Modelli di DRM e problemi tecnici

Fabio Vitali
Università di Bologna

Premessa

- Ci sono mille argomenti legali, economici, sociali e politici in favore e contro il supporto tecnologico della proprietà intellettuale
- Ci sono milioni di discussioni su come possa essere realizzato nel momento in cui i contenuti sono disponibili digitalmente.
- Io parlo solo dell'aspetto tecnico, e di rispondere alla domanda "È possibile proteggere da copia il contenuto digitale?"(ma non di "è legale", "è giusto" o "conviene" o altri aspetti non tecnici)

La protezione dalla copia (1)

- Nel mondo pre-digitale era facile:
  - I macchinari necessari per ottenere copie della stessa qualità dell'originale erano molto costosi, e richiedevano un'intenzione chiaramente criminosa e facilmente identificabile e punibile.
    - Tipografie pirata per i libri
    - Studi di duplicazione pirata di nastri per la musica
    - Studi di duplicazione pirata di pellicole per i film…
  - Altrimenti era possibile ottenere copie di qualità inferiore attraverso apparecchiature di basso costo e ampia diffusione
    - Fotocopie
    - Audiocassette collegate ad un giradischi
    - Videocassette collegate alla TV

La protezione della copia (2)

- È nel mondo digitale che le cose si fanno difficili
  - I macchinari per creare duplicati di CD audio sono disponibili ovunque
  - I macchinari per creare duplicati di DVD video sono disponibili ovunque
- La copia è ottima, e di fatto indistinguibile dall'originale (a parte stampe e marchi ufficiali).
- Non solo, ma esistono metodi per ottenere copie di qualità *leggermente* inferiore con notevoli vantaggi in termini di trasportabilità e occupazione di memoria (es. mp3 o DivX)
- Il costo di creare una copia è diventato bassissimo e i macchinari sono a disposizione di CHIUNQUE.
- Infatti, il possesso di questi macchinari non è neanche vagamente indicativo dell'intenzione di dolere del possessore.

Digital Rights Management

- Forse un eufemismo
  - Negli anni ottanta si riferiva soprattutto alla protezione del software, e si chiamava semplicemente “copy protection”.
  - Altri termini: “copy control” “technical protection measures”, etc.
- Digital Rights Management è il termine ora diffuso per la descrizione dei meccanismi attivi e passivi che controllano copia, accesso, distribuzione e ridistribuzione di contenuti protetti da copyright
- Giovedì scorso un top manager di HBO (un canale via cavo americano) ha proposto un nuovo termine, “Digital Consumer Enablement”

Meccanismi di DRM

- Encrypting o scrambling
  - Crittografazione del contenuto, in modo che solo chi è in possesso della chiave di decrittazione giusta (che viene venduta a parte) possa accedere al contenuto.
- Marcatura
  - Al contenuto viene aggiunta un'informazione semplice che indica il copyright e le modalità lecite di uso del contenuto
- Incompatibilità selettive
  - Al contenuto vengono aggiunti "errori" che rendono il supporto inutilizzabile eccetto che dalle apparecchiature autorizzate, che sanno evitare gli errori.
- Ibridi
  - Ad esempio un sistema di marcatura che permette di attivare la chiave di decrittazione solo da parte delle apparecchiature autorizzate.

Scrambling

- Tutto il contenuto è crittato da una singola chiave: molto fragile e la violazione è catastrofica (BOBE o *Break Once, Break Everywhere*)
- Ogni pezzo di contenuto è crittato da una chiave diversa: la perdita di segretezza di una chiave rende solo quel contenuto a rischio.
- Chiavi associate a contenuto e apparecchiatura. Usato dalle TV via satellite europee (la smart card contiene una chiave autorizzata a decrittografare il segnale che arriva dall'etere).
- Chiavi multiple (DVD): il contenuto viene crittato con chiavi multiple, e ogni device viene caratterizzato da una chiave ufficialmente licenziata al produttore.

Marcatura (marking)

- Labeling
  - il contenuto arriva dotato di un'informazione aggiuntiva che non modifica il segnale, e che indica le modalità d'uso del contenuto. Apparecchi conformi impediscono l'uso del contenuto in maniere difformi dal consentito.
- Watermarking
  - Un segnale subdolo che viene aggiunto ad una piccola parte del contenuto che non ne impedisce la copia o l'uso, ma si mantiene da una copia all'altra.
- Fingerprint
  - Un'informazione univoca che permette di identificare con certezza il pezzo di contenuto in questione, e verificare centralmente se l'apparecchio è autorizzato a riprodurlo.

