Sellos Digitales para Documentos[1]

Santiago Gómez, Víctor Bogarín y Benjamín Barán

Email: [bbaran, vbogarin] @cnc.una.py

Centro Nacional de Computación - CNC

Universidad Nacional de Asunción - UNA

Casilla de Correo:1439

Campus Universitario - San Lorenzo

Paraguay

Resumen

El presente trabajo especifica la definición de los "*Sellos Digitales para Documentos*", su alcance, ámbito de aplicación y limitaciones. A partir de esta conceptualización, se desarrolla un esquema de utilización que permite implementar un prototipo computacional.

El trabajo describe la implementación de un *kit* con varias opciones de "*Sellador*", que introduce marcas ya sea visibles o invisibles, en documentos digitales en formatos de uso corriente en las oficinas modernas. De esta forma, se puede reconocer la propiedad y el origen del documento así como otras posibles informaciones encubiertas, mediante un *"Reconocedor de Sello"*. Como experiencia inicial, se ha diseñado un prototipo aplicable al formato RTF pero con técnicas ampliables a otros formatos.

Palabras Claves: Seguridad Informática, Marcas de Agua (*Watermarking*), Escritura Encubierta (*Steganography*), criptografía, Formato de Texto Enriquecido (*Rich Text Format*, RTF), derecho de propiedad.

1. INTRODUCCIÓN

El acelerado proceso de digitalización de los últimos años, especialmente desde la cobertura masiva de Internet, ha expandido considerablemente nuestros conceptos clásicos de firmas, sellos o huellas, surgiendo la necesidad de aprovecharlos en los nuevos contextos computacionales, con fines similares de garantizar autenticidad, propiedad y fuente, de igual forma que sus equivalentes más tradicionales, de amplia difusión hasta nuestros días [4].

Dentro de esta nueva realidad tecnológica, se van proponiendo modernas técnicas digitales orientadas a la seguridad en el intercambio y almacenamiento electrónico de información. Así, se conocen diversas técnicas para asociar un documento con su autor utilizando criptografía (*encryption*); es decir, volviendo ilegible al documento para el que no posea la correspondiente clave de acceso (*password*) [5].

[1] Proyecto apoyado por la Dirección de Investigación, Postgrado y Relaciones Internacionales (DIPRI) de la Universidad Nacional de Asunción (UNA)

También existen mecanismos, como el *Message Digest*, que permiten asegurar que un documento electrónico no pueda ser alterado sin que esto pase inadvertido para quienes poseen acceso a la clave del documento. Las implementaciones mas conocidas son MD5 y SHA [10].

En la actualidad, se conocen varios mecanismos orientados a la seguridad, enfocados a diferentes aspectos como :

- Métodos de Clave Secreta, como el DES, IDEA y otros;
- Algoritmos de Clave Publica, como RSA;
- Protocolos de Autenticación basados en Clave Secreta Compartida como el de "Desafío-Respuesta";
- Formas de intercambio de claves como el de "Diffie-Hellman";
- Centros de Distribución de Claves con protocolos como el de "Desafío-Respuesta de Múltiples Vías" o variantes de esta como la de Otway y Rees, Kerberos, etc. [10];
- Autenticación usando clave pública;
- Firmas Digitales con clave secreta o pública,

así como programas que implementan combinaciones de estas funcionalidades para proveer privacidad en el intercambio electrónico de documentos, de los cuales los más representativos son PGP (*Pretty Good Privacy*) y PEM (*Privacy Enhanced Mail*) [10].

El trabajo está organizado de la siguiente manera: en la sección 2 se muestra un esquema de funcionamiento del programa, con sus módulos componentes y técnicas utilizadas; en la sección 3 se da una reseña sobre *watermarking*, en que consiste esta tecnología de uso creciente, junto con sus aplicaciones; la sección 4 describe otra técnica utilizada en este trabajo, cual es la *esteganografía*; la sección 5 se centra en detallar las características del Sellador-Reconocedor, junto con experiencias y resultados obtenidos; en la sección 6 se presentan los pseudocodigos de las diferentes opciones del prototipo; la sección 7 está dedicada a resaltar los aspectos novedosos y originales del trabajo y, por ultimo se tienen las conclusiones y trabajos futuros, en la sección 8, además de las referencias bibliográficas.

