DESDE 2013
https://repository.uaeh.edu.mx/revistas/index.php/icbi/issue/archive
Pädi Boletín Científico de Ciencias Básicas e Ingenierías del ICBI
ISSN: 2007-6363

Publicación Semestral Pädi Vol. 11 No. Especial 2 (2023) 175-182

Autenticación de audio digital mediante LSB mejorado y mapa logístico
Digital audio authentication using enhanced LSB and logistic map

M. T. Quiñonez-Carbajal[a,*], R. Reyes-Reyes[a], V. Ponomaryov[a], C. Cruz-Ramos[a]

[a] *Instituto Politécnico Nacional, ESIME Culhuacan, Av. Sta. Ana 1000, San Fco. Culhuacan, 04440, CDMX, México.*

Resumen

Con el avance en las comunicaciones se han desarrollado procesos para la protección del contenido digital y su transmisión segura utilizando múltiples herramientas de procesamiento de información que incluyen técnicas de marcas de agua, las cuales insertan información en una señal portadora sin degradar su calidad, permitiendo verificar la integridad del contenido multimedia y la detección de manipulaciones, a partir de los cambios detectados en la marca de agua extraída. Este articulo propone un algoritmo de marca de agua frágil para la autenticación de audio digital, mediante la inserción de una marca de agua empleando la técnica del bit menos significativo (LSB) a partir del análisis de los dos primeros bits más significativos (MSB). Los resultados obtenidos demuestran que la marca de agua es imperceptible, obtenido valores promedio de SNR y NC de 43.22 dB y 0.9980 respectivamente; adicionalmente, el método propuesto es resistente a diversos ataques como muteo, clonación y sustitución de muestras.

Palabras Clave: Marca de agua digital, autenticación de audio, LSB, detección de alteraciones, mapa caótico logístico.

Abstract

With the advancement of communication technologies, various processes have been developed to protect digital content and ensure its secure transmission. These processes involve the utilization of multiple information processing tools, including watermarking techniques. Watermarking allows for the embedding of information into a carrier signal without compromising its quality. It enables verification of the multimedia content's integrity and facilitates the detection of manipulations by analyzing changes in the extracted watermark. This article presents a fragile watermarking algorithm designed for digital audio authentication. The algorithm employs the least significant bit (LSB) technique and analyzes the first two most significant bits (MSB) for embedding the watermark. The results obtained demonstrate the imperceptibility of the watermark, with average signal-to-noise ratio (SNR) and normalized correlation (NC) values of 43.22 dB and 0.9980, respectively. Additionally, the proposed method exhibits resilience against various attacks such as mute, cloning, and sample substitution.

Keywords: Digital watermarking, audio authentication, LSB, tampering detection, logistic chaotic mapping.

1. Introducción

Con el crecimiento exponencial de nuevas herramientas computacionales que se pueden encontrar fácilmente en Internet o en las tiendas de aplicaciones para dispositivos móviles, existe la problemática de que hoy en día casi cualquier persona pueda manipular señales de audio digital de manera sencilla, sin dejar ningún rastro en las regiones alteradas provocando que los archivos de voz puedan pasar como auténticos. Dichas alteraciones pueden llevarse a cabo con el fin de hacer daños sociales, económicos o hasta legales. La necesidad de autenticar archivos de audio ha llevado a investigar y desarrollar esquemas para la detección de falsificaciones; esto ha derivado en el desarrollo de técnicas de marca de agua digital, que consisten en insertar un mensaje ya sea visible o invisible en un medio portador, que permitirá su futura autenticación. Estas técnicas son comúnmente utilizadas en aplicaciones como el seguimiento de transacciones (también conocidas como huellas digitales), prueba de propiedad, control de copia, entre otras (Cox *et al*., 2002). Por lo que, la implementación de modelos de marca de agua en esquemas de protección de información digital es un enfoque tradicional que se ha utilizado para abarcar procesos de autenticación de contenido.

Recientemente, dentro del campo de sistemas de marcas de agua en audio digital se tienen diferentes propuestas para su

*Autor para la correspondencia: mquinonezc1500@alumno.ipn.mx
Correo electrónico: mquinonezc1500@alumno.ipn.mx (María Teresa Quiñonez-Carbajal), rreyesre@ipn.mx (Rogelio Reyes-Reyes), vponomar@ipn.mx (Volodymyr Ponomaryov), ccruzra@ipn.mx (Clara Cruz-Ramos).

Historial del manuscrito: recibido el 18/03/2023, última versión-revisada recibida el 22/06/2023, aceptado el 28/06/2023, publicado el 11/09/2023. **DOI:** https://doi.org/10.29057/icbi.v11iEspecial2.10722

autenticación. Los algoritmos de marca de agua en el domino temporal tienen una mayor capacidad de inserción y un procesamiento más rápido, pero son menos robustos frente al ruido y ataques. El LSB convencional es uno de los enfoques más antiguos y sencillos utilizados para transmitir datos secretos a través de medios digitales. La capacidad de inserción de esta técnica es directamente proporcional a la frecuencia de muestreo del audio de cobertura. Por ejemplo, si la frecuencia de muestreo es de 44 kHz, la capacidad de inserción será de 44 kbps (kilobits por segundo) (Ahmed, *et al*., 2010). Este enfoque puede aplicarse fácilmente en la comunicación de audio en tiempo real y también puede fusionarse con otros enfoques de seguridad de datos digitales. Sin embargo, este método es muy sensible a ataques como la adición de ruido, la eliminación de LSB, la compresión, la amplificación, el remuestreo, etc.

