Aspectos prácticos de la protección de la propiedad intelectual en contenidos multimedia

Antoni Martínez Ballesté, Francesc Sebé y Josep Domingo-Ferrer

Universitat Rovira i Virgili, Dept. of Computer Engineering and Mathematics,
Av. Països Catalans 26, E-43007 Tarragona, Catalonia, Spain,
e-mail {anmartin, fsebe, jdomingo}@etse.urv.es

Resumen Internet se ha convertido en un entorno ideal para la distribución ilegal de archivos multimedia. La lucha contra la piratería se ha materializado en propuestas de prevención de copia, a base de cambios en el hardware o software o bien a base de licencias de reproducción. Dado que estos sistemas acaban siendo rotos por *hackers*, se presentan los sistemas de detección de copia, cuyos diseños se basan en la estenografía. En este artículo se hace una revisión del panorama de protección del copyright y se propone un sistema implementable para contenido digital y Internet.
Palabras clave: Comercio electrónico, protección del copyright, huella digital

1. Introducción

El creciente número de usuarios de Internet y la proliferación de conexiones de banda ancha han convertido la red de redes un entorno ideal para la distribución de contenidos multimedia. Se espera que Internet sea una plataforma para el comercio de música o vídeos en formato digital, pero el hecho de que muchos usuarios sean reacios a la hora de efectuar pagos por Internet está retardando considerablemente esta previsión.

Por otra parte, Internet se ha convertido en el canal de distribución ilegal por excelencia. Millones de usuarios usan programas P2P (*Peer-to-peer*, de igual a igual) para obtener versiones ilegales de películas y música.

Una de las características de los datos en formato digital es la posibilidad de realizar copias con una calidad idéntica a la del original. Es interesante, pues, el desarrollo de sistemas anticopia para contenido digital, puesto que las copias piratas obtenidas a partir de contenidos analógicos (como por ejemplo la grabación con video cámara en una sala de cine, técnica llamada *screener*) sufren una degradación de la calidad de sonido y/o imagen. En cambio, el proceso de obtención de una copia ilegal de un DVD usando la compresión DivX es totalmente

digital y se obtienen copias de alta calidad, por lo tanto en este caso deben hacerse esfuerzos para proteger la propiedad intelectual. La facilidad de copiar contenido multimedia sin perder calidad propicia el aumento de la piratería y distribución ilegal de copias.

1.1. Prevención de copia y detección de copia

Algunas formas de luchar contra la piratería se basan en evitar que un producto copiado se pueda reproducir en reproductores no autorizados, o bien que sólo pueda reproducir el contenido quien tenga derecho a hacerlo. Estos sistemas presentan dos inconvenientes. El primero es que suelen romperse al poco tiempo de ponerse en uso mientras que, por otra parte, algunos de estos sistemas requieren de ciertos cambios en el hardware o software de los equipos. Es de suma importancia considerar que estos sistemas requieren de consenso entre fabricantes de aparatos y distribuidoras de contenido.

Dado que los sistemas anticopia suelen fallar en un tiempo relativamente corto, una forma efectiva de controlar la distribución ilegal viene dada por las denominadas técnicas de detección de copia. Se basan en esconder una marca imperceptible en el contenido antes de venderlo. La posterior recuperación de la marca de una copia ilegal revelará quien fue el autor de la copia. Así como existen herramientas para sistemas anticopia, las técnicas de detección de copia se encuentran básicamente en fase de desarrollo teórico.

1.2. Contribución y estructuración

En este artículo se presenta un sistema de detección de copias ilegales asimétrico y anónimo, implementable en un entorno real. La Sección 2 presenta varias iniciativas para la protección del copyright, basadas en control de licencias o permisos sobre el contenido. En la siguiente sección se introduce el tema de la protección del copyright mediante técnicas de detección de copia. Finalmente, la Sección 4 presenta un sistema de asimétrico y anónimo basado en terceras partes de confianza.

2. Iniciativas para la protección del copyright

La mayoría de contenidos audiovisuales están realizados con el fin de que sus autores y distribuidores obtengan un beneficio comercial. Este beneficio se obtiene a través de licencias de reproducción, que también se

aplican a los programas de ordenador. Mediante la distribución de copias ilegales, el propietario del copyright no recibe beneficio alguno.

Algunos CD de música no son reproducibles en un PC, de modo que, en teoría, tampoco es posible hacer una copia digital mediante ordenador. Como se ha apuntado en la introducción, los sistemas que impiden la copia suelen romperse al cabo de un tiempo, como por ejemplo el caso del DVD [1].

