni e si inseriscono le backdoor ed infine il quarto passo è l'esecuzione finale con cui si ottiene il risultato desiderato.
Dal quarto passo si può tornare al primo nel caso in cui l'obiettivo sia recuperare informazioni per svolgere altri attacchi.
I tipi di attacchi che sfruttano l'ingegneria sociale sono:

- Phishing: attacco via email in cui è allegato un file malevolo o inserito un link a sito malevolo. Il testo spinge la vittima ad agire in tempo breve e/o sfrutta una finta posizione di forza
- Spear phishing: attacco di phishing targhetizzato sul profilo della vittima
- Whailing: attacco di phishing mirato a persone di alto profilo (CEO, CFO...)
- Viral hoaxe: post o video che stimolano la curiosità dell'utente per reindirizzarlo su un sito malevolo
- Virus hoax: alert falsi attestanti la presenza di virus sulla macchina vittima
- Vishing: "voice-phishing", attacco di phishing condotto al telefono. A volte viene inviato un messaggio in-app che chiede di richiamare il finto servizio clienti

- Impersonation: impersonificazione di autorità, persone più alte in grado o persone di cui ci fidiamo
- SMiShing: "SMS-phishing", attacco di phishing condotto via SMS
- Tailgating: accedere fisicamente ad un'area riservata senza essere notati. Per esempio rimanendo dietro un vero dipendente con il badge oppure fingendosi un corriere

5.2 Attacchi di Phishing

Il phishing è il tentativo di acquisire informazioni sensibili come nomi utente, password, dettagli delle carte di credito ed in alcuni casi indirettamente soldi, fingendosi, in una comunicazione elettronica, entità di cui fidarsi.
Uno dei siti più copiati per questi motivi è **Office 365**.
Quanto sono effettivi gli attacchi di phishing?

- L'88% delle organizzazioni nel modo ha riscontrato attacchi di phishing nel 2019
- Il 95% di tutti gli attacchi in reti aziendali sono il risultato di spear phishing di successo
- Il 22% dei data breaches nel 2019 ha coinvolto attività di phishing
- Il 97% degli utenti non può identificare un'email di phishing sofisticata
- Il 30% delle email sono lette dalle vittime
- Il 12% delle vittime ha cliccato il link malevolo o ha scaricato l'allegato malevolo contenuto nell'email
- Il 15% delle vittime sono contattate almeno un'altra volta nel corso dell'anno

Un caso molto rilevante nato da attività di phishing è il data breach di **Target** che ha portato al furto di dati di carte di credito di 41 milioni di utenti.
In alcuni casi l'email di phishing simula una condivisione di documenti Google Docs da parte di una persona che conosciamo, in altri è particolarmente evidente che si tratta di una truffa.
È importante fare attenzione al reale mittente così come al vero href del link inserito nell'email.

- Exploiting Authority: l'attaccante sfrutta il nome di un'autorità per essere più credibile
- Exploiting Scarcity: l'attaccante sfrutta un presunto tempo limitato prima che un evento accada (es: scadenza di una promozione o pagamento di una multa)
- Exploiting Commitment: l'attaccante sfrutta il fatto che la vittima ottemperi a norme sociali come ad esempio un avviso di comparizione
- Exploiting Linking: l'attaccante sfrutta la fiducia nelle relazioni di amicizia della vittima
- Exploiting Reciprocation: l'attaccante sfrutta il fatto che la gente tende a compiere un'azione se ottiene qualcosa in cambio
- Exploiting Social Proof: l'attaccante sfrutta il concetto che se molte persone hanno già compiuto un'azione allora c'è da fidarsi

La prof. ha condotto uno studio assieme ad un ragazzo della magistrale le cui conclusioni sono che: Il 31.4% degli impiegati ha cliccato il link nell'email di phishing, il 23.4% degli impiegati ha inserito le credenziali nel sito di phishing ed il 69.2% degli impiegati che ha cliccato il link di phishing ha inserito le credenziali nel sito di phishing. In pratica chi ha cliccato il link ha anche inserito le credenziali.
L'exploit che preme sull'urgenza è risultato il migliore (33.8% vittime di urgenza, 21.5% vittime di autorità, 15.6% vittime di phishing senza tecnica di persuasione).
Si è notato anche che i link e la firma dell'email non vengono ispezionati con attenzione. Se si vogliono testare le proprie capacità di riconoscimento delle email di phishing, esiste un servizio dedicato a questo link `www.phishingquiz.withgoogle.com`.

I siti di phishing d'altra parte presentano spesso un indirizzo poco legittimo o l'utilizzo di un sotto dominio fraintendibile o una cartella finale contenente la webapp con nome di un dominio legittimo, e mancano di certificato SSL, quindi di lucchetto HTTPS.
Il sito `https://phishtank.org/` permette di sottomettere un link per avere una valutazione indipendente sulla sua reale legittimità.
Per bloccare attacchi di phishing un'azienda dovrebbe adottare un approccio a livelli multipli:

- Livello 1: rendere difficile ad un attaccante raggiungere l'utente implementando controlli anti-spoofing in modo che sia più difficile per l'attaccante conoscere gli indirizzi email del personale, tenere conto delle informazioni che vengono rese pubbliche attraverso sito web e social oltre a informare anche i dipendenti che ciò rappresenta una vulnerabilità, filtrare e bloccare email di phishing
- Livello 2: Aiutare gli utenti ad identificare e riportare email sospette attraverso la formazione dei dipendenti e creando un sistema che permette ai dipendenti di chiedere aiuto attraverso un sistema di report semplice in una struttura che non li ferisce sentimentalmente in caso di richiesta d'aiuto o d'errore
- Livello 3: Proteggere l'organizzazione dall'effetto di una email di phishing non identificata utilizzando la Two Factor Authentication (2FA) ed una struttura di permessi gerarchica, proteggendo gli utenti da siti web malevoli, utilizzando un server proxy e un browser web aggiornato oltre che tenendo attivo un antivirus su ogni dispositivo
- Livello 4: Rispondere rapidamente agli incidenti definendo un piano di disposta per le differenti tipologie di incidenti includendo le responsabilità legali e rilevandoli velocemente incoraggiando i dipendenti a riportare le attività sospette

6 Cyber War and Attacks to Critical Infrastructures

Si parla di cyber war solo nel caso in cui l'attacco sia rivolto ad un'**infrastruttura critica**.
Le infrastrutture critiche sono quelle strutture, sistemi, siti, informazioni, persone, reti e processi indispensabili per uno stato poiché su di essi dipende la vita di ogni

giorno. Inoltre riguardano anche tutte quelle strutture, siti ed organizzazioni che pur non essendo indispensabili per mantenere un servizio, richiedono la massima protezione come siti chimici e nucleari.
Alcuni esempi di settori critici sono:

- Chimico
- Nucleare civile
- Comunicazioni
- Difesa
- Servizi di emergenza
- Energetico
- Finanziario
- Alimentare
- Governativo
- Medicale
- Spaziale
- Trasporto
- Idrico

Una compromissione di questi asset critici può comportare una mancanza, l'integrità o la ricezione di servizi essenziali ed impatti significativi nella sicurezza dello stato, nella sua difesa e nelle sue funzioni.
Attacchi ad infrastrutture critiche accadono ogni giorno; tra questi ne elenchiamo alcuni:

- 2019: ACSO Industries colpita da ransomware, Norsk Hydro attaccata, colpito il sistema di trattamento dell'acqua a Mosca, piattaforma Noya in Giappone
- 2018: Boing colpita da WannaCry, GreyEnergy APT, Shamoon colpisce Saipem e Fossenheim piattaforma nucleare
- 2017: Wolf Creek piattaforma nucleare, Triton, Energetic Bear, WannaCry, NoPetya
- 2016: Colpita la rete energetica di Kiev e attacco DDoS alla compagnia di riscalfamento Finlandese
- 2015: Energetic Bear attacca un'acciaieria Tedesca e la rete energetica Ucraina
- 2010: Stuxnet

Molte infrastrutture critiche sono controllate e gestite da un Industrial Control System (ICS). Gli ICS controllano molte delle funzionalità che utilizziamo quotidianamente senza che noi ce ne accorgiamo (es: accendere la luce, fare la doccia, bere acqua dal rubinetto...).
Il fatto che gli ICS si stiano sempre più connettendo ad internet li espone a nuovi rischi. Alcune delle vulnerabilità sono:

- Nessun programma di formazione sulla sicurezza per gli ICS
- OS e software installati potrebbero non essere più supportati o ricevere patch
- Utilizzo di configurazioni di default
- Dati non protetti nei dispositivi portatili
- Mancato utilizzo di password

- Applicazione di controlli d'accesso inadeguati
- Controlli di sicurezza fisici inadeguati per sistemi critici
- Buffer overflow
- Mancato tracciamento dei log

6.1 Stuxnet

La prima arma informatica riconosciuta è **Stuxnet**, scoperta nel giugno 2010. Un malware estremamente sofisticato che sfruttava 4 vulnerabilità zero-day e 2 rootkit. Creato presumibilmente dalla National Security Agency (NSA) statunitense assieme a Central Intelligence Agency (CIA) ed ai servizi segreti Israeliani prendeva di mira i Programmable Logic Controller (PLC) usati per il controllo e la gestione delle centrifughe per l'arricchimento dell'uranio ed il 70% delle vittime dell'infezione sono stati siti nucleari Iraniani.
Nello specifico si pensa che qualcuno abbia portato una USB infetta, contenente un "Autorun.inf" lanciato da un exploit della vulnerabilità "LNK", all'interno del sito nucleare perché il sito non è collegato in rete ma il malware ha anche la capacità di propagarsi attraverso le reti infettando macchine WinCC e propagandosi attraverso vulnerabilità di Windows che non erano ancora state scoperte.
Dato che Stuxnet non fa nulla a meno che non noti la presenza del programma specifico dei PLC delle centrifughe, è stato scoperto solo dopo molto tempo.
L'obiettivo è riprogrammare i PLC così da copiare i valori dei sensori in fase di corretta attività e successivamente attuare delle modifiche ai parametri delle funzioni delle centrifughe fornendo però come output i dati di funzionamento corretto registrati in precedenza.
Minime variazioni di pressione e di rotazione delle centrifughe porta allo scoppio delle stesse ed a danni rilevanti che comportano tempi di stop prolungati.
Si presume che in fase di scrittura del codice malevolo gli attaccanti abbiano fatto uso dei video propagandistici dello stato Iraniano scoprendo così dalle riprese alcuni dettagli dei software utilizzati e della struttura stessa in relazione alla disposizione delle centrifughe e da questa dei PLC utilizzati.
Stuxnet ha anche la capacità di collegarsi ad un server C2 (due server HTTP in ascolto su porta 80). Da questi server è capace di scaricare altri file come backdoor o una versione aggiornata del malware.