Los investigadores han propuesto varios enfoques LSB modificados para reducir las limitaciones del enfoque LSB convencional. (Asad *et al*., 2011) presentan dos técnicas esteganografías para la ocultación de información en los LSB de las muestras de un audio que la hace más robusta al esteganálisis. La técnica consiste en aleatorizar el número de bits del mensaje anfitrión utilizado para insertar el mensaje secreto, la segunda técnica consiste en aleatorizar el número de muestras que contiene el siguiente bit del mensaje secreto. Esto lo asigna a través de los Bits Mas Significativos MSB, sin embrago al modificar los últimos LSB lo hace frágil ante ataques de compresión de información. (Chetan *et al*., 2021) proponen un método de marca de agua en audio digital basado en la técnica LSB, se utiliza una imagen como marca de agua donde los bits de la imagen se ocultan en el último LSB de cada muestra de la señal de audio portadora, sin embargo, este sistema no es ciego puesto que necesita información adicional para poder recuperar la marca de agua. (Ghobadi *et al*., 2013) presentan un algoritmo que utiliza el método LSB. Se trata de una técnica de marca de agua de audio ciega para la detección y prevención de manipulaciones. Inicialmente, el archivo de audio se convierte en un vector que se normaliza y se convierte en una matriz. Después de calcular la media de cada columna por separado, se inserta utilizando el método LSB, insertando la marca de agua en el último bit menos significativo. Sin embargo, no realizan la detección de manipulaciones y su ubicación correspondiente.

Existen propuestas en el estado del arte que fusionan la técnica LSB en conjunto con el mapa logístico, para brindar mayor seguridad a la marca de agua. (Li et al., 2017) presentan un método de marca de agua frágil utilizando el mapa logístico e insertando la marca de agua en los bits LSB. (Wang *et al*., 2016) proponen un método de marca de agua, que utiliza el mapa logístico para aleatorizar a la señal, utilizan los bits de numero de bloque como marca de agua, insertándolos directamente en los valores de las muestras de la señal de audio después de realizar una cuantización. Su sistema detecta el área manipulada; sin embargo, no presentan métricas de SNR.

Los métodos del estado del arte analizados anteriormente presentan algunas desventajas. Las técnicas de autenticación basadas en la modificación de los últimos bits menos significativos de la señal portadora presentan el inconveniente de ser fácilmente detectados y manipulados por los atacantes siguiendo un patrón mediante esos bits. En el caso de las propuestas que fusionan las técnicas LSB en conjunto con el mapa logístico, presentan la misma problemática; ya que independientemente de insertar una marca de agua aleatoria siguen insertando la información en los mismos LSB.

Teniendo en cuenta los enfoques y desventajas mencionados anteriormente, se resumen las principales aportaciones de la propuesta para la autenticación y detección de manipulaciones en señales de audio: (1) El sistema propuesto es completamente ciego, ya que la información de autenticación de audio se extrae directamente de la señal con marca de agua, sin necesidad de información adicional o de un archivo externo. (2) La aleatoriedad de la información utilizando el mapa caótico logístico, así como la selección de bits LSB para la inserción de los bits de marca de agua en regiones diferentes de acuerdo a los bits MSB, permite que el sistema pueda detectar el área alterada inclusive si dos bloques o más consecutivos contienen silencios. (3) el sistema propuesto demuestra una buena imperceptibilidad de la marca de agua, obteniendo valores promedio de SNR superiores al estado del arte.

El resto del artículo se organiza de la siguiente manera. En la sección 2 se describen las características de las marcas de agua. El esquema detallado para la inserción de la marca de agua en la señal y su autenticación se describe en la sección 3. La sección 4 presenta las pruebas, resultados experimentales y el análisis de rendimiento comparado con otros métodos relacionados. Por último, concluimos nuestro trabajo en el apartado 5.

2. Marcas de agua

Las marcas de agua son comúnmente utilizadas para la autenticación de las señales de voz. Existen tres características generales que engloban el comportamiento de un esquema de marca de agua digital (Naskar *et al*., 2014):

- *Imperceptibilidad:* se basa en el comportamiento del sistema perceptivo humano. Una marca de agua es imperceptible, si la degradación que se produce en los archivos donde ha sido embebida es muy difícil de apreciar.
- *Capacidad:* se refiere al tamaño máximo que pude ocupar la marca de agua al ser embebida en el medio portador.
- *Robustez:* es la capacidad de persistencia de la marca de agua, que debe permanecer tras ataques intencionales (adición, edición o eliminación de contenido) o no intencionales (ruido digital, conversión de formato, compresión, etc.).

Sin embargo, el intentar mejorar uno de estos factores para un esquema de marca de agua generalmente deteriora a alguno de los otros. Por lo tanto, el comportamiento de cualquier técnica de marca de agua puede describirse visualizando un triángulo en el cual cada vértice corresponde a una de las tres características fundamentales (véase la Figura 2). Además,

Figura 1: Diagrama de bloques general de la técnica propuesta.

tales características ahora son requerimientos correlacionados que siempre deben existir.