Ante el gran problema que supone el auge de la piratería, han surgido varios sistemas que pretenden proteger la propiedad intelectual de los contenidos multimedia, centrándose en su distribución por Internet.

2.1. Digital Rights Management

Mediante plataformas para *Digital Rights Management* (DRM) se pretende controlar quién tiene copias legales de un producto [2]. Bajo esta definición se hallan varias propuestas de estándares, entre las cuales destacan: *Digital Objects Identifiers* [3], *Extensible Rights Markup Language* [4] y *Extensible Media Commerce Language* [5].

El software para DRM puede definir varios niveles de derechos sobre un contenido multimedia: para unos usuarios sólo se puede realizar la reproducción, mientras que para otros es posible la modificación del contenido. Se tiene en cuenta un amplio abanico de clientes, desde ordenadores personales y asistentes digitales (PDA), pasando por teléfonos móviles y equipos de televisión digital.

Su funcionamiento se basa en servidores de derechos (*right databases*) los cuales permiten, según los privilegios del usuario, obtener una clave para ver el contenido o acceder a éste para su modificación. La obtención de claves se lleva a cabo con el *DRM viewer* [6]. Las tarjetas inteligentes tendrán un papel importante en la implantación de los sistemas DRM.

A medida que se implanten sistemas DRM, es probable que la comunidad *hacker* vaya buscando métodos para anular su eficacia.

2.2. El caso de Microsoft

La compañía Microsoft ha desarrollado su propio sistema para DRM, el llamado *Windows Media DRM* [7]. Mediante este producto, incorporado en las últimas versiones de su sistema operativo, los proveedores y vendedores pueden distribuir contenidos multimedia protegidos contra la piratería.

Mediante el *Windows Media Rights Manager* se cifra el contenido multimedia con una clave. Esta clave se guarda en una licencia de uso. El

contenido multimedia especifica otra información como la localización en Internet de esta licencia. El archivo queda guardado usando los estándares Microsoft (*Windows Media Audio* o *Windows Media Video*).

Para reproducir un archivo protegido, el reproductor multimedia de Windows llevará el cliente a la web donde podrá adquirir la licencia. Como es habitual en este tipo de protecciones, en Internet se puede encontrar una forma de romper el sistema [8].

3. Técnicas de detección de copia

En este caso la idea no es impedir la copia, sino saber quién la hizo. Las llamadas técnicas de huella digital (*fingerprinting*) usan técnicas de marcas de agua (*watermarking*) para encastar una marca que identifica en cierta forma el comprador de la copia del contenido multimedia [9].

Lo que se hace habitualmente es esparcir la marca por el contenido: por ejemplo, poner una marca visible en una esquina de una imagen sirve de poco, puesto que su eliminación se puede llevar a cabo con un simple recorte. Las alteraciones que sufre el contenido durante el proceso de marcaje deben ser imperceptibles.

Además, la marca debe resistir a una serie de transformaciones (propiedad de robustez): en el mismo caso de la imagen, el mensaje encastado debe poderse extraer incluso si el contenido ha sido escalado o rotado. En un sistema de huella digital se debe tener en cuenta tanto el algoritmo de encaste de la marca como la forma de la marca. Estos elementos dependen, en cierta forma, del tipo de objeto a marcar y de su número potencial de clientes.

3.1. Proceso de marcado

Hay que tener en cuenta que la marca a encastar (habitualmente una secuencia de bits) tendrá una longitud limitada. Esta longitud dependerá de la capacidad del objeto donde se encaste la marca, así como del algoritmo usado para encastarla. La capacidad se define como el número máximo de bits de marca que se pueden encastar al sistema de forma imperceptible. Esta capacidad depende básicamente del algoritmo y del nivel de imperceptibilidad deseado.

Para aumentar la robustez del marcaje se hace que ésta contenga cierta redundancia para que pueda tolerar cierto número de errores. Supongamos un determinado objeto que admite una marca de 128 bits (con marcas superiores se detecta, cualitativamente, una degeneración de

su calidad). En este caso podemos escoger o bien una marca con 128 bits de información o bien una marca de 64 bits por duplicado o bien una marca de 32 bits encastada cuatro veces, etc. En resumen, a más redundancia, más robustez en la marca y, a más redundancia, menos bits de información. Existen diferentes propuestas para para el marcaje de imagen [10], vídeo [11] y audio [12].