6.2 Attacchi Sandworm

Il gruppo di criminali ha compiuto molteplici attacchi nel corso degli anni tra cui:

- 2015-16: governo Ucraino e infrastrutture critiche
- 2017: campagna di spearphishing contro il partito del presidente Francese Macron
- 2017: infezione di NotPetya
- 2017: campagna di spearphishing contro ospiti, partecipanti, partner e volontari dell'olimpiade invernale di PyeongChang
- 2017: Olympic Destroyer attacco al sistema IT delle olimpiadi invernali di PyeongChang

- 2018: investigazioni sull'avvelenamento di Novick
- 2018-19: compagnie Georgiane ed enti governativi

Alcuni dei criminali sono stati successivamente arrestati.

6.2.1 BlackEnergy

Il primo attacco è stato fatto ad una centrale elettrica Ucraina creando un blackout di 6 ore nella regione di Kiev. In questo attacco sono stati in grado di controllare i PLC o gli Remote Terminal Unit (RTU) e hanno aperto gli interruttori per bloccare il flusso di energia elettrica.
La metodologia d'attacco parte dalle email di phishing a persone nell'IT o amministrazione della compagnia per ottenere le credenziali alla VPN. Scaricato ed eseguito il file malevolo (contenente macro) veniva installato il malware **BlackEnergy 3** che apriva la comunicazione con il server C2 da cui gli attaccanti installavano **KillDisk**, un malware necessario per prevenire la ripresa del controllo da parte delle vittime perché cancella il Master Boot Record (MBR). Infine gli attaccanti eseguivano un attacco DoS al call center della compagnia per disturbare la possibilità per gli utenti di segnalare il disservizio.
BlackEnergy conteneva:

- Network scanner
- File stealer
- Password stealer
- Keylogger
- Screenshots
- Network discovery

Gli obiettivi dell'attacco erano in prima fase rubare le credenziali e nella seconda bloccare il sistema.

6.2.2 Industroyer

Un secondo attacco sempre nella regione di Kiev è stato portato a termine utilizzando il malware **Industroyer**.
Il blackout è durato solo 1 ore ed l'attacco è iniziato con una campagna di spearphishing mesi prima.
Industroyer è un malware molto sofisticato creato per interrompere il servizio di un Industrial Control System (ICS) nello specifico quelli utilizzati nelle sottostazioni elettriche. Implementa protocolli di comunicazione utilizzati negli ICS e crea dei DDoS contro i relé di protezione di Simens.
Ha la capacità di inserire backdoor multiple per evitare che scoperta una l'attaccante perda il controllo e permette di modificare il path dei servizi in uso nei registri di Windows rendendo la macchina non bootabile.

6.2.3 NotPetya

Il 27 giugno 2017 molte organizzazioni in giro per il mondo hanno subito l'attacco di un ransomware denominato NotPetya.
Intercettando il traffico verso i server che distribuivano l'aggiornamento di un software per la gestione delle tasse, si veniva reindirizzati ad un server malevolo che faceva scaricare ed installare il ransomware con la conseguenza di ritrovarsi i dati criptati con algoritmo AES 128.
La richiesta di riscatto era di $ 300 in Bitcoin ma era strano il fatto che venisse utilizzato lo stesso indirizzo Bitcoin a tutte le vittime e che venisse richiesto di inviare un'email di conferma dell'avvenuto pagamento.
L'attacco compromise la maggior parte delle macchine in rete locale in sole due ore.
In caso di mancato sistema di comunicazione della chiave non c'è modo di sapere quale chiave di decrittazione inviare alla vittima se si utilizza un id univoco per tutte e quindi si pensa che questo attacco mirasse ad altro e che il ransomware fosse solo una copertura.

7 Tipi di Malware

La parola "Malware" nasce dall'unione delle parole *Malicious* e *Software* e rappresenta un software o un firmware sviluppato per compiere azioni non autorizzate che causeranno un impatto negativo nella confidenzialità, integrità o accesso al sistema di informazioni.
I sistemi possono essere infettati attraverso:

- Accesso diretto al sistema vittima $->$ Dischi infetti, USB ecc...
- Ingegneria sociale
- Phishing, Spear-phishing, Whale-phishing
- Visitando un sito malevolo

7.1 Virus

Sono malware capaci di replicarsi da soli.
Richiedono un'azione umana per essere eseguiti.
Potrebbero infastidire l'utente o fare delle piccole modifiche alle macchine infette.
I software antivirus sono capaci di scovarli.
L'infezione può avvenire attraverso l'apertura di un file (pdf, word) contenente una macro malevola oppure mascherandosi come aggiornamenti di software o sistema.
Sono anche capaci di modificare il loro comportamento a seconda dell'OS.

7.2 Worms

Sono simili ai virus ma non infettano e non richiedono azioni da parte dell'utente.
Possono diffondersi in più dispositivi attraverso la rete.
Di solito sono più pericolosi di un normale virus e colpiscono server sfruttando le falle di configurazione dello stesso.
Il loro obiettivo è avere accesso alla macchina vittima.

7.3 Trojans

Sono software a tutti gli effetti ma vengono utilizzati per avere accesso alla macchina infetta per poi scaricare altri virus attraverso la backdoor aperta (es: TrickBot).
Possono essere utilizzati anche per rubare informazioni personali, file od anche trasformare la macchina vittima in uno zombie.

7.4 Rootkits

Questi malware di solito sono installati direttamente sul kernel e riescono quindi a mascherare le chiamate alle API dell'OS fatte da altri malware.
Permettono anche di avere accesso root alla macchina e vengono utilizzati spesso per mantenere l'accesso alla macchina vittima.
Alcuni non possono essere rimossi quindi l'unità deve essere distrutta.

7.5 Droppers Downloaders

Sono malware che contengono il vero e proprio malware quindi eseguibili con il compito di scaricare il malware una volta eseguiti sulla macchina vittima come ad esempio gli allegati malevoli.
Tipicamente vengono inviati attraverso malspam sotto forma di file Word o Excel.

7.6 Key loggers

In questo caso il malware salva su un file locale tutto ciò che viene scritto dalla tastiera della vittima e poi invia tale file all'attaccante.
L'obiettivo è scovare le credenziali ad account di vario genere o i dati delle carte di credito e a volte la comunicazione con l'attaccante viene criptata.
Tipicamente viene installato assieme ad altri applicativi malevoli per il furto di credenziali oppure inviati come allegato malevolo per email.
Key logger popolari sono **Refog**, **Revealer** e **KidLogger**.

7.7 Bots

I bot vengono utilizzati per creare delle macchine zombie a cui inviare comandi.
In questo modo nascono le botnet che sono utilizzate per eseguire attacchi DDoS o per spedire spam. Le botnet sono comandate da un bot master o più d'uno.
Tra le più note botnet ci sono **Mirai** e **Satori**.

7.8 Cripto Miners

I cripto miner sono malware utilizzati per minare attraverso le macchine infette inviando i profitti al portafoglio dell'attaccante. La maggior parte sono software per il mining open source modificati.
Proliferano grazie a botnet e malspam.