2. LA PROPUESTA

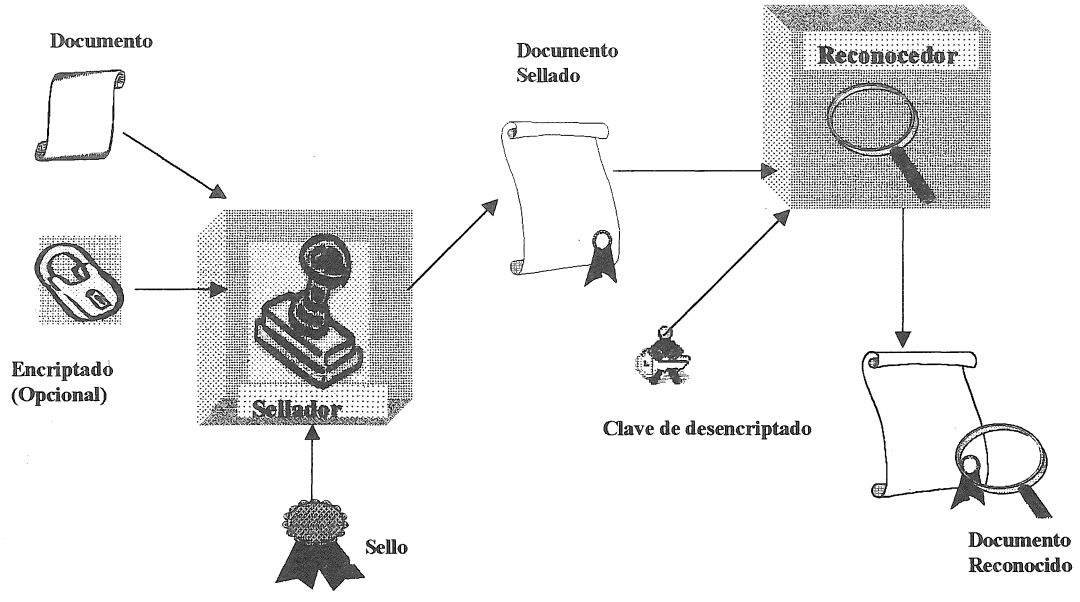

Figura 1. Esquema del prototipo implementado

La figura 1 ilustra el esquema de funcionamiento del prototipo implementado. Se cuenta con un sellador que inserta en documentos en formatos normalizados (RTF en el caso del prototipo desarrollado), unas marcas o patrones digitales que resultan totalmente desapercibidos ante la revisión del documento por medio de procesadores de textos o editores corrientes. De acuerdo al método utilizado, se puede contar con el encriptado del sello a través de una clave secreta, con lo cual solo los poseedores de esa clave pueden visualizar el texto del sello. En otros casos, este texto puede ser visible directamente, es decir, sin necesidad de la aplicación de mecanismos adicionales de seguridad.

Este documento sellado es reconocido si se dispone del "*Reconocedor de Sello*" (programa que hace las veces del equipo ultravioleta que reconoce la marca inserta en los billetes de cien dólares) que tiene la capacidad de reconocer el sello y cualquier otra información adicional que se haya incluido con el mismo.

Para implementar la presente propuesta se desarrolló un prototipo que realiza tres variedades de sellados que tienen características y funcionalidades diferenciadas, las que son explicadas a continuación:

- *Sellos Aplicados al Layout.* En este caso se comprimen o expanden de forma casi imperceptible los caracteres de un texto de modo que los mismos formen un patrón, correspondiente a un alfabeto paralelo, a través de cuya posterior lectura se puede descifrar el mensaje oculto. Esto puede hacerse en dos modalidades:
  - *Genérico:* en el cual un usuario del programa puede ejecutarlo para revisar cualquier documento y podrá visualizar el mensaje encubierto.
  - *Personalizado:* en este caso un usuario que no posea la clave de acceso necesaria podrá ver que existe un mensaje pero no podrá acceder a su contenido, ya que se encontrará encriptado.

- *Sellos Aplicados al Archivo.* Aquí lo que se aplican son marcas insertadas en el archivo en lugares predeterminados y no afectados por el contenido del documento ni por sus formateos. Se dispone de una clave que solo es conocida por el poseedor de dicho programa, y por consiguiente, asegura el origen del sello y su uso exclusivo por parte de su propietario.

- *Sellos Visibles*: En este caso se inserta un objeto ya sea textual, gráfico, imágenes, etc. como fondo de todas las paginas del documento, con fines generalmente estéticos. Tienen por objetivo contar con una biblioteca de sellos gráficos, de entre las cuales uno pueda elegir y aplicar al documento dado. No se aplican claves secretas ni otras formas adicionales de seguridad.

El énfasis del trabajo esta puesto en la seguridad de documentos de tipo "oficina", siendo las técnicas simples, fáciles de usar y relativamente resistente a ataques.

3. MARCAS DE AGUA (*WATERMARKING*)

Una interesante alternativa en cuanto a seguridad, pero en el contexto de proteger la propiedad intelectual de una obra, lo dan las marcas de agua (*Watermarking*) que son señales o patrones insertados en imágenes o documentos digitales para identificarlos en forma permanente e inalterable [3].