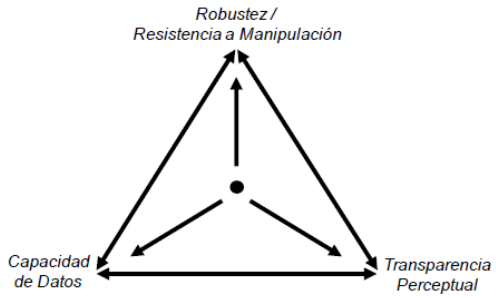

Figura 2: Relación entre los requerimientos clásicos de una marca de agua digital (Naskar *et al*., 2014).

En esta analogía, la posición dentro del triángulo representará el peso que se le ha dado a cada uno de los tres requerimientos, así, cuanta más importancia haya sobre uno, el rendimiento sobre los demás decaerá.

Las técnicas de marca de agua digital pueden categorizarse a partir de diferentes criterios, uno de ellos es de acuerdo con su reacción a los ataques maliciosos o no maliciosos. Existen tres tipos de marcas de agua; la robusta tolera los ataques que se le realicen, las frágiles son sensibles a cualquier modificación ya que estas son utilizadas principalmente para verificar la integridad de una señal, mientras que las semi frágiles toleran ciertos ataques, pero se destruyen ante ataques no reversibles (Vargas *et al*., 2016).

El mecanismo de extracción de una marca de agua puede ser ciego, que se refiere a que no es necesario la información original para determinar su autenticidad y los métodos semiciegos son los que requieren algunos datos originales para recuperar la marca de agua (Jeruchim, 1984).

En general existen dos métodos para la inserción de la marca de agua; el dominio temporal y el dominio frecuencial, en el domino temporal la información de marca de agua es insertada en los datos digitales sin aplicarles ningún procesamiento, en este dominio las técnicas desarrolladas para audio incluyen el bit menos significativo, la codificación de paridad, la ocultación de eco, la codificación de fase y el intervalo de silencio. Entre estas técnicas la ocultación de datos en los bits menos significativos (LSB) tiene como principal ventaja una alta capacidad de inserción de datos de información confidencial. Además, el método LSB tiene una baja complejidad computacional (Alwahbani *et al.*, 2018).

3. Sistema Propuesto

El diagrama de bloques general del sistema propuesto se muestra en la Figura 1, en donde se observa que se inserta la marca de agua $s[m]$ en una señal de audio $A[i]$; posteriormente el audio protegido $A'[i]$ puede ser compartido por un medio de transmisión. Para la verificación de la integridad del audio, la marca de agua se extrae de la señal de voz recibida para verificar la autenticidad del audio y en su caso localizar las modificaciones en la señal.

3.1. Mapa Caótico Logístico

En la teoría del caos (Sathishkumar *et al.*, 2011) se describe el comportamiento de ciertos sistemas dinámicos no lineales que, bajo condiciones específicas, muestran una dinámica sensible a las condiciones iniciales. En este trabajo, es empleado el mapa caótico logístico de 1D para producir la secuencia caótica y controlar el proceso de cifrado de la marca de agua.

3.2. Mapa Logístico

El mapa logístico es uno de los mapas caóticos unidimensionales más simples con un comportamiento caótico complejo. Se define mediante (1).

$$x_{n+1} = \mu x_n(1 - x_n), \quad (1)$$

donde μ es el parámetro del control con un rango de $\mu \in (0,4]$ y x_n es la secuencia caótica de salida (Park *et al.*, 2017).

3.3. Inserción de la Marca de Agua

En esta sección se describe el proceso de inserción de una marca de agua frágil en una señal de audio digital. El diagrama a bloques de este proceso se muestra en la Figura 3.

Sea $A[i]$, $i = 1,2,3, \dots l$ una señal de audio original, donde l representa el número total de muestras del audio digital.

La marca de agua está compuesta por 'ceros' y 'unos' $s[m] = 11110000$, con una longitud de $l_{s[m]} = 44100$.

Paso 1. $A[i]$ se divide en B_j bloques no traslapados de n muestras cada uno, donde j es el $j - ésimo$ bloque.

Paso 2. Los bits de marca de agua $s[m]$, $(1 \leq m \leq l_{s[m]})$ son permutados utilizando el mapa caótico logístico. Se denota $X_i = \{x_i | i = 1,2, \dots, l_{s[m]}\}$ como una secuencia pseudoaleatoria después de ser permutada. Los elementos de X son generados mediante (2).

$$x_{i+1} = \mu x_i(1 - x_i), \quad x_o = k, \quad 3.5699 \leq \mu \leq 4, \quad (2)$$

donde k es el valor inicial y llave del sistema propuesto.

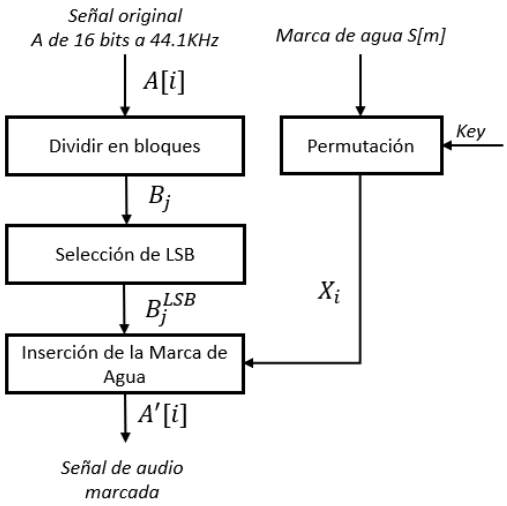

Figura 3: Diagrama a bloques de la inserción de la marca de agua.