7.9 Ransomware

Semplicemente criptano tutti i file nel sistema in cui vengono eseguiti e poi mostrano un messaggio all'utente in cui spiegano come fare per pagare il riscatto ed ottenere la chiave di decriptazione o il software di decriptazione.
Tipicamente viene accettato il Bitcoin come forma di pagamento per il fatto che è più conosciuto di altre criptovalute.
L'esempio più rilevante è il ransomware **WannaCry** che controllava la possibilità di condividere attraverso SMB, faceva leva sull'exploit **EternalBlue**, installava la backdor **DoublePulsar** ed il ransomware.
I ransomware si dividono in tipologie a seconda del comportamento che hanno e di come agiscono:

- Ransomware: cifrano i file
- Lockers: bloccano solo l'interfaccia
- Master Boot Record: cifrano o modificano l'MBR
- Wipers: cancellano tutti i dati

Si compongono di quattro principali componenti che sono il comportamento simile ai Trojan, la funzionalità di criptare e decriptare i file, il meccanismo di estrazione della chiave ed il modulo in interazione con l'utente.
Per raggiungere la vittima vengono infettati siti web, sfruttare vulnerabilità di OS e software oppure arrivano come allegati malevoli.
I tipi di cifratura possono essere a **chiave pubblica** (la stessa chiave di cifratura viene utilizzata per decifrare) oppure a **chiave asimmetrica** (si cifra con una chiave e si decifra con l'altra corrispondente).
È ovviamente importante che la chiave di cifratura venga eliminata dal sistema vittima in modo da non lasciare traccia.
Alcuni ransomware utilizzano la stessa chiave per ogni dispositivo, altri hanno la chiave scritta direttamente nel codice. In entrambi questi casi è possibile risolvere il problema utilizzando la tecnica di forza bruta fino a trovare la chiave di decriptazione.
Molto importanti sono invece i **Kill Switches** cioè dei sistemi implementati dagli attaccanti per evitare che le loro macchine vengano infettate o permettono di bloccare l'esecuzione del ransomware in modo da evitare che venga infettata la stessa macchina più volte oppure sono errori di codice scoperti dai ricercatori di sicurezza informatica. In ogni caso sono sistemi che si possono utilizzare per disabilitare il malware.
Nel caso di WannaCry, è stato scoperto che il malware faceva richiesta ad un dominio non registrato e se riceveva risposta allora si fermava e non infettava la macchina vittima. Un ricercatore ha quindi acquistato il dominio abilitando così il kill switch.

7.9.1 Bad Rabbit

Uno dei più noti MBR ransomware è **Bad Rabbit**, simile a 6.2.3 si finge un aggiornamento di Flash e si sparge attraverso il protocollo Server Message Block (SMB), un protocollo per la condivisione di file in una rete che permette alle applicazioni di una macchina di leggere e scrivere file.

Scoperto da ricercatori, il malware tenta di scrivere un file sul disco; se la scrittura fallisce allora il processo di criptazione si ferma.
In ogni caso il sistema rimane infetto e potrà contaminare altri dispositivi collegati fisicamente o in rete.

7.9.2 Hidden Tear

È un ransomware scritto in c# creato per motivi di studio ed open source ma viene utilizzato per attacchi reali.
Per approfondire andare al seguente link: `https://github.com/goliate/hidden-tear`.

7.10 Prevenzione dai malware

Per prevenire la ricezione di malware è opportuno adottare delle misure di sicurezza:

- Filtraggio delle email: assieme ad uno filtro anti-spam che in automatico blocca le email malevole e rimuove gli allegati malevoli
- Intercettazione dei proxy: in modo da bloccare i siti riconosciuti come malevoli
- Internet security gateways: sono in grado di ispezionare il contenuto in certi protocolli di comunicazione e scovare malware conosciuti
- Lista di navigazione protetta: all'interno del proprio browser web così da bloccare l'accesso a siti riconosciuti come pullulanti di contenuti malevoli

Per rallentare il diffondersi di malware è necessario:

- Utilizzare la 2FA
- Tenere aggiornati il OS ed i software installati
- Ridurre i privilegi non necessari
- Proteggere l'account degli admin creando un account dedicato per vedere le email ed uno per compiere azioni per cui è necessario avere i permessi d'amministrazione
- Fare formazione dei dipendenti
- Fare spesso un backup multiplo e sicuro disconnesso dalla rete eseguendo una scansione sugli stessi quando è necessario ripristinarli
- Aggiornare regolarmente i prodotti utilizzati per fare il backup

Nel caso in cui il malware abbia già infettato l'organizzazione:

- Scollegare immediatamente i dispositivi infetti
- Spegnere il Wi-Fi disabilitando ogni altra connessione di rete
- Resettare le credenziali d'accesso incluse le password
- Formattare con attenzione i dispositivi infetti e reinstallare il OS
- Controllare che il backup non contenga il malware
- Collegare il dispositivo ad una rete pulita per scaricare, installare ed aggiornare il OS e gli altri software
- Installare, aggiornare ed eseguire il software antivirus
- Ricollegarsi alla propria rete
- Monitorare il traffico di rete ed eseguire scansioni antivirus per identificare eventuali rimasugli di infezione

8 Attacchi alle Password

La determinazione dell'identità è normalmente basata su una combinazione di qualcosa che la persona conosce (es: password), qualcosa che la persona possiede (es: smart card) o in ciò che la persona è (es: impronte digitali).

8.1 Autenticazione basata su Token

L'autenticazione basata su token richiede che l'utente presenti un token per essere autenticato. Il token può essere di diverse tipologie tra cui:

- Codice a barre
- Dispositivi di One Time Password (OTP): il dispositivo ed il server di riferimento sono sincronizzati temporalmente così che il tempo è utilizzato da entrambi come seed per la generazione della OTP e per il controllo.
- Carte magnetiche strisciabili
- Smart Card: i certificati sono salvati nel chip ed il sistema di autenticazione funziona in questo modo:

  1. L'utente inserisce il pin
  2. Il lettore manda una "challenge B"
  3. La smart card genera un valore A e firma A e B con la chiave privata
  4. Il lettore verifica la firma con la chiave pubblica

  Hanno tuttavia lo svantaggio della possibilità di furto e copia dei dati sul chip.

8.2 Autenticazione Biometrica

La parola biometrico si riferisce a tutte le misure utilizzate per identificare unicamente una persona basandosi su un tratto biologico o fisico.
Di norma i sistemi biometrici incorporano un sistema per scansionare o leggere le informazioni biometriche per poi compararle con quelle di persone a cui permettere l'accesso salvate in memoria.
I requisiti per i sistemi di autenticazione biometrica devono essere univoci e non cambiare nel tempo.
Le impronte digitali autenticano con un certo margine d'errore ma esistono anche altri sistemi biometrici:

- Firma
- Impronta digitale
- Scansione della retina/iride
- DNA
- Analisi della firma
- Riconoscimento vocale
- Riconoscimento facciale
- Analisi della camminata

È stato anche proposto l'elettrocardiogramma come sistema biometrico ma ci sono varie problematiche tra cui l'errore di autenticazione del 9%. L'autenticazione

biometrica ha però delle forti limitazioni tra cui l'accuratezza degli algoritmi (falsi positivi che permettono l'accesso a persone non autorizzate e falsi negativi che la vietano a personale legittimo). Inoltre i tratti si possono replicare come le impronte digitali dato che le lasciamo in giro e c'è sempre una fetta di popolazione che non accetterebbe questo tipo di controllo.

8.3 Autenticazione basata su Password

L'autenticazione attraverso password è molto meno costosa ma pur sempre vulnerabile. È il sistema d'autenticazione più utilizzato poiché basta richiedere il nome utente e la password ed il sistema deve comparare la password con quella presente nel database.
Abbinando l'id ad un sistema di permessi è poi possibile garantire l'accesso solo a personale autorizzato.
D'altra parte però il 63% dei data breache è causato da password rubate o molto deboli.
La botnet **Mirai** funziona scansionando continuamente alla ricerca di dispositivi IoT che sono accessibili da internet, provando ad inserire le credenziali di default dei dispositivi o nomi utente e password standard e se riesce ad entrare li infetta con il malware forzandoli a fare riferimento ad un server C2 che li rende dei bot utili per attacchi DDoS.
Il problema è che gli utenti hanno molti account online e spesso utilizzano la stessa password per l'accesso inoltre molti utenti utilizzano password deboli come combinazioni numeriche semplici, caratteri in sequenza nella tastiera e pattern prevedibili.

8.4 Attacchi alle Password

Ci sono quattro tipologie di attacchi per rubare le password:

- Attacchi offline: accedendo al file o al database delle password
- Attacchi online attivi: accedendo da remoto al file o database delle password (exploit, malware, SQLInjection)
- Attacchi online passivi: intercettando il traffico contenente la password (key logger, Man In The Middle (MITM))
- Attacchi non tecnici: utilizzando l'ingegneria sociale

Le password possono quindi essere "crackate" in vari modi tra cui il sistema a forza bruta utilizzato principalmente offline, spiando la vittima mentre la inserisce, cercando le stesse dall'interno, utilizzando un key logger, indovinandola, rubando documenti fisici in cui è scritta, convincendo le vittime a rivelarla, intercettandole attraverso la rete.
Le password nei sistemi operativi sono salvate in file specifici:

- Windows:
    - Macchine locali: SAM database
    - `C:\Windows\System32\config`

    - Montato come HKLM/SAM
- Linux:
    - Macchine locali: `etc/shadow`
- Apple:
    - `/var/db/dslocal/nods/default/users`
    - `user.plist`

e sono salvate con uno user di riferimento, un SID ed un hash. Riguardo gli hash ci sono varie tipologie di hashing tra cui LM che salva password fino a 14 caratteri, nel caso i caratteri della password siano meno vengono inseriti caratteri vuoti, poi la password è divisa in due set da 7 caratteri, criptata e ricomposta e NTLM.
Gli **attacchi di forza bruta** possono essere eseguiti su ricerca esaustiva tentando ogni possibile combinazione di simboli fino ad una preimpostata lunghezza oppure assumendo una lunghezza della password probabile e combinando tutte le lettere maiuscole, minuscole, i numeri e i simboli comuni per un totale di 96 caratteri che per una lunghezza di 8 caratteri della password diventano $96^8 = 7.2 * 10^{24}$ cioè 7.2 quadrilioni di possibilità.
Prima di tentare un attacco del genere conviene tentare con un **attacco a dizionario** cioè un attacco di forza bruta che tenta tutte le combinazioni presenti in un file "dizionario" appunto il quale contiene password comuni e combinazioni probabili.
Gli **attacchi ibridi** uniscono le tipicità del brute force standard con quelle dell'attacco a dizionario prendendo le parole del dizionario e tentando variazioni delle stesse con caratteri speciali e numeri (es: p4ssw0rd, c!@o).
In ogni caso questi attacchi sono lenti e richiedono molto tempo per ottenere una risposta positiva o negativa che sia. Un modo per rendere più veloce il sistema è utilizzando le **Rainbow Tables** cioè tabelle con gli hash di molte password già computati cosi da controllare se quello della password vittima è presente ed ottenere il corrispettivo non computato. Lo svantaggio è che occupano molta memoria.
Il più famoso strumento per il cracking delle password è **John the Ripper**. Supporta molte tecnologie di cifratura comuni su sistemi UNIX e Windows e prevede tre modi di operare: single crack, wordlist e incremental.
Per stimare quanto tempo occorre per crackare una password esistono siti web dedicati in cui inserendo la password in chiaro otteniamo il tempo stimato per crackarla.
Sistemi diversi di valutazione delle password possono ritornare valutazioni opposte perché alcuni sistemi valutano positivamente password che hanno un tot numero di caratteri, lettere maiuscole, minuscole, numeri e caratteri speciali, altri riconoscono la vulnerabilità di impostazioni del genere che portano l'utente a seguire uno standard (maiuscola all'inizio, numero alla fine...) quindi preferiscono password lunghe.