En lugar de asegurar la autenticidad o integridad de un documento, como lo harían las firmas digitales u otros mecanismos similares, las diferentes formas de *watermarking* buscan identificar ya sea el origen, autor, propietario, derechos de uso, distribuidor o usuario autorizado, de un documento digital, incluso si el mismo ha sido procesado y/o distorsionado [12].

El *watermarking* se aplica tanto a imágenes, video, audio, como a algunas formas de texto; es decir, en casi todos los formatos de archivos empleados en la actualidad, obviamente con técnicas diferentes para cada caso, pero con el mismo objetivo.

El *watermarking* también puede ser aplicado a imágenes de textos. Básicamente se conocen tres métodos:

- Codificación de línea. En este caso, las líneas de texto de un documento son desplazadas imperceptiblemente hacia arriba o hacia abajo.
- Codificación de espaciado de palabras. Aquí es alterado el espaciado entre palabras en una línea de texto justificado.
- Codificación de caracteres. Envuelve alteraciones menores a las formas de los caracteres [5].

Una persona interesada en quebrar estos mecanismos de seguridad podría hacerlo simplemente reespaciando las líneas ya sea uniforme o aleatoriamente, el espaciado entre palabras o la forma de los caracteres. Las marcas colocadas en un texto, usando cualquiera de estas técnicas pueden ser siempre removidas por el retipeo del documento. Incluso este esfuerzo puede ser automatizado mediante dispositivos de reconocimiento de caracteres [11] como *scanners* con OCR (*Optical Character Recognition*).

Las técnicas citadas son aplicadas a representaciones de documentos en forma de imágenes, que describen cada página de un documento como un arreglo de pixeles, como también a archivos de formateo de documentos. Estos últimos son archivos digitales que describen el contenido del documento y la disposición de las páginas usando lenguajes estándar de descripción de formatos como Postscript, Tex, troff, etc. [5]. Estos métodos se aplican principalmente en las publicaciones electrónicas ya que las mismas se están incrementando movidas por el costo decreciente del procesamiento computacional y la alta calidad de impresoras y monitores, a más de la mayor disponibilidad de comunicaciones de bajo costo y alta velocidad [12].

Otras aplicaciones podrían ser: la correspondencia confidencial o de distribución limitada, los reportes técnicos, y cualquier documento del que se quiera preservar la autoría, y cuyos créditos o ganancias si hubieren, quedaran así fuera de disputa.

En el caso de textos hechos con procesadores de palabras de amplia difusión, como por ejemplo las últimas versiones de Microsoft Word (7.0 en adelante), marcas de agua pueden ser insertadas en archivos, pero solo con fines estéticos, es decir, un objeto ya sea gráfico o textual puede ser impreso como fondo del documento haciendo las veces de un papel con sello de agua, pero en estos casos la marca es fácilmente removible.

Aunque el *watermarking* digital ha empezado a ser utilizado comercialmente, hay aún varias barreras que evitan que esta tecnología se vuelva efectiva y de uso difundido. La misma se halla aún en sus desarrollos iniciales. El mayor desafío técnico es desarrollar una protección segura, pero manteniendo las marcas ocultas. Robustez absoluta es prácticamente imposible, pero queda aún lugar para mejoras. Hoy en día ninguno de los sistemas de *watermarking* puede clamar que puede sobrevivir todas las operaciones de procesamientos de imágenes [11], ya que los requerimientos que deben ser satisfechos compiten entre sí y son muchas veces contrapuestos y conflictivos [3].

4. ESCRITURA ENCUBIERTA (*STEGANOGRAPHY*)

Steganography es el arte de encubrir información de manera a evitar la detección de mensajes ocultos. *Steganography* deriva del griego y literalmente significa "escritura encubierta". Incluye un vasto juego de métodos de comunicación secreta que ocultan la propia existencia del mensaje. Estos métodos incluyen tintas invisibles, micropuntos, ordenamiento de caracteres, canales encubiertos, y comunicación de espectro expandido (*spread spectrum)* [8].

La diferencia entre *watermarking* y *steganography* es básicamente la intención. La forma tradicional del segundo implicaba el ocultamiento de la información, en cambio el *watermarking* extiende la información y llega a ser un atributo del documento sellado. En *steganography*, el objeto de la comunicación es el mensaje oculto y la envoltura solo el medio de enviarlo. En *watermarking*, el objeto de la comunicación es la envoltura y el mensaje oculto solo hace referencia a esa envoltura con fines de garantizar autoría, derechos de propiedad o utilización, etc.