Paso 3. Cada muestra de $A[i]$ de cada B_j se convierte a binario de 16 bits, reservando el primer bit MSB para el bit de signo.

Paso 4. Posteriormente se seleccionan los bits LSB para el ocultamiento de los bits de marca de agua permutados X_i, como se describe en la sección 3.4.

Paso 5. Una vez insertada la marca de agua en la señal portadora se obtiene la señal de audio marcada $A'[i]$.

3.4. Selección de los Bits LSB

Para confundir al intruso o atacante, se recomienda no utilizar siempre los mismos bits LSB de una muestra para insertar la información de marca de agua. Por lo tanto, se propone que a partir de los dos primeros bits más significativos (MSB) de una muestra, se determina en qué bits LSB de la misma muestra serán insertados los bits de la marca de agua. De esta manera la aleatoriedad se produce seleccionando bits diferentes en cada muestra para ocultar la marca de agua. La Tabla 1 describe la asignación de los bits LSB para el ocultamiento de la marca de agua, a partir de los valores de los bits MSB.

Tabla 1: Selección de bits para ocultamiento en cada muestra de audio

1er MSB	2do MSB	Bits de ocultamiento seleccionados	
0	0	10mo LSB	12vo LSB
0	1	11vo LSB	13ro LSB
1	0	10mo LSB	12vo LSB
1	1	11vo LSB	13ro LSB

De la Tabla 1 se puede observar que, si los dos primeros MSB de una muestra son iguales a 00 ó 10, el décimo y doceavo LSB se sustituirán por el bit de marca de agua. Si los dos primeros MSB son iguales a 01 ó 11, se sustituirán el onceavo y décimo tercero LSB por el bit de marca de agua.

3.5. Extracción de la Marca de Agua y Autenticación

En esta sección se describe el proceso de extracción y autenticación de una señal de audio digital. El diagrama a bloques de este proceso se muestra en la Figura 4. A continuación se describen estos pasos de manera detallada:

Denota $A^r[i]$ como una señal de audio marcada recibida.

Paso 1. $A^r[i]$ se divide en bloques con n muestras cada uno, donde cada $j-ésimo$ bloque se denomina como B_j^r.

Paso 2. Las muestras de cada B_j^r bloque se convierten a binario de 16 bits reservando el primer MSB para bit de signo.

Paso 3. De acuerdo con los dos primeros MSB se seleccionan los LSB para la extracción de los bits de marca de agua $B_j^{r\ LSB}$ de cada muestra de audio, tal como se describió en la sección 3.4, para con estos reconstruir la marca de agua permutada $s_p'[m]$.

Paso 4. Los bits extraídos de marca de agua permutada $s_p'[m]$ son reordenados utilizando la permutación inversa descrito en el paso 2 de la sección 3.3, empelando la misma llave de usuario k.

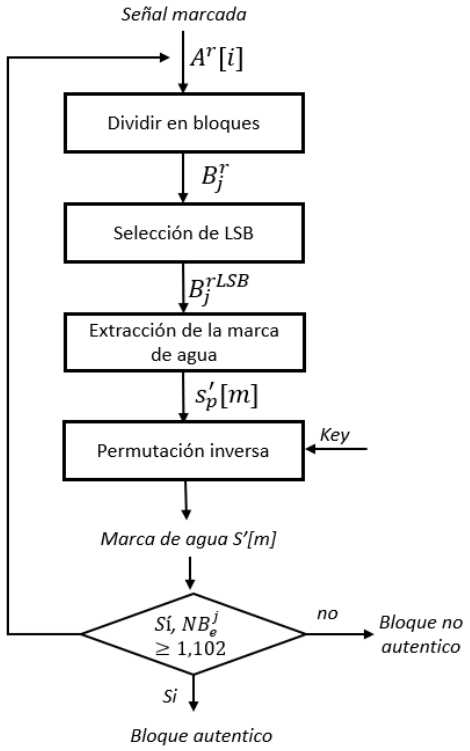

Figura 4: Diagrama a bloques de la extracción de la marca de agua.

Paso 5. La duración de cada bloque extraído B_j^r representa 0.5 segundos del audio total, tiempo en los que en ocasiones no es posible detectar cambios en el contexto de audio. Para determinar si un bloque ha sido modificado o no, se emplea la condición mostrada a continuación:

Si el *j-ésimo* bloque analizado tiene un numero de bits erróneos (NB_j^e) en la marca de agua extraída menor al 5% del tamaño total, es decir $NB_j^e < 1{,}102$, se considera que el bloque es autentico; en caso contrario, el bloque se considera alterado.

4. Pruebas y Resultados

Para realizar las pruebas, se utilizaron 100 señales de audio. Las características de los audios digitales utilizados son: Frecuencia de muestreo, $Fs = 44.1\ KHz$, bits de cuantización $Q = 16$ bits con una duración de 5 segundos, en formato .wav, tomadas de la base de datos (THCHS-30 speech library). Los parámetros que se utilizaron para este trabajo son los siguientes: $l = 220{,}500,\ j = 10,\ n = 22050,\ k = 0.068,\ \mu = 3.9728$.

Todas las simulaciones fueron ejecutadas en un sistema operativo Windows 10 utilizando la plataforma de MATLAB R2022b.