8.5 Possibili Contromisure

Le contromisure per rendere più difficile scoprire le proprie password sono l'aggiungere un numero random, restringere l'accesso ai file delle password solo ad utenti con privilegi e tenere le password cifrate in hash divise dagli ID degli utenti.

Un'attaccante intelligente che prende di mira una persona ricerca possibili affetti, hobby ecc in modo da utilizzare quei nomi per comporre ed indovinare le password dopodiché tenta di crackarla attraverso un dizionario e con password popolari.
Le password policies possono essere completamente controproducenti perché spingono l'utente e utilizzare dei pattern che rendono più facile l'individuazione della password.
Contromisure effettive sono:

- Bloccaggio dell'account dopo un tot di tentativi di accesso (es: 10)
- Tempo fisso prima di riprovare l'inserimento successivamente ad un tentativo errato
- Controllo di attività di login inusuali con l'opzione di blocco dell'account
- Controllo della password scelta. Se già presente nei database di password crackate allora costringi l'utente a scegliere una nuova password
- Utilizzo di 2FA con OTP o sistemi biometrici

Per evitare l'intercettazione utilizzare certificati di comunicazione criptata per evitare spoofing e MITM. Ad esempio nei servizi web va utilizzato il protocollo HTTPS invece del normale HTTP. Infine è importante fare attenzione agli attacchi di ingegneria sociale come il phishing, il shoulder-surfing (spionaggio), il dumpster-diving (ricerca di informazioni nella spazzatura) prendendo le giuste contromisure come la formazione e l'aggiornamento.

9 Identità e Gestione Accessi

Per evitare il problema delle password ripetute e quindi averne una diversa e randomica per ogni account, esistono gli Identity Management Platform (IMP) che si occupano di contenere sia le credenziali che eventuali dati dell'utente ed inserirli automaticamente nelle pagine di login.
Oltre alle password e nomi utente gli IMP contengono anche altri dati come il numero del passaporto e della carta d'identità, l'indirizzo ecc...
In questo modo il sistema verifica le credenziali del soggetto e permette o nega l'accesso ai servizi.

L'utilizzo dei Single-Sign On (SSO) invece permette all'utente di eseguire un unico accesso ad uno dei servizi di un'azienda per poter fruire anche di tutti gli altri senza la necessità di ripetere il login.
Questo sistema viene implementato attraverso token (Gmail $->$ YouTube) e può allargarsi anche a più aziende in che collaborano in una rete di reciproca fiducia (easyJet $->$ Booking.com).
Oltre ad essere una comodità per l'utente, questa tecnologia permette anche di ridurre i costi di mantenimento del servizio di autenticazione per le aziende che così facendo non devono implementare singolarmente una struttura di registrazione, mantenimento dei dati ecc...

9.1 SPID

Anche lo Sistema Pubblico di Identità Digitale (SPID) è un sistema che si basa sullo stesso principio (ha già generato più di 26 milioni di identità digitali). Nel caso specifico dello SPID, possiamo identificare vari responsabili per porzioni differenti del servizio:

- Agenzia per l'Italia Digitale (AgID): l'entità che monitora e autorizza i fornitori di SPID
- Identity Provider: fornitore pubblico o privato che ccertificato da AgID ha il permesso di verificare l'identità dell'utente e rilasciare lo SPID (es: Poste, Aruba...)
- Service Provider: entità pubblica o privata che offre un servizio online
- Attribute Provider: entità che rilascia all'utente l'attributo di qualifica
- Utente: possessore dello SPID che lo utilizza per accedere al servizio online

Lo spid ha tre livelli di sicurezza

- Livello 1: permette di accedere con nome utente e password
- Livello 2: permette di accedere con le credenziali di livello 1 e una OTP generata dall'applicazione per smartphone e tablet
- Livello 3: permette di accedere con i passaggi dei livelli 1 e 2 e l'aggiunta di un dispositivo fisico (es: smart card) garantito dall'identity provider

9.2 SAML

Lo Security Assertion Markup Language (SAML) permette lo SSO e la federazione delle identità fornendo uno standard per la rappresentazione delle asserzioni degli attributi e delle autenticazioni.
Gli SAML Asserting Party o Identity Provider verificano l'identità di un utente e rilasciano un'asserzione di autenticazione.
L'utente può presentare al provider del servizio l'asserzione dell'autenticazione senza dover ripetere l'autenticazione.
In poche parole è un protocollo Extensible Markup Language (XML) con cui gli identity provider si scambiano informazioni ad esempio Dropbox che permette l'autenticazione attraverso Google o **Shibboleth**, un consorzio che permette alle università di condividere risorse e ricerche attraverso confini istituzionali con un protocollo simile allo SAML. Le asserzioni dello SAML sono tre:

- Authentication statement: descrive lo scopo utilizzato per l'autenticazione del soggetto
- Attribute statement: elenca gli attributi posseduti dal soggetto
- Authorization statement: definisce i permessi del soggetto

Gli elementi comuni in un'asserzione sono:

- Emittente e timestamp di emissione
- ID dell'asserzione
- Soggetto (nome e dominio di sicurezza)
- Condizioni per le quali l'asserzione è valida (condizioni come la validità per un periodo di tempo)

10 Controllo degli Accessi

L'elemento centrale della sicurezza informatica è il controllo degli accessi che previene un accesso non autorizzato alle risorse ed anche l'utilizzo in modo non autorizzato di una risorsa.
Ovviamente coinvolge l'autenticazione al sistema e l'assegnazione dei permessi di accesso a certe risorse del sistema.
I principi del controllo degli accessi sono l'**Autenticazione** che verifica l'identità rivendicata da o per un'entità di sistema, l'**Autorizzazione** che garantisce i permessi ad un'entità di sistema di accedere alle risorse e l'**Accountability**/Responsabilità che monitora e processa l'accesso degli utenti alle risorse.
Il sistema di controllo degli accessi si sviluppa attraverso:

- Policies: coinvolgono i soggetti cioè le entità che possono accedere all'oggetto, gli oggetti cioè la risorsa ad accesso controllato (file, cartelle...) e il diritto di accesso (o permessi) cioè il modo in cui i soggetti accedono all'oggetto (scrittura, lettura...).
- Modelli:
  - Mandatory Access Control (MAC) $->$ accesso basato sull'etichetta di sicurezza dell'oggetto e sul livello di autorizzazione del soggetto; utilizzato principalmente in ambito militare
  - Discretionary Access Control (DAC) $->$ accesso basato sull'identità del soggetto.
  - Role Based Access Control (RBAC) $->$ accesso basato sul ruolo impersonato dal soggetto; il più utilizzato.
  - Attribute Based Access Control (ABAC) $->$ accesso basato sugli attributi del soggetto, dell'oggetto e del contesto; molto utilizzato
  - In alcuni casi come per esempio quello famoso di **Edward Snowden**, i dipendenti possono abusare dei loro permessi per compiere azioni illecite come la fuga di notizie o la vendita di informazioni riservate. Ci sono poi casi in cui la specificità della situazione non permette l'implementazione dei modelli sopra descritti (es: accesso IoT ad abitazione per baby sitter, impresa delle pulizie, abitanti, amici...)
- Meccanismi

10.1 Modello DAC

Prevede regole esplicite che stabiliscono chi può o non può eseguire quali azioni e su quali risorse. L'amministratore può modificare/revocare i permessi degli utenti. Implementata spesso come **matrice degli accessi** che tuttavia è molto dispendiosa a livello di memoria (per ogni soggetto, per ogni risorsa, il soggetto ha il permesso di...).
Esistono anche altri sistemi come la **lista del controllo accessi** in cui per ogni risorsa si ha la lista di chi può accedere e come ma ha lo svantaggio di dover scorrere tutte le liste in caso di eliminazione di un utente per rimuovere tutti i permessi oppure la **capability list** cioè per ogni utente si ha la lista delle risorse che può

vedere e di quali sono i permessi per ognuna tuttavia è difficile avere un'anteprima dei permessi garantiti ad un oggetto.
La gestione delle policy è un compito complesso in sistemi elaborati ed è difficile avere la visione generale dei permessi concessi agli utenti per le risorse.