Como hemos visto, el *watermarking* para textos está limitado solo a ciertos formatos y son fácilmente removibles por un reformateo de los documentos. Por otro lado, *steganography* tiene otra finalidad, siendo a su vez implementado principalmente sobre imágenes.

El presente trabajo plantea una combinación de ambas técnicas en cuanto a concepto y a su aplicación a un formato no considerado por ninguna implementación anterior, de la siguiente manera:

- proveer las informaciones sobre el documento como habitualmente lo propone el *watermarking*;
- hacerlo de forma oculta, es decir que pase desapercibido para un usuario no enterado de su existencia, conforme se hace con *steganography;*
- en caso de que fuera sospechada la existencia de información oculta, que la misma no pueda ser removida con relativa facilidad; y que
- la implementación sea realizada sobre un formato comercial accesible masivamente, como el RTF (*Rich Text Format*).

5. EL SELLADOR-RECONOCEDOR

A continuación se explica en detalle las características principales de cada método, sus fortalezas y debilidades, así como comparaciones entre las mismas.

5.1 Sellos aplicados al layout

Su funcionamiento se basa en que la aplicación del sello no es fácilmente perceptible a simple vista y es relativamente trabajoso encontrarlo por una búsqueda a través de un procesador de textos como Word.

Si se quisiera buscar el sello a través de la visualización interna del documento con un editor corriente, este debería ser de buena capacidad para los casos de archivos de cierto tamaño, ya que el Notepad de Windows por ejemplo, tiene la posibilidad de editar solo archivos pequeños.

Asumiendo que la existencia del sello es desconocida por otra persona que no sea el autor, permitiría hacer un seguimiento de párrafos o documentos enteros copiados a partir de otro. La utilidad de esto esta dada no solamente por la inclusión de *watermarking,* con datos como tema, autor, propiedad, derechos de uso, o cualquier otra información que se puede incluir normalmente en un documento, sino por la posibilidad de utilizarlo para *Esteganografía,* es decir, como un canal de comunicación encubierto sin tener que recurrir a encriptarlo para volverlo ilegible y sin hacer evidente con ello que se están enviando documentos con contenidos ya sea confidenciales, informativos, de alerta o de otra índole.

Como ya se mencionó, se dispone de 2 tipos de sellos aplicados al layout:

Sellador-Reconocedor Genérico

El texto del mensaje oculto puede ser obtenido si se ubican las posiciones en las que se insertó el sello y se lo descifra conociendo el patrón de compresión-expansión empleado (que se convierte de esta manera en una especie de alfabeto de sustitución) o utilizando el Sellador-Reconocedor genérico ya que el uso de un método de criptografía simple y públicamente conocido como el aquí implementado, posee la ventaja de poder contar con un descifrador de uso masivo con el cual una persona pueda insertar un texto, mensaje o sello en cualquier documento y reconocer cualquier documento sellado por otra persona. Para estos casos, el Sellador-Reconocedor genérico constituye una alternativa práctica y de fácil utilización.

Sellador-Reconocedor Personalizado

Si el énfasis esta dado en la confidencialidad, es decir, el usuario no desea que cualquiera pueda enterarse del contenido de su mensaje encubierto, ya sea por revisión manual o a través del Sellador-Reconocedor genérico, se tiene la opción de utilizar uno Personalizado, en el cual se posee un *password* para reconocer lo insertado, de forma que solo el autor y los poseedores de ese *password* tengan acceso a ese mensaje.

Y se puede continuar volviendo al Sellador-Reconocedor tan sofisticado como se quisiera, es decir, agregándole Claves Compartidas, Claves Privadas y Publicas, etc., ya que son tres las técnicas principales componentes de esta

propuesta: *Watermarking*, *Esteganografia* y *Criptografia*, de las cuales esta última es la que posee más alternativas, siendo a su vez aplicables de acuerdo a la finalidad primaria perseguida, ya sea *Watermarking*, en la cual el objeto de la comunicación es el documento y el sello se desea que agregue información sobre dicho objeto, o *Esteganografia*, en donde el documento actúa solo como un canal o envoltorio y lo que realmente se desea transmitir es el mensaje contenido en el sello.

Experiencias y resultados

Una vez aplicado un sello a través del kit implementado, el documento resultante, en caso de que fuera editable, mantiene el sello aun si:

- se copia todo el documento a un documento vacío o a otro con texto;
- se copia un párrafo sellado a otro documento sin sello o con sello;
- se cambian tipos de letras, tamaños de caracteres, colores, etc. en el párrafo con sello;
- se insertan espacios, caracteres y textos nuevos entre medio de textos sellados;
- se sella un documento varias veces, ya sea con el mismo o diferentes sellos, es decir, permite sellos acumulativos;
- se copian párrafos sellados de 2 documentos en un tercero, en cuyo caso este tendrá los dos sellos.