3.2. Asignación de marcas a usuarios

El principal problema de la técnica de la huella digital viene dado por el hecho de que cada copia vendida lleva una marca distinta. Esto conlleva dos problemas: en primer lugar se debe tener en cuenta que si el número de posibles compradores es elevado, el algoritmo utilizado para marcar debe tener capacidad suficientemente grande. Por otro lado, el hecho de que todas las copias sean distintas hace posible los denominados ataques por confabulación.

En estos ataques, dos o varios compradores deshonestos comparan bit a bit sus respectivas copias y utilizan la información sobre las diferencias encontradas para componer una nueva copia cuya marca no les identifique.

Para hacer que estos ataques no sean exitosos, la cadena de bits que identifica a un comprador es una palabra de un código resistente a confabulaciones [13]. La particularidad de estos códigos es que, mientras la confabulación no supere el límite de confabulados que permite el código, la copia generada en una confabulación contendrá una marca que identificará como mínimo a uno de ellos.

3.3. Sistemas de huella digital asimétrico y anónimo

En los primeros sistemas de huella digital, llamados simétricos, es el vendedor quien marca el contenido. El problema de esto es que un vendedor deshonesto puede distribuir ilegalmente una copia que éste ha marcado para un determinado usuario. Si se encuentra la copia ilegal, se culpará al usuario. Por otra parte, un cliente fraudulento puede distribuir ilegalmente su copia y luego acusar al vendedor de la distribución.

Para evitar este problema es necesario que durante el proceso de marcado el vendedor no tenga acceso a la copia marcada. Los sistemas que cumplen con este requisito son los denominados asimétricos [14]. En ellos, el proceso de marcaje consiste en un protocolo en el que intervienen vendedor y comprador.

En la literatura existen distintas propuestas de sistemas asimétricos de huella digital. En [14] se utiliza computación multiparte, que es muy

compleja computacionalmente. La propuesta [16] utiliza la herramienta de transferencia inconsciente [17]. El problema de esta propuesta, pese a ser teóricamente correcta y realizable, es el coste de cálculo y de comunicación, siendo impracticable para objetos grandes. En [18] se usa una primitiva de prueba de conocimiento nulo cuya implementación no está claro que sea posible.

Anonimato En comercio electrónico el anonimato es un aspecto fundamental. Es un derecho para los usuarios poder comprar un determinado producto sin que su identidad sea conocida, tal y como ocurre con el pago en efectivo. El anonimato debe seguir preservándose aun usando huella digital. El anonimato se consigue generalmente mediante el uso de pseudónimos [15].

4. Implementación un sistema de huella digital

A continuación se propone un sistema de huella digital. La implementación de un sistema asimétrico y anónimo de huella digital debe tener en cuenta dos aspectos: por una parte el control de clientes y encaste de marca y, por otra parte, la localización de copias ilegales y la revelación del cliente deshonesto. El sistema es asimétrico en el sentido de que el vendedor no tiene conocimiento de la copia marcada del producto.

4.1. Escenario

El entorno donde opera el sistema de huella digital está formado por las siguientes entidades:

- **Vendedor.** Se trata de una tienda virtual que vende, a través de la web, contenidos multimedia (imágenes, música y vídeo) descargables de la red.
- **Cliente.** Compra contenido multimedia y puede distribuirlo ilegalmente. Para este fin dispone de varias posibilidades. Una de ellas es distribuir el archivo mediante una herramienta P2P tal como Kazaa [19] o eDonkey [20]. Otra forma consiste en copiar el contenido multimedia en un soporte físico del cual podrá sacar copias para la venta o distribución. Es posible que varios compradores se reúnan y realicen un ataque por confabulación, tal y como se ha descrito anteriormente.

El sistema de huella digital está formado por dos entidades en quien se confía:

- **Autoridad pública de huella digital.** Esta parte de confianza lleva a cabo el proceso de marcaje del objeto multimedia.
- **Autoridad de registro.** Esta parte de confianza proporciona pseudónimos a los clientes, para que estos puedan realizar compras.