10.2 Modello RBAC

Questo sistema si divide esso stesso in tre tipologie differenti:

- $RBAC_0$: l'RBAC di default
- $RBAC_1$: somma tra $RBAC_0$ e la gerarchia dei ruoli cioè un sottoruolo eredita i permessi del sopraruolo
- $RBAC_2$: $RBAC_1$ + vincoli che possono essere statici Static Separation of Duty (SSD) cioè un utente non può essere assegnato a più di n ruoli nell'insieme dei ruoli o dinamici Dynamic Separation of Duty (DSD) cioè un utente non può attivare più di n ruoli in un insieme di ruoli nella stessa sessione)

Nel caso del $RBAC_2$ è da considerare qual è il ruolo di un utente nel momento specifico "session_role" (es: un professore non può collegarsi contemporaneamente come studente). Il vantaggio del modello RBAC è l'efficienza nell'amministrare e monitorare i permessi poiché l'assegnamento dei permessi non è manuale ma automatico con l'assegnazione del ruolo è tuttavia difficile implementare correttamente il controllo degli accessi quando ci sono migliaia di ruoli che cambiano spesso e si è notato che dal 50% al 90% degli impiegati in grandi organizzazioni hanno più permessi di ciò che gli compete.

10.3 Modello ABAC

Il modello ABAC permette o impedisce al soggetto di compiere un'operazione in base agli attributi assegnati al soggetto, all'oggetto ed alle condizioni specifiche del momento oltre che dall'insieme di regole che sono impostate sulla relazione tra attributi e condizioni.
Ad esempio i medici del reparto di cardiologia possono visualizzare i record dei pazienti cardiopatici utilizzando il computer dell'ospedale all'interno del reparto ospedaliero.

10.4 Standard XACML

L'eXtensible Access Control Markup Language (XACML) è uno standard OASIS[9] che fornisce un linguaggio per la definizione di policy basato su XML, un linguaggio di richiesta e risposta basato su XML, tipi di dati, funzioni e algoritmi di combinazione standard, un'architettura che definisce la maggior parte dei componenti in un'implementazione, un profilo di privacy ed un profilo RBAC.

- 1) Il **policy enforcement point** è l'entità che protegge le risorse (file, cartelle ecc) e che conduce i controlli d'accesso prendendo le decisioni sulle richieste, rinforzando le decisioni di autorizzazione ed eseguendo gli obblighi

[9]https://www.oasis-open.org

- 2) Il **policy decision point** riceve ed esamina le richieste, recupera le policies applicabili, valuta le policy applicabili e risponde la decisione presa
- 3) Il **policy administration point** crea le policies di sicurezza e le archivia nel repository
- 4) Il **context handler** converte le richieste da formato nativo a XACML e converte le decisioni sulle autorizzazioni dal formato XACML a quello nativo
- 5) Il **policy information point** rappresenta il sorgente dei valori degli attributi o i dati richiesti per la valutazione delle policies

Il flusso di dati di questo modello inizia con 3 che scrive le policies e gli insiemi di policy rendendoli accessibili a $2 ->$ il richiedente l'accesso manda una richiesta a $1 ->$ che la inoltra a 4 il quale la traduce in XACML e $->$ la invia a 2 il quale richiede ogni attributo aggiuntivo $->$ a 4 che li richiede a $->$ 5 il quale recupera gli attributi e $->$ li ritorna a 4.
Il sistema è applicabile sia in rete locale che in cloud.
I componenti chiave del linguaggio XACML sono:

- $< PolicySet >$: è la chiave di alto livello dell'elemento che aggrega altri elementi PolicySet o Policy
- $< Policy >$: elemento composto principalemnte da $< Target >$, $< Rule >$ e $< Obligation >$ ed è valutato dal policy decision point per produrre ed accedere alla decisione
- $< Rule >$: elemento che fornisce le condizioni che testano gli attributi rilevanti all'interno della $< Policy >$
- $< Target >$: elemento utilizzato per abbinare le risorse richieste con una Policy applicabile. Utilizza gli elementi $< AnyOf >$, $< AllOf >$ e $< Match >$
- **Algoritmi di Combinazione**: utilizzati per riconciliare più risultati delle valutazioni delle policies in un'unica decisione. Il risultato in caso di algoritmi "First-applicable" è la prima regola/policy il quale target è valutato `True` .

10.5 OAuth

Il protocollo OAuth è un protocollo di autorizzazione standard che permette ad applicazioni di terze parti di proteggere risorse ospitate su server HTTP. Richiede un **access token** da un server di autorizzazione così che le applicazioni di terze parti possano usarlo per richiedere un accesso alle risorse riservate.
Gli attori in gioco in questo modello sono:

- Resource Owner: entità che garantisce l'accesso ad una risorsa riservata (es: Mario)
- Resource Server: il server in cui sono archiviate le risorse richieste dal proprietario (es: Facebook)
- Authorization Server: il server che distribuisce il token di accesso al cliente dopo aver autenticato il proprietario della risorsa e ottenuto la sua autorizzazione (es: Facebook)
- Client: un'applicazione di terze parti che richiede accesso a risorse riservate a nome del proprietario e con il suo permesso (es: Spotify)

Ci sono tre tipi differenti di autenticazione attraverso OAuth:

- **Authorization Code Grant Flow**: questo è il sistema applicato in tutti quei contesti in cui chi gestisce l'autenticazione differisce da chi gestisce il servizio (es: utente $->$ autenticazione con Google $->$ accesso a Spotify). Chi gestisce il servizio può essere un'applicazione web piuttosto che un un'applicazione nativa (app Android) o basata interamente su browser
- **Resource Owner Password Grant Flow**: in questo caso invece è lo stesso gestore a servire sia l'autenticazione che il servizio (es: utente $->$ login Google $->$ Gmail e GApps)
- **Client Credential Grant FLow**: questo sistema permette ad applicazioni di eseguire in autonomia l'autenticazione in vece dell'utente (es: Google Doc per accedere ai file deve autenticarsi su Google Cloud Storage da cui poter recuperare i file dell'utente). Il processo è sempre legato all'access token ma in questo caso di tipo applicativo e non utente.

11 Introduzione alla Privacy

Sul TIME Zuckerberg ha detto che "la privacy è morta".
Effettivamente paragonando il mondo offline con quello online le nostre abitudini sono cambiate molto:

Tabella 1: Tabella comparativa azioni Offline ed Online.

Offline	**Online**
Conversazioni faccia a faccia	Messaggistica istantanea
Lettere postali	Email
Archivi cartacei	Cloud
Pagamenti in contanti	Carte di credito
Seguire le persone fisicamente	Tracciamento della posizione
Conoscere i propri amici	Social network
Cercare informazioni nei libri	Ricerca Google

Le informazioni sono difficili e costose da ottenere, mantenere e ricercare nel mondo offline.
Il mondo online ha permesso di introdurre il concetto di "Surveillance Capitalism" - Shoshama Zuboff.
Emergono però nuove problematiche tra cui il furto di dati, cosa che accade costantemente ogni giorno.
Nel 2021 ha fatto scalpore il **Data Breach di Facebook** che ha rivelato le informazioni di 1.5 miliardi di utenti (circa la metà degli utenti di Facebook) tra cui nomi, cognomi, email, indirizzo ecc... Nel 2013 il caso di **Edward Snowden** ha messo in risalto l'aspetto legato alla sorveglianza attiva di massa dunque non più solo la profilazione per scopi di marketing.
A causa di tutto ciò gli organismi politici internazionali stanno applicando nuove regole per la gestione e trasmissione dei dati come per esempio il General Data Protection Regulation (GDPR), il protocollo di tracciamento senza dati dell'utente

tramite Bluetooth Decentralized Privacy-Preserving Proximity Tracing (DP3T) così come nascono nuovi progetti tra i quali la **rete Tor** per l'anonimato ed il **Privacy Badger** che previene l'online tracking.
Il tentativo di creare delle leggi per il trattamento dei dati personali ha fatto emergere il problema della definizione di che cos'è un dato personele. Per quanto riguarda il GDPR ogni informazione relativa ad una persona identificata o identificabile direttamente o indirettamente è un dato personale la tal persona invece è definita **Data Subject**. Quindi anche l'indirizzo IP è considerato dato personale o qualsiasi numero identificativo.
Un'altra definizione necessaria è quella del controllore dei dati **Data Controller** cioè la persona o figura legale, autorità pubblica, agenzia o qualsiasi altro corpo che singolarmente o in gruppo con altri determina i motivi della lavorazione dei dati personali.

Il significato di privacy è cambiato durante il tempo ed essendo un concetto soggettivo è stato definito in vari modi durante il tempo Quanto detto da Westin

Tabella 2: Definizioni di privacy.

Anno	**Autore**	**Definizione**
1890	Warren & Brandeis	Il diritto di essere lasciati da soli
1970	Westin	Il diritto dell'individuo di decidere che informazioni riguardanti sè stesso dovrebbero essere comunicate agli altri ed in quale circostanza
1970	Solove	La tassonomia della privacy è nociva
2001	Agre & Rotenberg	La libertà alla costruzione della propria identità da vincoli irragionevoli
2004	Nissenbaum	Privacy come Integrità Contestuale
2018	GDPR	Trasparenza, scopo, proporzionalità, responsabilità

nel 1970 è alla base delle regole di privacy attuali, Agre & Rotenberg invece hanno un approccio più psicologico sul come un utente si esprime sapendo di essere ascoltato, Nissenbaum vede l'utente come una persona che condivide informazioni a seconda del contesto (es: certe informazioni le dici al medico ma non le pubblicheresti sul web), infine il GDPR considera il problema anche dal punto di vista di chi raccoglie i dati il quale ha il compito di dire come, per quale ragione e se li condividerà con terzi e con questo tenere traccia di come vengono raccolti, utilizzati e trasmessi a terzi.