De lo expuesto, se puede notar que los sellos son bastante resistentes y confiables y que la única forma de removerlos sin afectar el contenido del documento es aplicando un espaciado entre caracteres ya sea uniforme o arbitrario. Esto podría ser realizado a través de las funcionalidades propias del Word o se podría también editar el documento con un editor común y remover todos los espaciados que no correspondan al *default*, pero esta es una alternativa muy trabajosa.

Otros mecanismos complementarios de protección

Si se quisiera un nivel mayor de seguridad, se podría evitar que un documento pueda ser modificado, a través de la funcionalidad de "Proteger Documentos" que provee el Word de forma estándar. De esta manera, el documento no puede ser alterado, editado, borrado, etc., e incluso al regrabar un archivo de solo lectura con otro nombre, el mismo conserva las características de protección del original, con lo cual no se podría modificar ninguno de los dos si no se conociera el *password* correspondiente.

Sin embargo, en el caso del Microsoft Word, la debilidad está dada por el hecho de que al guardar el nuevo archivo se puede elegir el formato del mismo, pudiendo utilizarse por ejemplo el RTF, que conserva todas las características del formato .DOC, excepto la seguridad, que es totalmente perdida al grabar y volver a abrir el documento, por lo que el procesador de textos debería tener una opción de protección que impida el regrabado de un archivo protegido.

Esta vía hace que la seguridad sea casi nula en este procesador de textos, siendo el password para la apertura de archivos la única que funciona, pero vuelve inaccesible al documento, por lo que no es útil cuando lo que uno desea es publicar un texto y que a la vez el mismo permanezca *congelado* como seria el caso de una publicación distribuida por medios electrónicos.

5.2 Sellos aplicados al archivo

El texto asociado a cada documento es definido por el usuario. Normalmente incluiría por lo menos el nombre del autor y adicionalmente el del destinatario de la copia autorizada si lo quisiese individualizar, los derechos de uso y cualquier información adicional que se quisiera agregar hasta el límite establecido para esta versión, de 70 caracteres.

Para la implementación computacional del '*Sellador Digital Encubierto*' se incluye en el cuerpo del documento una señal que identifica al autor, además de otras informaciones adicionales, con las siguientes características:

- no puede ser detectado por simple inspección del documento cuando se lo visualiza con algún paquete de software aplicativo como Word, Excel, etc.;
- no es removible fácilmente con cambios menores al documento (obviamente, si se re-escribe todo el documento, la nueva versión no estará sellada);

- no permite que un intruso pueda utilizar el método del 'Texto Plano Elegido' [8], para encontrar el patrón seguido por el '*sellador*' original. En otras palabras, un intruso no puede utilizar su propio '*sellador*' para encontrar el patrón de un sello ajeno. Para esto, la forma de sellar es variable de un usuario a otro, según sea su clave secreta;
- es suficientemente complicado como para ser detectable por inspección interna de la codificación del documento, de forma a evitar que un intruso pueda detectar y/o remover un '*Sello*';
- es bastante difícil crear un "Removedor de Sellos" que automáticamente remueva (borre) todos los sellos existentes en un sistema de documentos;
- es suficientemente complejo, combinando parámetros y métodos diferentes, de forma a evitar la probabilidad de falsificaciones por fuerza bruta [10];
- los datos que se agregan adicionalmente y que constituyen el '*Sello*' no aumentan excesivamente el tamaño del archivo original;

Sin embargo, el método tiene sus limitaciones, por ejemplo si una persona supiera de la existencia del sello en un documento, podría editarlo, borrar su contenido y reescribir otro texto. El problema aquí es que se obtendría un sello ajeno, a disposición de ser utilizado por alguien que no es el verdadero propietario de dicho sello. Esto se supera guardando también el tamaño del archivo original y otros datos mas, o incluso la cadena resultante de aplicar un Message Digest, con lo cual el reconocedor alertaría que el sello no coincide con el archivo original al cual fue aplicado.

Otra forma de detectar el sello seria dejar el documento en blanco, solo con el sello y crear otro documento en blanco propio, obviamente sin sello, editar ambos y las diferencias encontradas serian el sello.

Otra característica es que si se copia solo uno o varios párrafos, el sello no se transporta.

5.3 Sellos Visibles

La información inserta automáticamente puede ser el logotipo de la empresa, un slogan, una dirección web, de correo electrónico, etc. La ventaja es que este dato se encuentra inserto una sola vez, pero es visible y accesible desde cualquier pagina.