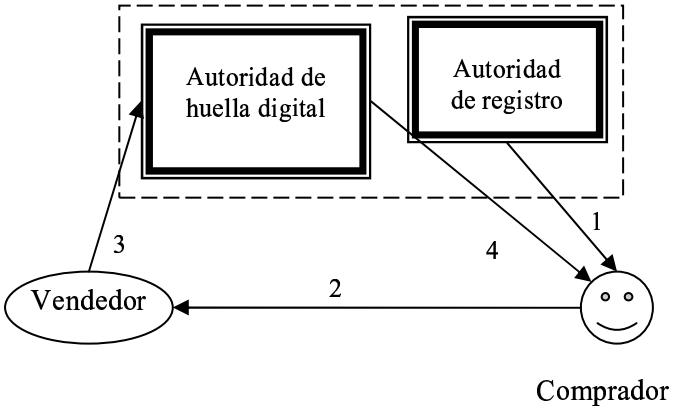

Figura 1. Proceso de compra

4.2. Proceso de compra

Mediante el protocolo de registro de vendedor se establece un acuerdo para un vendedor V que quiere usar los servicios de la autoridad de huella digital AHD. Usando técnicas criptográficas convencionales, se establecerá un canal de comunicación seguro entre ambas partes.

La Figura 1 ilustra el proceso de compra. En resumen, el proceso de compra empieza con el registro del cliente, si es que éste no está registrado. El cliente usa un pseudónimo para hacer sus compras, el cual se obtiene de la autoridad de registro (1). En (2) el comprador pide un producto. El vendedor pide a la autoridad de huella digital que marque el producto (3). Después del marcado, la autoridad de huella digital entrega el producto marcado al comprador (4). Véase que el vendedor no ve la copia marcada.

Registro del cliente La autoridad de registro proporciona un pseudónimo al cliente. Este protocolo requiere que el cliente tenga una clave pública certificada, que se usa para autenticar su identidad. La autoridad de registro genera el pseudónimo de forma aleatoria y lo guarda

en su base de datos, juntamente con la identidad del comprador. Es conveniente que la información que contiene el certificado contenga datos personales, como la dirección o el teléfono, que serán de utilidad en caso de identificar al usuario deshonesto.

Compra de contenido multimedia El cliente visita la web del vendedor y se interesa por un producto en concreto. Si decide comprar el producto X, se obtiene su descripción. El cliente C compone un mensaje *msg* que contiene: su pseudónimo ps_C, la descripción del producto, $desc_X$, y la clave pública del cliente, PK_C.

El mensaje se cifra con la clave pública de la autoridad de huella digital, PK_{AHD}. Nótese que de esta forma el pseudónimo sólo será conocido por *AHD* y el vendedor no podrá usarlo para hacer compras sin conocimiento del cliente. El cliente C se pone en contacto con el vendedor y le manda una orden de compra firmada: $ord = SIG_{SK_C}(E_{PK_{AHD}}(msg), desc_X, fecha, hora)$.

C realiza una operación para pagar el producto. Para preservar el anonimato, este pago debe hacerse mediante un sistema que lo preserve.

Una vez el vendedor V comprueba la validez del pago y de *ord*, manda el objeto a marcar X y la información del cliente $E_{PK_{AHD}}(msg)$ hacia *AHD*.

AHD descifra *msg*, coge el pseudónimo y comprueba que $desc_X$ concuerda con X. La autoridad elige una palabra código w de un código resistente a confabulaciones:

- Si X no se ha marcado todavía, *AHD* construye un código resistente a confabulaciones Γ_X. La construcción de este código está sujeta a los parámetros indicados por el vendedor. Se asigna la primera palabra del código a la compra *msg*.
- Si no, se asigna a *msg* la primera palabra no usada de Γ_X.

AHD encasta w en el objeto X, usando un sistema robusto de marca de agua, que dependerá del tipo de objeto comprado. La autoridad guarda el pseudónimo ps_C y la palabra w en su base de datos. La copia marcada, $\bar{X}$, se cifra usando PK_C y se manda hacia el cliente.

4.3. Identificación de clientes deshonestos

Como se ha comentado anteriormente, el cliente deshonesto $\breve{C}$ puede introducir su copia en un entorno como eMule [21]. Por otra parte, $\breve{C}$ puede copiar $\bar{X}$ en un soporte digital y distribuir varias copias. En

este caso, una serie de clientes deshonestos podrían extraer el contenido digitalmente y poner el objeto $\bar{X}$, aun con el nombre del fichero cambiado, en un entorno P2P.

También se puede introducir en el entorno P2P una copia $\hat{X}$, generada a partir de la confabulación de varios compradores. Para que el vendedor pueda detectar posibles copias ilegales, deberá disponer de clientes P2P. Una vez encontrada y descargada una posible copia fraudulenta, se procederá a la recuperación de identidad.

Recuperación de identidad En este caso el vendedor V procede como sigue: V manda la copia ilegal a AHD, que recupera la marca m de la copia ilegal. Se decodifica m de tal forma que se obtiene una lista de palabras código. En caso de no haber confabulación se obtendrá sólo una palabra código: la que identifica al comprador que ha distribuido ilegalmente su copia.