Principi del watermarking

- Un buon sistema di watermarking deve essere:
  - ***Impercettibile all'utente***: accedere al contenuto protetto non deve dare un'esperienza differente dal contenuto sprotetto.
  - ***Percettibile alla macchina***: un player autorizzato deve essere in grado di leggere con affidabilità il marker
  - ***Difficile o impossibile da rimuovere***: deve essere impossibile rimuovere il marker a meno di compromettere irrimediabilmente la qualità del segnale, almeno nella posizione del marker.
- Molti sistemi di watermarking sono incompatibili coi sistemi di compressione, che essendo *lossy* (cioè perdono segnale) non sono in grado di ricostruire esattamente il marker posto nel contenuto originale.
- Infatti, il marker viene "nascosto nel margine", e il margine è proprio il luogo che un meccanismo di compressione elimina per primo.

Principi del Fingerprint

- Un buon sistema di fingerprinting deve essere:
  - ***Unico o almeno preciso***: Due contenuti che sembrano diversi ad una persona debbono avere fingerprint diversi (la mia esecuzione di *Per Elisa* al pianoforte deve risultare differenziabile da quella di Arturo Benedetti Michelangeli)
  - ***Difficile o impossibile da rimuovere***: deve essere impossibile rimuovere il marker a meno di compromettere irrimediabilmente la qualità del segnale, almeno nella posizione del marker.
  - ***Persistente***: il contenuto deve essere identificabile anche in seguito a modifiche lecite al contenuto.
- Il fingerprint è un identificativo univoco del contenuto che deve essere verificato presso il server centrale per l'autorizzazione.

La persistenza del fingerprint

- Ci sono alcune azioni che pur essendo lecite su contenuti legalmente acquistati ne cambiano sottilmente o grandemente la natura tecnica:
  - Compressione del contenuto (come la generazione di un MP3)
  - La registrazione analogica del contenuto (suono la musica, e la riprendo con un microfono e la registro)
  - La modifica timbrica o di velocità o di colore o di nota in sottili maniere non distinguibili dall'utente
  - L'aggiunta di segnale ulteriore al contenuto (es. il simbolo di una stazione TV)
- Per sopravvivere anche a modifiche al contenuto, è necessario identificare quello che rende unico quello specifico pezzo di contenuto.

Comportamento delle apparecchiature

- Per far funzionare un sistema di marcatura, è necessario che le apparecchiature seguano le direttive contenute nel marker.
- Inoltre debbono sparire le copie non marcate di contenuto
- Debbono sparire tutti i device che non controllano l'esistenza del marker:
  - Quelli creati prima che il meccanismo di marker fosse stato introdotto
  - Quelli progettati e prodotti senza badare ai marker (per esempio in paesi in cui la legge di adesione al sistema di marking non è stata approvata o non è attuabile)
- Tre approcci al problema:
  - Ignorare: chi ha l'apparecchio vecchio o quello cinese continua allegramente ad accedere ai contenuti protetti, gli altri no.
  - Obsolescenza imposta: chi vuole accedere al contenuto si compra un apparecchio nuovo, altrimenti niente
  - Obsolescenza programmata: il vecchio contenuto è accessibile ai vecchi apparecchi, quello nuovo no.

Il buco dell'analogo

- Il buco dell'analogo (analog hole) è la ultima e definitiva vulnerabilità di qualunque sistema di protezione da copia.
- Salvo operazioni legislative (che effettivamente esistono) non c'è intrinsecamente difesa dal buco dell'analogo che non implichi un inaccettabile peggioramento della esperienza di accesso ai contenuti anche da parte di utenti legittimi.
- Il senso del buco dell'analogo è che ***prima o poi*** il contenuto protetto e crittato viene trasformato in un contenuto percepibile dall'utente, e quindi intrinsecamente non protetto.
- In quel momento, nel peggiore dei casi in formato analogico, esso può essere ricatturato e memorizzato in maniera totalmente sprotetta.
- Ovviamente, la cattura del contenuto sprotetto può avvenire anche solo all'ultimo stadio, quando è già in forma analogica e quindi di peggior qualità. Ma comunque paragonabile a quella dello schermo in cui è visualizzato!