Permite tener siempre visible la Marca del propietario, tal como hacen los canales de televisión a fin de que si sus imágenes son tomadas queden claros los derechos intelectuales o de propiedad, pero siendo editable, el sello puede ser removido, por lo que su finalidad es básicamente estética más que de seguridad.

5.4 Comparaciones

En el sello aplicado al archivo se requiere que el usuario del programa tenga archivos donde almacene sus mensajes ya que lo que se inserta en el documento es solo una clave de acceso (o string corto) a la base de datos donde realmente se encuentran los mensajes.

Como ejemplo, se podría tomar el tamaño de los memorándums o notas de oficina, que normalmente estarían entre unas pocas decenas de Kbytes, mientras los códigos adicionados como parte de los sellos a su vez estarían en alrededor de unas pocas decenas de bytes, lo cual hace que el crecimiento sea en estos casos inferior al 0,1 %. Cuanto más grandes sean los archivos, menor será la relación de overhead, ya que el mismo no es proporcional al tamaño del archivo.

En el caso aplicado al layout el overhead es de aproximadamente 20 bytes por cada carácter insertado, con lo cual el archivo puede crecer tanto como la longitud del mensaje inserto o la redundancia que el usuario quisiera agregar, por ejemplo haciendo que todos los párrafos o la mayoría estuviesen sellados.

Otra característica importante de esta alternativa es que permite que el mensaje se encuentre autocontenido en el documento, es decir, no se precisa de la base de datos auxiliar con los mensajes del autor para poder reconocerlo.

Los sellos visibles son mas bien estéticos y mientras persistan las debilidades mencionadas, constituyen solamente una facilidad operativa y, para el caso presente, ilustran una de las modalidades de sellados existentes, junto con su aplicabilidad en un entorno de oficina.

6. EL PROGRAMA

Básicamente el programa se constituye en un kit de Sellados, con prestaciones y utilidades diversas, enfocadas en distintos aspectos de la técnica del Marcado de Agua. La figura 2 muestra un esquema de los tipos de sellados, con sus módulos componentes y las opciones que provee cada uno de ellos.