11.1 Proprietà della Privacy

Ci sono due macro possibilità su come i dati vengono forniti e dunque su come agiscono i protagonisti del sistema in oggetto.
La **Hard Privacy** prevede la minimizzazione dei dati; il soggetto fornisce meno dati possibili. Riduce il più possibile la necessità di fiducia con altre entità. I dati infatti potrebbero già essere cifrati durante l'invio al Data Controller.
La **Soft Privacy** si verifica quando il Data Subject ha già perso il controllo dei suoi

dati per cui è molto difficile per il Data Subject verificare come i suoi dati sono archiviati e processati. Ciò accade quando l'utente si fida del Data Controller e quindi è il Data Controller che deve tutelare i dati dell'individuo. In questo modo diventa più difficile per l'individuo avere controllo di come i propri dati vengono processati.
L'**Anonimity** è definita da Pfitzmann come l'incapacità per un attaccante di identificare il soggetto all'interno di un insieme di soggetti detto anonimity set. Questo in linea di principio si ottiene nascondendo il collegamento tra l'identità e l'azione o pezzo d'informazione.
Un modo per ottenere l'anonimato è l'utilizzo di uno **Pseudonimo** tuttavia tendenzialmente l'utente utilizza sempre lo stesso pseudonimo ovunque e questo si traduce in tracciabilità dell'utente fintanto alla sua individuazione.
L'**Unlinkability** o Scollegamento si verifica quando un attaccante non riesce a distinguere se due o più oggetti di interesse sono in relazione tra loro o no (es: due o più email). Anche in questo caso nascondendo il collegamento tra le azioni, le identità e i pezzi di informazione è la soluzione.
L'**Irrintracciabilità** si verifica quando l'attaccante non riesce a distinguere se un oggetto appartiene o meno ad un insieme di oggetti.
Una proprietà molto importante è anche la **Plausible Deniability** cioè l'impossibilità di provare che un utente conosce, ha fatto o detto qualcosa (es: utilizzato per il voto online).
La **Confidenzialità** è la proprietà di conservazione delle restrizioni autorizzate all'accesso e alla divulgazione delle informazioni, compresi i mezzi per proteggere la privacy personale e le informazioni proprietarie.
La **Compliance** è legata alla legislazione sulla protezione dei dati; la GDPR specifica i principi per la gestione dei dati personali all'interno dell'Unione Europea.
Infine l'**Awareness** è la proprietà per la quale l'utente dovrebbe essere messo a conoscenza delle conseguenze alla condivisione delle proprie informazioni.

11.2 Minacce alla Privacy

Solove ha definito quattro tipi di azioni che possono essere legate al concetto di privacy:

- Information collection: sorvegliare l'utente guardando, ascoltando e registrando audio, video. Questionari con domande inappropriate
- Invasion: intrusione nella vita di una persona (giochi AR che direzionano l'utente in luoghi privati). Inferenza nelle decisioni di una persona
- Information processing: aggregazione, insicurezza, identificazione, uso secondario, esclusione
- Information dissemination: rompere la confidenzialità, pubblicazione di dati privati, amplificazione dell'accessibilità ai dati di una persona, ricattare per la cancellazione dei dati, utilizzare l'identità altrui per coinvolgere altri nel prodotto, pubblicare false informazioni di una persona

La sorveglianza è un concetto che è spesso sottovalutato se si considera che per esempio uno smart meter/hub casalingo rappresenta un modello di sorveglianza molto alto perché può fornire dati sui consumi, su quando la persona si sta lavando o sta

cucinando, quando è in casa e quando no.
Il caso delle smart tv che spiano l'utenza di fatto è un evento realmente accaduto che ha comportato una sanzione di $ 2.2 milioni all'azienda produttrice che registrava dati quali la preferenza sui contenuti guardati per vendere pubblicità mirata.
Angry Birds invece è stato preso di mira dalla National Security Agency (NSA) e dal Government Communications Headquarters (GCHQ) per ottenere i dati degli utenti.
Il **probing** è il sistema di furto di informazioni adottato nelle campagne di phishing.
Per quanto riguarda l'information processing, un caso di aggregazione dei dati è quello di **Target** che è riuscito a sapere prima della ragazza che la stessa era incinta consigliando l'acquisto di prodotti per donne in maternità.
Casi di identificazione invece sono quelli in cui dai click fatti su di un sito è possibile determinare l'identità di un individuo mentre un caso eclatante di uso secondario è quello di **Cambridge Analytica** che con i dati di profili Facebook creava dei profili psicologici delle persone per spingerle a votare a favore di un candidato in particolare.
Per quanto riguarda l'information processing alcuni casi sono il breach della confidenzialità accaduto a **Equifax**, azienda di recupero crediti che è stata hackerata con conseguente furto di dati sensibili per la predizione della capacità di pagamento del debito dei clienti, oppure il caso di exposure che ha coinvolto molte celebrità cadute vittime di campagne di phishing. Casi di appropriazione sono stati riscontrati nei social network dove criminali rubano l'identità delle persona per sfruttarla in altri social network e compiere truffe e casi di disseminazione sono per esempio i ransomware con doppia richiesta di riscatto (decriptare i dati e non renderli pubblici). La distorsione riguarda più i cosiddetti troll che in alcuni casi sono anche stati condannati a pene pecuniarie. Casi di invasion, nello specifico intrusion si sono registrati con frequenza grazie ai social network dove alcuni utenti hanno sfruttato le informazioni sulla geo-localizzazione di altri per atti di stalking ed anche crimini peggiori. Anche il cyberbullismo si alimenta di queste informazioni.

11.3 Privacy Enhancing Technologies (PETS)

Riguarda tutti gli strumenti, meccanismi e architetture che ambiscono a mitigare la preoccupazione sulla privacy pur permettendo agli utenti di gioire dei benefici delle tecnologie moderne.
PETS può essere applicata alle comunicazioni o a database esistenti, sia da utenti individuali che da organizzazioni.
In poche parole sono quelle tecnologie e comportamenti a protezione della privacy.
Quando il Data Controller è ritenuto affidabile è possibile applicare alcune tecnologie per la protezione dei dati tra cui:

- Criptazione dei dati sia in trasmissione che in archiviazione
- Autenticazione e autorizzazione di dipendenti che gestiscono dati personali
- Login sicuro per l'accesso ai dati
- Cancellazione sicura dei dati (diritto all'oblio)
- Controllo degli accessi basato sullo scopo

In questo caso il Data Controller si protegge da attacchi di terzi ma i dati rimangono vulnerabili se è il Data Controller stesso a decidere di utilizzarli in modalità non etiche o illegali.

Tutte le tecnologie che permettono all'utente di scegliere se, come e in quale circostanza divulgare i propri dati personali ricadono nel macro gruppo definito **User Awareness Technologies**. Esse aiutano l'utente a fare scelte consapevoli riguardanti la protezione della propria privacy.
Alcuni esempi di queste tecnologie sono:

- Privacy friendly defaults: settaggi della privacy di default impostati a privato
- Impostazioni della privacy modificabili, feedback contestuali
- Interfacce per l'esercizio sui diritti di accesso dell'utente
- Privacy policies chiare, concise e comprensibili
- Privacy nudges: sistemi di consapevolezza (es: app che mostrano all'utente quante volte è stata condivisa con app terze la loro posizione)

Privacy Bird è un esempio di tecnologia di questo tipo che mostra se le privacy policy di un sito sono in linea con le preferenze impostate dall'utente.

Le **Anonimity Technologies** sono tutte quelle tecnologie che assicurano l'anonimato dell'utente come ad esempio:

- Anonimizzazione dei database: k-anonimity, l-diversity, t-closeness
- Comunicazioni anonime: Mixnet, Onion routing, Tor
- Credenziali anonime: Idemix (IBM) è un sistema che prova al provider di avere determinati attributi senza rivelare la propria identità. Il provider dunque non conosce chi è l'utente ma riceve solo la conferma che l'utente ha determinati attributi.

Esistono poi altre tecnologie che stanno venendo sviluppate per migliorare la privacy come gli archivi privati remoti, ricerca per parole chiave di file cifrati su cloud (quando i file sono salvati vengono etichettati per parole chiave), computazione dei dati che preserva la privacy.

12 Protezione dei Dati

Nel 1995 in Europa ogni stato decideva regole sulla privacy per conto proprio creando così difficile implementare regole trasversali per siti web visibili ovunque. L'European data protection board ha imposto ad ogni nazione di creare un proprio organo di sorveglianza sulla tutela della privacy e nel 2018 la GDPR ha permesso di uniformare i diritti dell'utente, il concetto di dato personale, le regole sul tracciamento dell'utilizzo dei dati e multe salate per i trasgressori. Dall'articolo 82 della GDPR "Ogni controllore coinvolto nel processo deve essere responsabile per i danni causati dalla loro elaborazione nel caso in cui infranga questo Regolamento. L'elaboratore deve essere responsabile per i danni causati dall'elaborazione solo quando non ha ottemperato agli obblighi che questo Regolamento prevede relativamente all'elaborazione o nel caso in cui abbia agito contrariamente alle istruzioni della legge o del controllore". E ancora "Quando più di un controllore o elaboratore o entrambi i

controllori ed elaboratori sono coinvolti nella stessa elaborazione e nell'archiviazione, visti i paragrafi 2 e 3, sono responsabili per ogni danno causato dall'elaborazione, ogni controllore o elaboratore deve essere ritenuto responsabile per l'intero danno in modo da assicurare l'effettivo compenso al soggetto dei dati". Per **Dati Personali** con la GDPR si intendono anche dati genetici, biometrici e di posizione geografica. I dati però possono essere considerati personali in alcuni contesti e non personali in altri.
Nel caso in cui un sito web registri l'indirizzo IP degli utenti per identificare e reagire ad attacchi l'indirizzo IP è considerato un dato personale.
Mario Rossi è un dato personale? Non sempre perché è un nome molto comune.
L'uomo alto di mezza età che possiede un Labrador, guida una Fiat Punto e vive al numero 15 è considerabile un dato personale perché permette di distinguerlo dalla massa.
Dunque anche l'indirizzo può essere considerato in alcuni casi un dato personale perché ad esempio nelle Pagine Gialle identifica una persona.
In sostanza gli stessi dati possono essere considerati dati personali in un caso e dati non personali in un altro. Quando un'informazione è legata ad un individuo allora si è in presenza di un dato personale.