Trusted Computing (1)

- Un'iniziativa per creare un'architettura di computer completamente affidabile (trusted):
  - Non soggetta a virus e malware
  - Non soggetta a furti di dati
  - Non soggetta a furto di identità
  - Affidabile in presenza di errori hardware e software
  - In grado di verificare correttamente e continuamente il possesso dei permessi per svolgere una determinata funzione
- Meccanismi di base:
  - Chiavi di attivazione: crittografia molto sofisticata e chiavi individuali
  - I/O sicuro: percorso dati dall'input alla memoria non intercettabile
  - Velatura della memoria: totale inaccessibilità di parti della memoria
  - Attestazione remota: l'hardware genera un certificato sul software che sta eseguendo, che può essere approvato o bloccato in caso di manipolazioni pirata

Trusted Computing (2)

- Vantaggi:
  - Inaccessibilità dei dati di un HD in seguito a furto
  - Digital Rights Management
  - Protezione dal furto di identità
  - Protezione da virus e spyware
  - Protezione dai bari nei giochi online
- Svantaggi
  - Non personalizzabilità del software (l'intero castello dell'open source non può usare trusted computing e non può neanche esere eseguito in un ambiente trusted computing)
  - Mancanza di controllo delle attività di un computer
  - Mancanza di controllo sui dati posseduti dall'utente
  - Censura
  - Perdita di anonimità

Alcuni casi importanti

- DeCSS (1999)
- Secure Digital Music Initiative Challenge (2000)
- Sony CD Rootkit (2005)
- Longest Suicide Note of History (2006)
- AnyDVD (2007)
- I numeri illegali: 09 F9 11 02 9D 74 E3 5B D8 41 56 C5 63 56 88 XX

DeCSS (1999)

- Il Content Scrambling System è un sistema di DRM attualmente usato su tutti i DVD di prima generazione
- Introdotto nel 1996, nel 1999 è stato violato da un giovane norvegese, Jon Johansen (noto dome DVD Jon) con il software DeCSS
- Secondo CSS, ogni device autorizzato è in possesso di una chiave CSS. Ogni DVD commerciale è in possesso di un'altra chiave. La coppia delle chiavi viene utilizzata per accedere e decodificare il contenuto.
- La violazione della sicurezza di un device implica la rimozione dell'autorizzazione da tutti i device che usano quella chiave (ce ne sono circa 400).
- In seguito ci si accorse che la piccola dimensione della chiave utilizzata (40 bit) ammetteva soluzioni "di forza bruta" e da allora sono sorti più di cento software diversi che usano tecniche diverse per accedere ai contenuti di un DVD.

La sfida SDMI

- Nel settembre 2000, la Secure Digital Music Initiative pubblicò un sfida in cui si chiedeva alla comunità di hacker, esperti di sicurezza e studiosi accademici a testare e violare il loro schema di watermarking digitale inserito su file di musica in formato MP3.
- I quattro differenti modelli di watermarking avevano richiesto cumulativamente vari anni di lavoro per ideazione, realizzazione e test.
- Un gruppo di ricercatori dell'università di Princeton e Rice riuscirono a violare tutti e quattro i metodi in meno di tre settimane, creando file musicali in cui il watermark era stato rimosso senza perdita inaccettabile di qualità musicale.
- Poiché la soluzione trovata non era connessa con la specifica implementazione, ma con i concetti stessi utilizzati per generare il watermark, che causarono la fine delle attività pubbliche di SDMI fin dalla metà di maggio del 2001.

Sony CD Rootkit (2005)

- La azienda First 4 Internet, nel 2005, produsse per conto di alcune major musicali un pacchetto chiamato Extended Copy Protection (XCP).
- Sony BMG distribuì più di 250 titoli (svariati milioni di CD) con XCP installato.
- CD musicali che avevano XCP installato potevano essere suonati normalmente su CD player tradizionali, ma non su CD di computer, perché risultavano in un suono distorto (ma solo su Windows: su Macintosh e Linux suonavano normalmente).
- Per suonare normalmente su Windows, l'utente doveva installare un programma proprietario.
- Questo venive eseguito in maniera protetta sul computer e che in pratica aprivano la porta per la creazione di software maliziosi che prendevano possesso del computer, nonché consumavano risorse in grande quantità e provocavano sporadici crash di sistema.

Sony CD Rootkit (2005)

- Un software innocuo che apre la porta per attacchi da parte di applicazioni maliziose si dice rootkit. Quindi XCP funzionava di fatto da rootkit per qualunque tipo di virus, cavallo di troia, spyware e ogni altro malware che ne sfruttasse (banalmente) i difett idi progettazione
- Sony venne denunciata da almeno tre stati americani (New York, California e Texas), produsse un disinstallatore che apriva ancora più buchi del rootkit originale, dovette richiamare quasi 4 milioni di CD e subì una contro-campagna di boicottaggio notevolissima e che dura ancor'oggi.
- First 4 Internet cambiò nome nel novembre 2006 diventando Fortium technologies ed è stata denunciata da pià case discografiche.