Para evaluar la imperceptibilidad auditiva de la señal con marca de agua, se calcula la relación señal a ruido (SNR) (Kabal, 2002) empleando (4).

$$SNR = 10 \cdot \log_{10}\left[\frac{\sum_{n=1}^{N} x(n)^2}{\sum_{n=1}^{N} |x(n)-y(n)|^2}\right], \quad (4)$$

donde $x(n)$ es la señal de voz original, y $y(n)$ es la señal de audio marcada.

El valor SNR indica la cantidad de distorsión producida por la inserción de la marca de agua en la señal de audio original. Los valores SNR superiores a 40 dB indican que la calidad de la señal de audio con la marca de agua se conserva, mientras que los valores SNR inferiores a 20 dB denotan una señal de audio ruidosa.

El coeficiente de correlación normalizado compara la similitud entre la señal original y la señal marcada mediante (5). Si el valor NC se aproxima a 1 $A[i]$ es muy similar a $A'[i]$, de lo contario si son muy diferentes el valor se aproxima a 0.

$$NC = \frac{\sum_i A[i]A'[i]}{\sum_i A^2[i]}, \quad (5)$$

Para evaluar la calidad de la marca de agua extraída, se utiliza la tasa de error de bits (BER), la cual se calcula mediante (6), este parámetro indica el número de bits erróneos recibidos con respecto al total de bits enviados durante un intervalo de tiempo (Jeruchim, 1984).

$$BER = \frac{bits\ erróneos}{Total\ de\ bits}. \quad (6)$$

En la Tabla 2 se muestran los resultados obtenidos en SNR y NC para diferentes señales de audio, donde se modificaron los LSB 10, 12, 11 y 13, tomando en consideración la Tabla 1 para insertar la marca de agua de acuerdo con los dos primeros MSB de cada muestra.

Tabla 2: Resultados SNR de las diferentes señales de audio marcadas.

Archivo de audio	SNR (dB)	NC
1	42.18	0.9978
2	42.09	0.9979
3	43.27	0.9980
4	44.44	0.9982
5	44.16	0.9982
Promedio	43.22	0.9980

En la Tabla 2, *Archivo de audio* es la señal marcada evaluada, SNR es la relación señal a ruido entre la señal de audio original y la señal de audio marcada y NC es la correlación normalizada entre la señal original y la señal con marca de agua.

El método propuesto es evaluado con diferentes señales de audio digitales, la Tabla 3 presenta una comparación del método propuesto con respecto al estado del arte, obteniendo

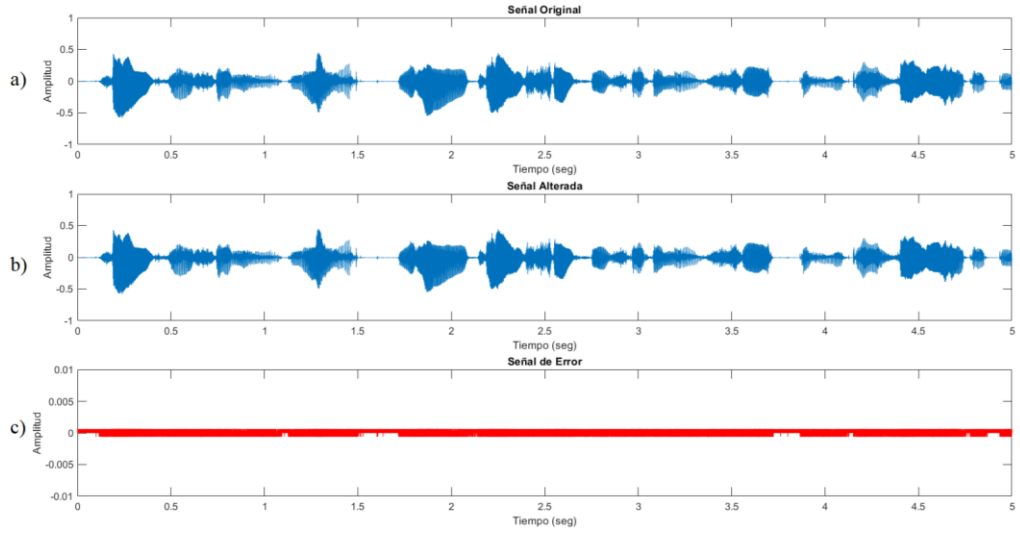

Figura 5: Señal original vs señal marcada. a) señal original, b) señal marcada, c) señal de error.

un SNR promedio de 43.22 dB, superando a los esquemas de (Chetan et al., 2021) y (Ghobadi *et al*., 2013), los resultados mostrados en la Tabla 3 indica que la marca de agua es auditivamente imperceptible, lo que también se comprueba al observar la Figura 5, donde la señal original $A[i]$ y la señal marcada $A'[i]$ son perceptiblemente idénticas, puesto que la señal de error entre ellas es muy poco perceptible.

Las pruebas realizadas para la evaluación de la integridad de la señal marcada, se llevó a cabo mediante modificaciones durante Tm segundos realizadas al inicio, en medio, al final y en diferentes partes del audio, los resultados obtenidos se muestran en las Tablas 4, 5, 6 y 7 donde se realizaron ataques intencionales, como el muteo (intervalo de muestras puestas en 0), la sustitución (las muestras de audio son sustituidas por muestras de otra señal distinta) y la clonación (cierto número de muestras del mismo audio son sustituidas en otra parte de la misma señal).