La lista de palabras código se manda a la autoridad de registro para revelar la identidad de los compradores culpables, para que el vendedor pueda emprender acciones legales.

5. Conclusiones

En este artículo se ha hablado de sistemas para la protección del copyright de contenidos multimedia. Por una parte, se han presentado algunas propuestas cuyo objetivo es la prevención de copia. En contrapartida, se han introducido temas referentes a la detección de copia y se ha justificado su uso frente a los métodos de prevención. Los sistemas de prevención de copia han dado lugar a multitud de publicaciones en la literatura científica. En este artículo se ha presentado un sistema de detección de copias ilegales asimétrico y anónimo, implementable en un sistema real. Se ha ilustrado un escenario de uso del sistema.

Las entidades de confianza autoridad de huella digital y autoridad de registro deben contar con el consenso y la aceptación general similar a la que requieren las autoridades de certificación. Ciertamente, nadie puede dudar de la honestidad de una autoridad de certificación. Lo mismo ocurre con la honestidad de la autoridad de huella digital. Grandes empresas de distribución audiovisual podrían llegar a desempeñar este papel.

Referencias

1. http://www.lemuria.org/DeCSS/

2. S. M. Cherry, "Making Music Pay" en *IEEE spectrum*, vol. 38, no. 10, pp. 41–46, oct. 2001.
3. Digital Objects Identifiers, `http://www.doi.org`
4. Extensible Rights Markup Language, `http://www.xmrl.org`
5. Extensible Media Commerce Language, `http://www.xmcl.org`
6. S. Marks, "Staking out digital rights", en *Network World*, feb. 2002, `http://www.nwfusion.com/ecomm/2002/rights/rights.html`
7. Microsoft DRM `http://www.microsoft.com/windows/windowsmedia/drm.aspx`
8. Microsoft's Digital Rights Management Scheme - Technical details, `http://cryptome.org/ms-drm.htm`
9. S. Katzenbeisser y F. Petitcolas, *Information Hiding. Techniques for Stenography and Digital Watermarking*, Artech House, 2000.
10. F. Sebé, J. Domingo-Ferrer y J. Herrera-Joancomartí "Spatial-domain image watermarking robust against compression, filtering, cropping and scaling" en *Information Security - LNCS 1975*, pp.44-53, Berlín, 2000. *LNCS*, vol. 2200, pp. 420-432, oct. 2001. Vol. *Information Security*, eds. G. Davida y Y. Frankel, Berlin: Springer-Verlag.
11. F. Hartung y B. Girod, "Watermarking of uncompressed and compressed video" en *Signal Processing*, vol. 66, no. 3, pp. 283–301, 1998.
12. L. Boney, A. H. Tewfik y K. N. Hamdy, "Digital Watermarks for Audio Signals" en *International Conference on Multimedia Computing and Systems*, pp.473-480, 1996.
13. D. Boneh y J. Shaw, "Collusion-secure fingerprinting for digital data", a *Advances in Cryptology - CRYPTO'95*, LNCS 963, Springer-Verlag, 1995, pp. 452-465.
14. B. Pfitzmann y M. Schunter, "Asymmetric fingerprinting", a *Advances in Cryptology - EUROCRYPT'96*, LNCS 1070, Springer-Verlag, 1996, pp. 84-95.
15. B. Pfitzmann y M. Waidner, "Anonymous fingerprinting", a *Advances in Cryptology - EUROCRYPT'97*, LNCS 1233, Springer-Verlag, 1997, pp. 88-102.
16. J. Domingo-Ferrer, "Anonymous fingerprinting based on commited oblivious transfer" en *Public Key Cryptography, PKC'99 - LNCS 1560*, pp.43-52, Springer-Verlag, 1999.
17. C. Crépeau, J. van de Graaf y A. Tapp, "Commited oblivious transfer and private multi-party computation" en *Advances in Cryptology - CRYPTO'95 - LNCS 963*, pp. 110-123, Springer-Verlag, 1995.
18. J. Domingo-Ferrer y J. Herrera-Joancomartí, "Efficient smart-card based anonymous fingerprinting" en *Smart Card Research and Applications, CARDIS'98 - LNCS 1820*, pp. 231-238, Springer-Verlag, 2000.
19. Kazaa Media Desktop, `http://www.kazaa.com`
20. eDonkey, `http://www.edonkey2000.com`
21. The eMule Project, `http://www.emule-project.net`