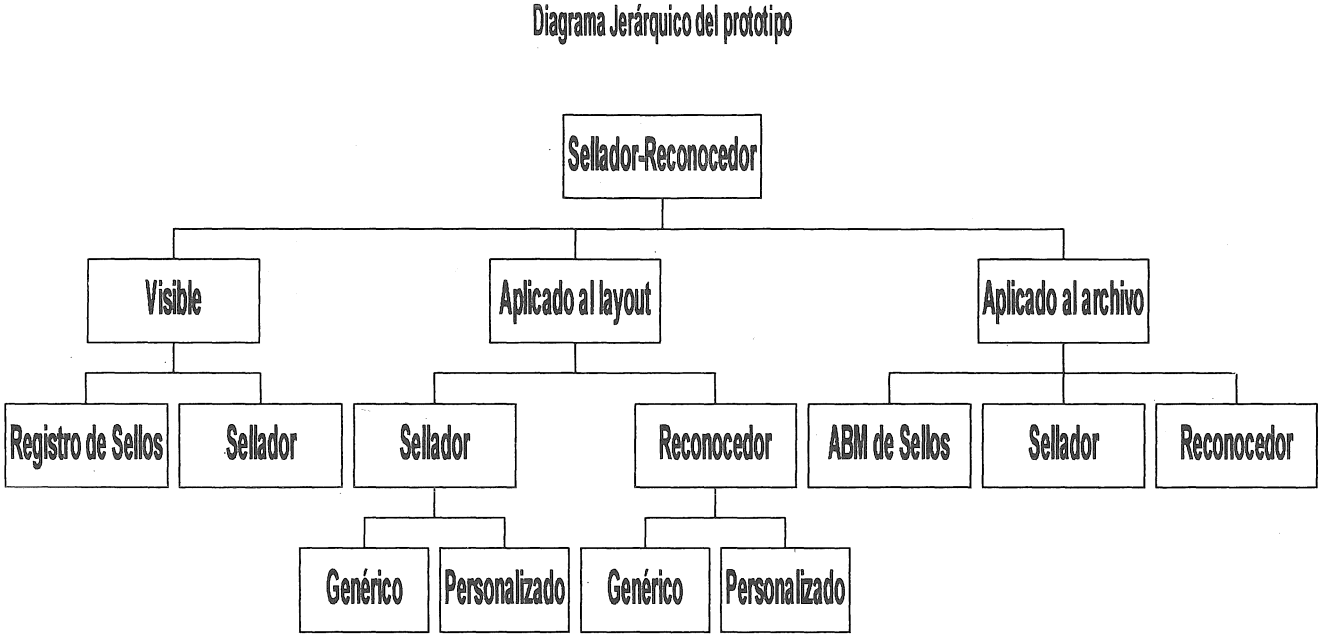

Figura 2. Diagrama jerárquico de opciones del prototipo

A continuación se presentan pseudocodigos resumidos de las lógicas de los sub programas selladores, pudiendo entenderse intuitivamente a partir de estos, la lógica de los reconocedores.

```
Escribir mensaje a insertar.
Elegir documento a sellar.
Verificar que el documento tenga por lo menos una línea completa de texto o sea de longitud mayor que el mensaje
que se quiera insertar.
Recorrer la codificación interna del documento hasta encontrar el primer lugar disponible y libre para insertar el
sello.
Si ya hubieran mensajes previos, agregar a continuación del último mensaje, para lo cual
        Mientras hayan letras del mensaje a insertar
                Tomar letra del mensaje
                Buscar los comandos de formateo correspondiente a esa letra según el alfabeto de sustitución
                Insertar dichos comandos previos al carácter del párrafo que se modificará.
                Ir al siguiente carácter del párrafo
        Fin Mientras
Regrabar el documento con los formateos incluidos.
```

Pseudocodigo 1. Sellos aplicados al layout

```
Si el sello aun no fue registrado,
        Dar de alta el nuevo sello o mensaje en la base de datos de sellos, siendo su identificador tomado de acuerdo
        a la clave introducida por el autor al inicio de la sesión y habilitada al instalar el programa.
Elegir el sello  a utilizar.
Elegir el documento a aplicar.
Verificar que no exista un sello previo.
Si existe sello previo del mismo autor
        Avisar, con opción a
                Reemplazar el sello anterior,
                Mantener el sello existente o
                Eliminar el sello existente.
Si  existe algún otro sello ajeno
        Avisar y no permitir la sobreescritura.
Si no estuviera aun sellado
        Buscar, los comandos de formateo entre los que se insertara el identificador de acceso a la base de datos de
        sellos.
        Insertar este identificador.
        Regrabar el documento.
```

Pseudocodigo 2. Sellos aplicados al archivo

En caso de que el autor de un documento deseara aplicarle otro sello, por ejemplo para distribuir múltiples copias de un documento, identificando en cada sello a los receptores autorizados, podría hacer las correspondientes copias primeramente y aplicarle luego el sello diferenciado a cada copia.

```
Sellar un documento manualmente por medio de los comandos del procesador de texto.
Entrar al programa Sellador-Reconocedor.
Indicar cual es el documento que posee el sello
Extraer el sello de dicho documento y almacenarlo en la biblioteca de sellos.

Posteriormente, cuando se desee sellar otro documento con el mismo sello registrado.
        Ir a la opción de sellado visible
        Indicar el documento a sellar
        Elegir el sello visible a aplicar
        Insertar el sello
        Regrabar el documento.
```

Pseudocodigo 3. Sellos Visibles

La versión actual esta desarrollada en Visual Basic 5.0 y no tiene limitaciones en cuanto al tamaño de los documentos a sellar, lo que era un inconveniente en el prototipo inicial, desarrollado en Clipper 5.0.

7. CARACTERISTICAS RESALTANTES

Entre las principales características de las técnicas aquí presentadas, se puede mencionar que son nuevas en cuanto a sus concepciones e implementaciones en los siguientes aspectos:

Como ya se mencionó, en el *watermarking* aplicado a imágenes de textos existen tres métodos propuestos: codificación de línea, codificación de espaciado de palabras y codificación de caracteres. El presente trabajo implementa otras variantes: espaciado entre caracteres, insertado en el archivo y watermarking visible.

Las técnicas conocidas en el *watermarking* tradicional, y que se mencionaron previamente, también se pueden agregar al programa pero esto ya seria con fines ilustrativos, ya que no agregaría nada sustancial a lo ya desarrollado.

La mayoría de los métodos publicados hasta la fecha requieren del documento no sellado. Y a partir de la comparación con el mismo se detecta la información encubierta contenida. En esta propuesta no se precisa contar con los documentos originales.

Generalmente se basan en procesamiento de imágenes de textos, ya sea resultantes de conversiones a partir de los originales o escaneados a partir de copias impresas. En este trabajo se consideran los archivos en sus formatos originales, es decir, como textos no como imágenes.

Existen trabajos publicados que aplican las técnicas de marcado a formatos PostScript. Sin embargo en el presente proyecto el formato utilizado es el RTF, reconocido por procesadores como el Microsoft Word y otros, es decir, de uso mucho mas extendido, principalmente en entornos comerciales o empresariales.

Comúnmente el watermarking existente tiene algún componente de Esteganografia. En este caso, las técnicas presentadas son aplicables tanto para watermarking puro como para fines netamente esteganográficos, o combinaciones de ambos.