Longest Suicide Note (2006)

- *A Cost Analysis of Windows Vista Content Protection* è un saggio di novembre 2006 (ultimo aggiornamento, aprile 2007) di Peter Gutmann, neozelandese, in cui si analizzano le caratteristiche di trusted computing presenti nell nuovo sistema operativo Microsoft.
- MS Vista Content Protection è una serie di iniziative software e di certificazione hardware che permettono di garantire ai proprietari di diritti su contenuti commerciali (in particolare HD-DVD e Blue-Ray) che sono usi autorizzati dei contenuti possono essere fatti su computer che usano Vista.
- In particolare, non è possibile intercettare e salvare diversamente i contenuti una volta decrittati (cioè non è possibile in Vista craccare HD-DVD e BlueRay)

Longest Suicide Note (2006)

- Questo avviene creando un canale sicuro di accesso al contenuto in cui non solo il software, ma anche l'hardware deve essere certificato e verificato in continuazione.
- Hardware precedenti a Vista non possono essere usati, hardware non licenziati da Microsoft non possono essere usati, hardware che, anche se licenziati da Microsoft, si dimostrano con il tempo vulnerabili non possono essere usati.
- In caso di piattaforma non sicura, il sistema operativo autonomamente degrada la qualità del segnale fino a livelli simili ai DVD tradizionali.
- Poiché questo va contro a trend storici (unificazione dei driver, evoluzione continua di nuovi device innovativi, aumento dei costi di upgrade e di nuovo acquisto, ecc. si è coniato il termine "Longest suicide note in history" per le specifiche del Vista Content Protection.

AnyDVD (2007)

- Nell'ottobre del 2006 sono usciti sul mercato i primi lettori di DVD ad alta definizione, nei due standard in competizione HD-DVD e Blue-Ray.
- Entrambi si basano su AACS, un derivato di CSS con chiavi molto più lunghe e una sicurezza molto maggiore. Lo sviluppo di AACS ha richiesto svariati anni di lavoro e milioni di dollari di sviluppo.
- AACS fu votato come una delle tecnologie più probabilmente destinate al fallimento nel gennaio del 2005 dalla rivista IEEE Spectrum.

AnyDVD (2007)

- Furono trovati tre modelli di attacco funzionanti:
  - Print Screen: salvare ogni singolo fotogramma come appare a schermo e ricostruire il film in un altro formato
  - Memory space snooping: spiare nella memoria per le chiavi di crittografazione e salvarle.
  - AnyDVD: nel gennaio del 2007 la versione nuova di AnyDVD (un descrittatore di DVD tradizionali) incluse un attacco generalizzato a qualunque chiave di AACS direttamente da disco e in maniera generalizzata.
  - A marzo 2007 tutte le chiavi esistenti di AACS sono state revocate e il 22 maggio è prevista la prima uscita di nuovi titoli HD-DVD e Blu-ray con chiavi rinforzate.
  - **Oggi** è uscita una nuova versione di AnyDVD che è in grado di superare le chiavi rinforzate in uscita tra 5 giorni.

Numeri illegali

09 F9 11 02 9D 74 E3 5B D8 41 56 C5 63 56 88 XX

- Secondo DMCA (la legge che regola i meccanismi di protezione dei contenuti digitali) è illegale possedere non solo contenuto copiato, ma anche informazioni utili per la realizzazione di copie illecite.
- Il numero qui sopra è un numero fondamentale per la violazione della protezione AACS. Il suo possesso è ILLEGALE negli Stati Uniti per merito di DMCA.
- Nell'aprile del 2007, questo numero è stato pubblicato su un sito di materiale di hacker, ed è stato subito denunciato dalla Motion Picture Association of America come violazione di DMCA.
- Da allora più di tredici milioni di siti al mondo riportano questa chiave, sotto forma di testo, immagini, file musicale, fotografia di testo, fotografia di tatuaggio, fotografia di gente disposta sulla spiaggia, ecc.

Riferimenti

- *http://www.publicknowledge.org/pdf/citizens_guide_to_drm.pdf*
- *http://en.wikipedia.org/wiki/Technical_protection_measures*
- *http://en.wikipedia.org/wiki/Digital_Rights_Management*
- *http://en.wikipedia.org/wiki/Trusted_Computing*
- *http://www.cs.auckland.ac.nz/~pgut001/pubs/vista_cost.html*
- *http://en.wikipedia.org/wiki/2005_Sony_BMG_CD_copy_protection_scandal*
- *http://en.wikipedia.org/wiki/Secure_Digital_Music_Initiative*
- *http://en.wikipedia.org/wiki/AnyDVD*
- *http://en.wikipedia.org/wiki/DeCSS*