Para las tablas 4, 5, 6 y 7, A es la señal analizada, Tm es la modificación de la señal marcada medida en segundos, BER es la tasa de bits erróneos, BE es el número de bits erróneos, BT son los bits totales, % es el porcentaje de la tasa de error y BA son los bloques que se etiquetaron como alterados, donde se observa que si los bits modificados no superan el umbral establecido en la sección 3.5, el bloque se considerará como autentico. En la Tabla 4, se muestran 5 señales de audio marcadas, donde la modificación se realiza al inicio del audio.

En la Tabla 4, se observa que para el ataque de sustitución de muestras de 0.1 y 0.5 segundos la mayoría de los bloques afectados por las modificaciones se consideran auténticos, ya que las modificaciones son menores al umbral establecido.

En la Tabla 5, se observa que para el ataque de clonación y sustitución se realizaron modificaciones durante un segundo en diferentes partes de la señal de audio, los resultados obtenidos demuestran que para el ataque de sustitución el porcentaje de error es mínimo, sin embargo, aunque la modificación es mínima el sistema lo detecta. Para el ataque de clonación, se puede observar que el sistema detecta de manera adecuada los 2 bloques que fueron modificados (clonados).

En la Tabla 6, se muestran 5 señales de audio marcadas, donde la modificación se realiza a la mitad del audio. Se puede observar que el ataque de sustitución es donde se encuentra un menor % de error, esto se debe a que en algunas muestras del audio marcado coinciden con las muestras del audio sustituido, sin embrago, aunque pase esto, si el umbral que se estableció anteriormente es mayor, el sistema etiquetara ese bloque como no autentico.

Tabla 3: Comparación de la imperceptibilidad de la marca de agua del método propuesto con el estado del arte.

Archivo de audio	SNR de las señales marcadas (dB)		
	Chetan et al., 2021	Ghobadi *et al*., 2013	Método propuesto
1	22.30	39.41	42.18
2	28.22	42.96	42.09
3	20.94	36.23	43.27
4	20.94	39.36	44.44
5	16.60	39.28	44.16
Promedio	21.80	39.44	43.22

Tabla 4: Resultados obtenidos mediante ataques intencionales al inicio de la señal marcada.

A	Tm	Modificación	BER			BA
			BE	BT	$\%error$	
1	0.1	Muteo	4,394	441,000	0.99	1
		Sustitución	462	441,000	0.10	0
	0.5	Muteo	17,591	441,000	3.98	1
		Sustitución	1,166	441,000	0.26	1
	1	Muteo	44,100	441,000	10.0	2
		Sustitución	1,977	441,000	0.44	1
2	0.1	Muteo	4,379	441,000	0.99	1
		Sustitución	553	441,000	0.12	0
	0.5	Muteo	17,612	441,000	3.99	1
		Sustitución	1,105	441,000	0.25	1
	1	Muteo	44,089	441,000	9.99	2
		Sustitución	1,992	441,000	0.45	1
3	0.1	Muteo	4,421	441,000	1.0	1
		Sustitución	578	441,000	0.13	0
	0.5	Muteo	17,624	441,000	3.99	1
		Sustitución	1,154	441,000	0.26	1
	1	Muteo	44,091	441,000	9.99	2
		Sustitución	2,010	441,000	0.45	0
4	0.1	Muteo	4,370	441,000	0.99	1
		Sustitución	550	441,000	0.12	0
	0.5	Muteo	17,600	441,000	3.99	1
		Sustitución	1,053	441,000	0.23	0
	1	Muteo	44,097	441,000	9.99	2
		Sustitución	2,032	441,000	0.46	0
5	0.1	Muteo	4,381	441,000	0.99	1
		Sustitución	485	441,000	0.10	0
	0.5	Muteo	17,576	441,000	3.98	1
		Sustitución	1,080	441,000	0.24	0
	1	Muteo	44,090	441,000	9.99	2
		Sustitución	1,927	441,000	0.43	0

Tabla 5: Resultados obtenidos mediante ataques intencionales en distintas partes de la señal marcada.

A	Tm	Modificación	BER			BA
			BE	BT	$\%error$	
1	1	Clonación	44,256	441,000	10.03	2
		Sustitución	2,470	441,000	0.56	2
2	1	Clonación	44,094	441,000	9.99	2
		Sustitución	2,489	441,000	0.56	2
3	1	Clonación	44,155	441,000	10.01	2
		Sustitución	2,516	441,000	0.57	2
4	1	Clonación	44,232	441,000	10.02	2
		Sustitución	2,511	441,000	0.56	2
5	1	Clonación	44,198	441,000	10.02	2
		Sustitución	2,440	441,000	0.55	2

En la Tabla 7, las señales A fueron modificadas al final del audio. Como se puede observar la métrica BER aumenta conforme aumenta el ataque malicioso, demostrando la fragilidad de la marca de agua ante dichas modificaciones. Las Figuras 6, 7, 8 y 9 muestran en marcas de color naranja, los segmentos detectados y localizados que tienen alteraciones de muteo, sustitución y clonación de muestras respectivamente, de acuerdo con las Tablas 4 a la 7.

Tabla 6: Resultados obtenidos mediante ataques intencionales en la mitad de la señal marcada.