12.1 Doveri nel trattamento dei dati

Sia il data controller che il data processor hanno dei doveri per quanto riguarda la raccolta, l'elaborazione e l'archiviazione dei dati:

1. Devono avere una base legale per elaborare i dati
2. Devono elaborare i dati per motivi dichiarati e specifici
3. Devono collezionare solo i dati necessari per lo scopo
4. Devono tenere i dati solo per il tempo necessario
5. Devono tenere solo dati accurati
6. Devono tenere i dati al sicuro
7. Devono permettere ai data subject di esercitare i propri diritti
8. Devono mantenere un registro delle attività di elaborazione

Sei possibili basi legali per elaborare i dati sono:

- Consenso: i data subject hanno dato il loro permesso all'elaborazione dei dati personali per una o più motivazioni
- Contratto: l'elaborazione è necessaria per la prestazione di un contratto in cui il data subject è parte o per rispondere alla richiesta del data subject prima che stipuli il contratto
- Obblighi legali: l'elaborazione è necessaria per proteggere l'interesse del data subject o di altre persone fisiche
- Interesse vitale: l'elaborazione è necessaria per proteggere l'interesse vitale del data subject o di altre persone fisiche
- Interesse pubblico: l'elaborazione è necessaria per la prestazione del servizio di pubblico interesse o nell'esercizio di autorità ufficiali investite dal data controller

- Interesse legittimato: l'elaborazione è necessaria per lo scopo di interesse legittimato perseguito dal data controller o da parti terze ad eccezion fatta per quegli interessi che sono sovrascritti da interessi o diritti fondamentali di libertà del data subject che richiedono la protezione per dati personali in particolare quando il data subject è un bambino

Per **consenso** del data subject si intende liberamente concesso (non dovrebbe esserci una precondizione per registrarsi ad un servizio), specifico (deve essere richiesto il consenso per ogni motivo di elaborazione e attività condotta sui dati), informato (spiegato in modo chiaro e conciso) e non indicato ambiguamente (silenzio assenso, checkboxs pre-abilitati o inattività) attraverso una chiara un'azione affermativa.
Per **Trasparenza** si intende che il data controller deve informare gli utenti su come vengono elaborati i loro dati in modo semplice da capire. Quando i dati sono collezionati il data controller deve fornire una notifica di privacy con i dettagli del caso.
Ci sono anche delle limitazione delle finalità. I dati personali devono essere collezionati per specifiche, esplicite e legittime finalità e nessun'altra elaborazione può essere fatta per filanità incompatibili con quelle dichiarate. Finalità compatibili sono archiviazione per interesse pubblico, scientifico, statistico o di ricerca storica.
La minimizzazione dei dati è il concetto che prevede che i dati siano adeguati, rilevanti e limitati a ciò che è necessario per fornire il servizio ed il data controller deve accertarsi che lo siano.
I dati personali devono essere accurati e tenuti aggiornati se necessario. Vanno prese tutte le misure del caso per eliminare i dati non accurati o rettificarli senza ritardi.
I dati personali vanno archiviati in modo da permettere l'identificazione del data subject non oltre il necessario per la fornitura del servizio. Possono tuttavia essere archiviati più a lungo solo per motivi di interesse pubblico, scientifico, statistico o di ricerca storica.
I dati personali devono essere elaborati attraverso modalità che assicurano adeguata sicurezza compresa la protezione da elaborazioni non autorizzate o illegali e la perdita accidentale, il danneggiamento o la distruzione utilizzando appropriate misure tecniche e organizzative.
Il data controller deve essere in grado di dimostrare di attenersi agli obblighi della GDPR e per fare ciò dotarsi di sistemi che permettono l'accountability. Inoltre deve rispettare tutto ciò detto sopra per esempio adottando un approccio di tipo data protection by design and default e incaricando un data protection officer.

12.2 Diritti del Data Subject

Il data subject ha un certo numero di diritti:

1. Diritto di essere informato
2. Diritto di accedere ai propri dati
3. Diritto di rettificare i propri dati
4. Diritto di far eliminare i propri dati se elaborati in modo non legalmente conforme
5. Diritto di obiettare a certe elaborazioni se non basate sul consenso

6. Diritto di non essere soggetto a decisioni completamente automatizzate
7. Diritto di richiedere i propri dati in formato consultabile

La GDPR richiede anche sistemi e metodologie per denunciare eventuali violazioni ai dati personali; tutte le organizzazioni devono denunciare determinati personal data breaches all'autorità si supervisione relativa. Questo deve essere fatto nelle 72 ore successive all'essere venuti a conoscenza del breach per quando possibile.
Se il breach è ad alto rischio devono anche essere informate gli individui a rischio senza alcun ritardo.
Deve esserci in ogni caso un sistema robusto per scovare eventuali breach, investigarli e riportarli attraverso procedure dedicate in modo da rendere più facile il processo decisionale.
Deve anche essere mantenuto un registro di tutti i personal data breach a prescindere che richiedano la notifica agli enti preposti.

La GDPR prevede due livelli di sanzioni:

1. Infrazioni non gravi: possono risultare in una sanzione massima di €10 milioni o del 2% dei guadagni mondiali annuali (es: non aver riportato il data breach alla data protection authority, non aver incaricato un data protection officer, aver elaborato i dati illegalmente)
2. Infrazioni gravi: possono risultare in una sanzione massima di €20 milioni o del 4% dei guadagni mondiali annuali (es: violazione dei principi della protezione dei dati, violazione dei diritti del data subject, trasferimento di dati personali ad organizzazioni internazionali o a paesi terzi)

13 Privacy per Design

Ci sono 7 principi fondamentali:

1. Proactive not Reactive: prevenire non rimediare
2. Privacy come impostazione di default
3. Privacy incorporata nel design
4. Funzionalità complete
5. Sicurezza end-to-end e protezione a vita
6. Trasparenza
7. Rispetto per la privacy degli utenti

13.1 Metodologia LINDDUN

La metodologia Linkability Identifiability Non-Repudiation Detectability Disclosure of Information Unawareness Non-Compliance (LINDDUN), basata sulla conoscenza, nasce con lo scopo di aiutare gli ingegneri del software con poche esperienze di privacy ad introdurre il concetto di privacy all'inizio dello sviluppo.
La prima parte corrisponde alla creazione del diagramma del flusso di dati e alla descrizione di tutti i dati, poi vengono mappate le minacce ed i rischi al diagramma di flusso dei dati identificandole attraverso l'albero delle minacce, infine vengono

gestite le minacce. LINDDUN aiuta nella fase di riconoscimento delle minacce dando supporto documentale nel mappare le tabelle e con la tassonomia delle minacce così come nella fase di gestione delle minacce fornendo documentazione sulla tassonomia delle strategie di mitigazione delle minacce e sulla classificazione delle soluzioni per la privacy.
Ognuna delle parole che compongono l'acronimo LINDDUN corrisponde ad una minaccia da considerare e gestire. La stesura di card che considerano le minacce permette di riassumere il tutto in tabella e successivamente di esportare la tabella come grafico tenendo presente le relazioni di fiducia tra elementi se si assume che si comportino come ci si aspetta.
Gli step da compiere sono dunque:

1. Creare un modello del sistema
    - Entità $->$ rettangolo
    - Processo $->$ ellisse
    - Archivio di dati $->$ parole tra due linee orizzontali
    - Flusso di dati $->$ freccia che punta dalla sorgente al ricevente
    - Confine di fiducia $->$ rettangolo tratteggiato che racchiude vare entità, processi ecc...
2. Mappare il diagramma di flusso dei dati degli elementi come categorie LINDDUN $->$ tabella con le "X" in caso di minaccia
3. Ricavare e documentare le minacce ad esempio usando una struttura ad albero
4. Assegnare una priorità alle minacce riscontrate calcolando un punteggio dato da probabilità ed impatto
5. Documentare una strategia di mitigazione per tutti i casi di rischio alto o critico utilizzando la tabella comparativa LINDDUN
6. Scegliere soluzioni avanzate per la tutela della privacy

14 Anonimizzazione dei Dati

Una delle domande più difficili a cui rispondere è come elaborare e realizzare funzioni di un dataset contenente informaizioni personali e sensibili proteggendo tuttavia la privacy individuale?
La domanda nasce dal fatto che i dati in possesso di aziende come Google, Facebook ecc... vengono utilizzati anche da altre aziende non solo private ma anche governative per scopi di analisi ecc...
È quindi importante che questi dati siano resi anonimi.