Una limitación encontrada ya sea con este método como con los demás de *watermarking*, es que todas las marcas son removidas si se convierten los archivos a formato texto (TXT), pero en este caso se pierden además todos los formateos, *fonts*, *layouts*, etc.

8. CONCLUSIONES

Los principales objetivos alcanzados con este proyecto son:

- Reconocer como propios los documentos sellados de un autor;
- Identificar copias no autorizadas que se hubieran hecho de sus documentos sellados;
- Detectar en ciertos casos que un archivo contiene partes de un documento sellado pertenecientes a un autor;

Asimismo, existe un gran potencial de aplicaciones en las oficinas automatizadas modernas (que manejan principalmente archivos computacionales en lugar de los tradicionales documentos impresos en papel) y en las comunicaciones electrónicas utilizando Internet.

En el presente proyecto, no se pretende todavía elaborar un mecanismo de 'Seguridad Máxima' ya que para ello haría falta una mayor cantidad de recursos computacionales y una compleja combinación de diversos métodos y tecnologías [1].

En consecuencia, se presentó la implementación de un prototipo aplicable a situaciones de uso cotidiano donde una 'Seguridad Suficiente' satisface las necesidades de los usuarios al desalentar la utilización indebida de documentos no autorizados, y cubre un área no considerada hasta ahora pero que es mayoritaria, como ser el mundo de las aplicaciones comerciales para computadoras personales y redes de computadoras que no operan en ambientes colaborativos, tales como Lotus Notes y otros aplicativos similares [2].

Si bien el objetivo inicial del trabajo fue la utilización de *watermarking* para preservar derechos autorales, el prototipo desarrollado va un poco mas lejos al permitir un uso más flexible, que abarca incluso a la *esteganografía*, satisfaciendo también los alcances de esta disciplina.

Entre los trabajos futuros se encuentra el insertar directamente las funcionalidades desarrolladas al Microsoft Word, en forma de Macro programada en *Visual Basic for Aplication* y disponibles ya sea como opciones de menú de este programa o como botones agregados a ciertas cajas de dialogo estándar, con lo cual se integraría el programa Sellador-Roconocedor con el procesador de texto, aunque siempre podrá también ser ejecutado por separado.

También de esta manera se podría generalizar los sellos para otros formatos usados por Word, ya que si se cambia el formato del documento a otro, por ejemplo .DOC, y luego se lo vuelve a reconvertir a RTF, los sellos permanecen en el documento. Vale decir, que una forma práctica de utilizar el kit existente para cualquier formato sería grabar el documento inicialmente como RTF, sellarlo y regrabarlo como .DOC. Cuando se quisiera reconocer el sello, se podrían convertir todos los archivos de interés a formato RTF y allí aplicar el reconocedor. Estos pasos se estarían automatizando en la versión programada con Visual Basic for Aplication. Para el prototipo actual se consideran exclusivamente los archivos que se encuentren en el formato RTF, tanto al sellar como al reconocer.

Otro de los formatos considerados para trabajos posteriores es el HTML, utilizado en internet, ya que su estructura y formato son similares a los utilizados en este proyecto.

REFERENCIAS

[1] G. Aldegani., "La amenaza del Hacking" , *Compu Magazine*, Ago-97.

[2] J. Bach, "Good Enough Quality : Beyond the Buzzword", *IEEE Computer* , Ago-97.

[3] H. Berghel, L. O'Gorman, "Protecting ownership rights through digital watermarking", *IEEE Computer* , Jul-96.

[4] H. Berghel, "Watermarking Cyberspace", *ACM Communications* , Nov-97.

[5] J. Brassil, S. Low, N. Maxemchuk, L. O'Gorman, "Electronic Marking and Identification Techniques to Discourage Document Copying", *IEEE Journal on Selected Areas in Communications,* Oct-95

[6] D. Chadwick , A. Young , "Merging and Extending the PGP and PEM Trust Models - The ICE-TEL Trust Model", *IEEE Network* , May/Jun-97.

[7] S. Garfinkel, "Public Key Cryptography", *IEEE Computer*, Jun-96.

[8] N. Johnson, S. Jajodia, "Exploring Steganography: Seeing the Unseen", *IEEE Computer,* Feb-98

[9] P. Neuman,"Crypto Key Management", *ACM Comnunications*, Ago-97.

[10] A. Tanenbaum, *"Computer Networks"* , Prentice Hall PTR New Jersey - 1996.

[11] S. Low, N. Maxemchuk, "Performance Comparison of Two Text Marking Methods", *IEEE Journal on Selected Areas in Communications,* May-98

[12] J. Zhao, "Look, It's Not There", *Byte*, Ene-97.