A	Tm	Modificación	BER			BA
			BE	BT	$\%error$	
1	0.1	Muteo	4,384	441,000	0.99	1
		Sustitución	584	441,000	0.13	0
	0.5	Muteo	22,049	441,000	4.99	1
		Sustitución	1,105	441,000	0.25	1
	1	Muteo	42,087	441,000	9.54	2
		Sustitución	2,179	441,000	0.49	2
2	0.1	Muteo	4,320	441,000	0.97	1
		Sustitución	602	441,000	0.13	0
	0.5	Muteo	21,123	441,000	4.78	1
		Sustitución	1,023	441,000	0.23	1
	1	Muteo	41,312	441,000	9.36	2
		Sustitución	2,056	441,000	0.46	2
3	0.1	Muteo	4,503	441,000	1.02	1
		Sustitución	525	441,000	0.11	0
	0.5	Muteo	22,128	441,000	5.01	1
		Sustitución	1,109	441,000	0.25	1
	1	Muteo	42,056	441,000	9.53	2
		Sustitución	2,230	441,000	0.50	2
4	0.1	Muteo	4,025	441,000	0.91	1
		Sustitución	596	441,000	0.13	0
	0.5	Muteo	22,029	441,000	4.99	1
		Sustitución	10,302	441,000	2.33	1
	1	Muteo	42,632	441,000	9.66	2
		Sustitución	1,998	441,000	0.45	2
5	0.1	Muteo	4,008	441,000	0.90	1
		Sustitución	608	441,000	0.13	0
	0.5	Muteo	21,398	441,000	4.85	1
		Sustitución	1,251	441,000	0.28	1
	1	Muteo	41,852	441,000	9.49	2
		Sustitución	2,357	441,000	0.53	2

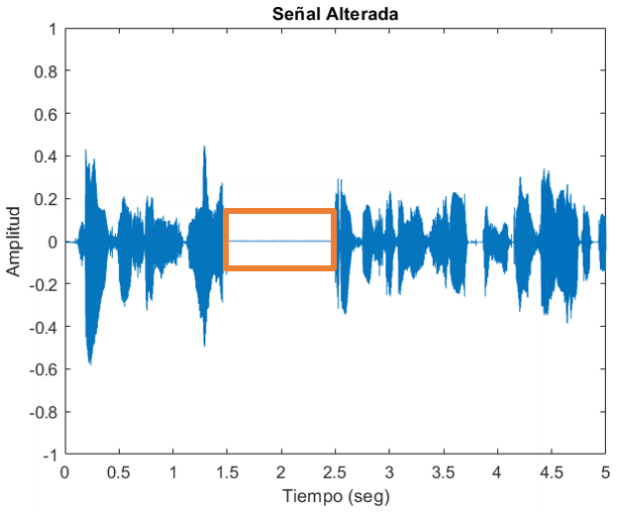

Figura 6: Detección y localización de modificaciones del ataque de *muteo* en 1 segundo de la señal de audio marcada.

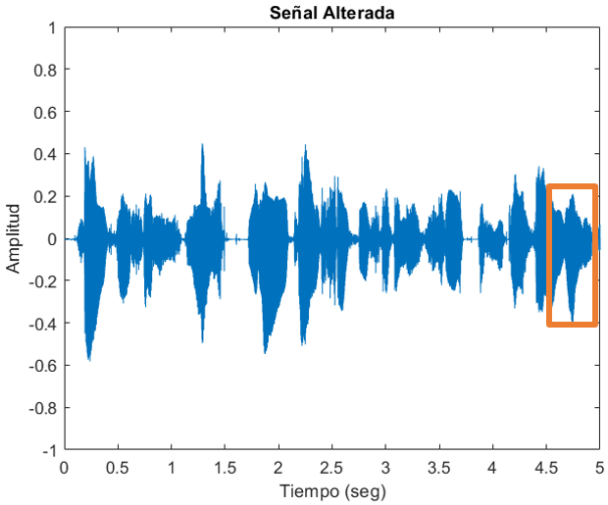

Figura 7: Detección y localización de modificaciones del ataque de *sustitución* en 0.5 segundos de la señal de audio marcada.

Tabla 7: Resultados obtenidos mediante ataques intencionales al final de la señal marcada.

A	Tm	Modificación	BER			BA
			BE	BT	$\%error$	
1	0.1	Muteo	0	441,000	0	0
		Sustitución	0	441,000	0	0
	0.5	Muteo	3	441,000	0	0
		Sustitución	9	441,000	0.002	0
	1	Muteo	22,043	441,000	4.99	1
		Sustitución	1,107	441,000	0.25	1
2	0.1	Muteo	0	441,000	0	0
		Sustitución	0	441,000	0	0
	0.5	Muteo	0	441,000	0	0
		Sustitución	4	441,000	0	0
	1	Muteo	22,047	441,000	4.99	1
		Sustitución	1,079	441,000	0.24	0
3	0.1	Muteo	0	441,000	0	0
		Sustitución	0	441,000	0	0
	0.5	Muteo	3	441,000	0	0
		Sustitución	4	441,000	0	0
	1	Muteo	22,039	441,000	4.99	1
		Sustitución	1,166	441,000	0.26	1
4	0.1	Muteo	0	441,000	0	0
		Sustitución	0	441,000	0	0
	0.5	Muteo	1	441,000	0	0
		Sustitución	4	441,000	0	0
	1	Muteo	22,042	441,000	4.99	1
		Sustitución	1,095	441,000	0.24	0
5	0.1	Muteo	0	441,000	0	0
		Sustitución	0	441,000	0	0
	0.5	Muteo	2	441,000	0	0
		Sustitución	6	441,000	0	0
	1	Muteo	22,058	441,000	5.0	1
		Sustitución	1,099	441,000	0.24	0