Per provare a rispondere a tale domanda è bene classificare gli attributi come segue:

- Identificatori espliciti: identificano un utente come il nome, il cognome, il numero di passaporto...
- Quasi-identificatori: data di nascita, età, codice postale, numero di telefono...
- Attrivuti sensibili: malattie, salari... Sono anche quegli attributi di cui i ricercatori hanno bisogno e che quindi sono sempre rilasciati direttamente

Un sistema per la protezione degli identificatori espliciti è la **Tokenizzazione** cioè generare un token unico per ogni riferimento oppure adottare la **Sostituzione** sostituendo randomicamente un attrivuto con attributi inventati (nomi a caso).
Questi sistemi purtroppo non sono abbastanza per proteggere i dati poiché filtrando i dati attraverso gli altri attributi oppure collegando diversi dataset è possibile comunque risalite all'identificatore esplicito.

14.1 K-Anonimity

In caso di attributi quasi-identificatori, un dato deve essere indistinguibile da almeno $k-1$ altri dati.
Inoltre, sempre per quanto riguarda gli attributi quasi-identificatori, ogni classe di equivalenza deve contenere almeno k dati che hanno lo stesso valore.
Per ottenere questi due principi vengono adottate due tecniche, la **Generalizzazione** e la **Soppressione**.
Nello specifico la prima rimpiazza quasi-identificatori specifici con valori meno specifici (es: l'età invece di essere 23 è $2*$ oppure ≥ 20) finché non ci sono almeno k valori identici.
La seconda invece viene adottata quando la generalizzazione provoca troppa perdita di informazioni e non fa altro che rimuovere un'intera colonna che non è utile allo scopo statistico (es: la religione o il nome).
Inoltre ci sono altri algoritmi che possono essere applicati per raggiungere lo scopo ma rimane il problema che questo sistema provoca errori statistici perché i casi limite vengono generalizzati per ottenere l'offuscamento.
Di fatto più si generalizza, più si perdono dettagli.
La k-anonimity non garantisce la privacy se i valori sensibili in una classe di equivalenza mancano di diversità (**homogeneity attack**: può accadere quando l'attributo sensitivo è uguale per tutti i k-record nella classe di equivalenza e quindi a prescindere da chi sia il soggetto in ogni caso ha quel valore di attributo) oppure quando l'attaccante ha delle conoscenze di base (**background attack**: sapendo che una nazionalità ha un tatto di affetti da una patologia, si deduce per probabilità che il soggetto è colui che ha quella patologia conoscendo la sua nazionalità come dato di partenza).

14.2 L-Diversity

L'l-diversity richiede che ogni classe di equivalenza abbia abbastanza valori sensibili diversi ma anche che la distribuzione delle differenze dei valori sensibili sia uniforme.
L'entropia della distribuzione dei valori sensibili in ogni classe di equivalenza deve essere almeno $\log(l)$:

$$Entropia(E) = -\sum_{s \in S} p(E,s) \log p(E,s)$$

Cioè la somma per tutti i valori che l'attributo sensitivo può assumere (S) della frazione dei record della classe di equivalenza (s) per il logaritmo della stessa funzione.
Questo sistema rappresenta una miglioria ma presenta comunque delle problematiche poiché è suscettibile agli **attacchi di inferenza probabilistica**. Se l'attaccante

conosce alcuni attributi del soggetto potrebbe risalire allo stesso attraverso tali attributi. Inoltre la l-diversity non considera la distribuzione generale dei valori sensibili quindi in un caso d'esempio in cui gli attributi sensibili siano HIV+ (1%) e HIV- (99%), la l-diversity non fa differenza tra le due classi di equivalenza generando una possibile violazione della privacy.

14.3 T-Closeness

Il sistema t-closeness permette di risolvere i problemi di inferenza visti in sezione 14.2 poiché le distribuzioni degli attributi sensibili di ogni gruppo di quasi-identificatori deve essere simile alla loro distribuzione nel database originale.

Ci si accorge così che a prescindere dal sistema di anonimizzazione utilizzato, è possibile risalire ad un soggetto attraverso gli attributi quasi-identificatori.
Ribaltando quindi l'idea di database potremmo pensare ad implementare un sistema di query che fornisce direttamente il risultato richiesto senza fornire i dati in sé ma è dimostrabile che anche questo sistema ha gli stessi problemi poiché con la giusta sequenza di query è comunque possibile rilevare dati sensibili.

14.4 Differential Privacy

Ogni rischio relativo alle informazioni di una persona non dovrebbe cambiare significativamente l'output delle informazioni della persona stessa o, quanto meno, non nell'analisi.
Se nel mondo reale il risultato dell'analisi è ottenuto con tutti i dati e nel mondo ideale lo è senza i dati di una persona, la differenza tra i due risultati è al massimo ϵ. Lo scopo della differential privacy è quello di rendere $\epsilon = 0$ cioè che i risultati non abbiano differenze.
Per fare ciò non espongo i dati dell'individuo ma le probabilità ottenute con essi (es: A ha un cancro $->$ A probabilmente ha un cancro). Un algoritmo A si dice **differential private** se soddisfa la seguente condizione: date le probabilità P calcolate da A su entrambi i dataset D e D', il valore assoluto del rapporto tra le due probabilità è $\leq \varepsilon$.

$$sup_t |\log \frac{p(A(D) = t)}{p(A(D') = t)}| \leq \varepsilon$$

L'invarianza dopo l'elaborazione è la proprietà per la quale qualsiasi algoritmo eseguito dopo l'anonimizzazione non genera rischi per la privacy. Se si utilizzano diversi algoritmi di analisi sui dati con ϵ diverse, il rischio sarà al più la sommatoria delle ϵ. Con la privacy differenziali è possibile calcolare statistiche descrittive, compiti di machine learning supervisionati e non supervisionati (classificazioni, regressioni, clustering, apprendimento su distribuzioni ecc...) e generazione di dati sintetici.
L'**US Census Bureau 2020** utilizza l'analisi differenziale per esporre i propri dati e anche Google utilizza questa tecnica. Apple la utilizza ma senza rivelare le modalità ed il codice utilizzato, inoltre non sembrerebbe un algoritmo molto sicuro.
Google **TensorFlow** e Facebook **Opacus** possiedono delle implementazioni all'algoritmo differenziale.

Acronimi

2FA Two Factor Authentication. 18, 25, 29

ABAC Attribute Based Access Control. 31, 32
AgID Agenzia per l'Italia Digitale. 30
API Application Programming Interface. 10, 23
APK Android application PacKage. 7, 8
AWS Amazon Web Services. 10

C2 Command and Control. 1, 12, 14, 15, 20, 21, 27
CIA Central Intelligence Agency. 20

DAC Discretionary Access Control. 31
DDoS Distributed Denial of Service. 3, 4, 6, 7, 10, 19, 21, 23, 27
DoS Denial of Service. 21
DP3T Decentralized Privacy-Preserving Proximity Tracing. 35
DSD Dynamic Separation of Duty. 32

GCHQ Government Communications Headquarters. 37
GDPR General Data Protection Regulation. 34–36, 38–41

HTTP Hypertext Transfer Protocol. 12, 29, 33
HTTPS Hypertext Transfer Protocol Secure. 12, 18, 29

ICS Industrial Control System. 19, 21
IMP Identity Management Platform. 29
IoT Internet of Things. 8–10, 27, 31

LINDDUN Linkability Identifiability Non-Repudiation Detectability Disclosure of Information Unawareness Non-Compliance. 41, 42

MAC Mandatory Access Control. 31
MBR Master Boot Record. 14, 21, 24
MFT Master File Table. 14
MITM Man In The Middle. 27, 29

NSA National Security Agency. 20, 37

OS Operating System. 15, 19, 22–25
OTP One Time Password. 26, 29, 30

PETS Privacy Enhancing Technologies. 2, 37
PLC Programmable Logic Controller. 20, 21

RAT Remote Access Trojan. 6
RBAC Role Based Access Control. 31, 32
RTU Remote Terminal Unit. 21

SAML Security Assertion Markup Language. 30
SMB Server Message Block. 12, 24

SPID Sistema Pubblico di Identità Digitale. 30
SSD Static Separation of Duty. 32
SSL Secure Sockets Layer. 7, 18
SSO Single-Sign On. 29, 30

TCP/IP Transmission Control Protocol/Internet Protocol. 9

VPN Virtual Private Server. 21

XACML eXtensible Access Control Markup Language. 32, 33
XML Extensible Markup Language. 30, 32

Elenco delle figure

Elenco delle tabelle

Riferimenti bibliografici

[1] Satya Gupta, VIRSEC https://www.virsec.com/resources/learning-center/white-paper-taxonomy-of-the-attack-on-solarwinds-and-its-supply-chain

[2] I. Lella, M. Theocharidou, E. Tsekmezoglou, A. Malatras - European Union Agency for Cybersecurity, S. Garcia, V. Valeros - Czech Technical University in Prague https://www.enisa.europa.eu/publications/threat-landscape-for-supply-chain-attacks, Luglio 29, 2021

[3] C. Zaboeva https://securityintelligence.com/posts/german-task-force-for-covid-19-medical-equipment-targeted-in-ongoing-phishing-campaign/, Giugno 8, 2020

[4] J. Szurdi, Z. Chen, O. Starov, A. McCabe, R. Duan https://unit42.paloaltonetworks.com/how-cybercriminals-prey-on-the-covid-19-pandemic/, Aprile 22, 2020

[5] Dr. J. Scott Brennen, Felix Simon, Dr Philip N. Howard, Prof. Rasmus Kleis Nielsen https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation, Aprile 7, 2020