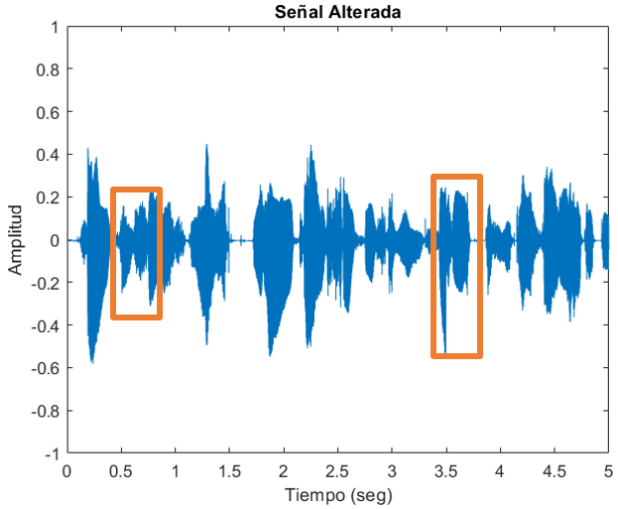

Figura 8: Detección y localización de modificaciones del ataque de *sustitución* en 1 segundo de la señal de audio marcada en distintas regiones.

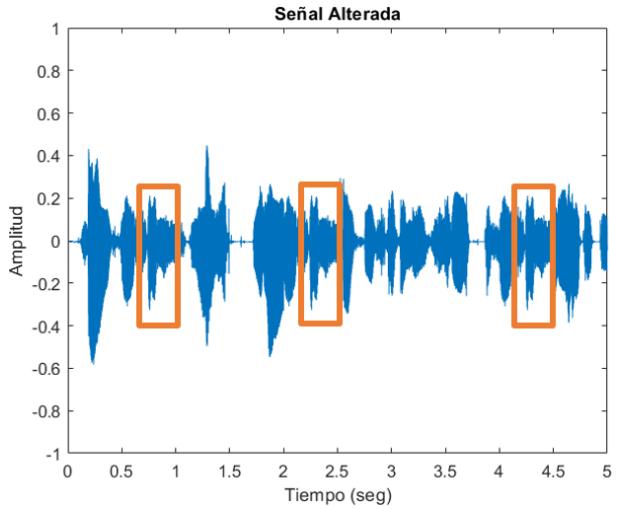

Figura 9: Detección y localización de modificaciones del ataque de *clonación* en 1 segundo de la señal de audio marcada en distintas regiones.

5. Conclusiones

En este trabajo se presenta un método de verificación de integridad y autenticidad de una señal de audio digital en donde se detectan y localizan las regiones que fueron modificadas mediante una marca de agua frágil ciega embebida en el dominio temporal, donde a través de los dos primeros MSB se hace una selección de los LSB donde se ocultará la marca de agua. El método propuesto se ha evaluado con 100 señales de audio tomadas de la base de datos (THCHS-30), demostrando una alta robustez e imperceptibilidad. Los resultados objetivos para la evaluación del método demuestran que la marca de agua es auditivamente imperceptible obteniendo un valor de SNR y NC promedio de 43.22 dB y 0.9980 respectivamente, superando los trabajos propuestos del estado del arte. El sistema permite detectar y localizar correctamente modificaciones intencionales como el *muteo*, *sustitución* y *clonación* de muestras. Adicionalmente, la técnica de marca de agua basada en el método mejorado LSB presenta una baja complejidad computacional, lo que permite implementar el método en diferentes ordenadores o dispositivos móviles.

Como trabajo a futuro, se pretende realizar un método de marca de agua que resista ataques intencionales como el recorte o adición de información, así como ataques no intencionales como la compresión.

Agradecimientos

Los autores agradecen al Instituto Politécnico Nacional (IPN-México), a la Comisión de Operación y Fomento de Actividades Académicas (COFAA-IPN) y al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT), por el apoyo brindado para la realización de este trabajo.

Referencias

Ahmed, M. A., Kiah, M. M., Zaidan, B. B., & Zaidan, A. A. (2010). A novel embedding method to increase capacity and robustness of low-bit encoding audio steganography technique using noise gate software logic algorithm. Journal of Applied Sciences, 10(1), 59-64. DOI: 10.3923/jas.2010.59.64

Alwahbani, S. M., & Elshoush, H. T. (2018). Chaos-based audio steganography and cryptography using LSB method and one-time pad. In Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016: Volume 2 (pp. 755-768). Springer International Publishing. DOI: 10.1007/978-3-319-56991-8_54

Asad, M., Gilani, J., & Khalid, A. (2011, July). An enhanced least significant bit modification technique for audio steganography. In International Conference on Computer Networks and Information Technology (pp. 143-147). IEEE. DOI: 10.1109/ICCNIT.2011.6020921

Chetan, M., Bhat, P. P., Shet, V., Husenbhai, S. B., & Bhat, A. (2021, December). Audio watermarking using modified least significant bit technique. In 2021 International conference on circuits, controls and communications (CCUBE) (pp. 1-5). IEEE. DOI: 10.1109/CCUBE53681.2021.9702715

Cox, I. J., & Miller, M. L. (2002). The first 50 years of electronic watermarking. EURASIP Journal on Advances in Signal Processing, 2002, 1-7. DOI: 10.1155/S1110865702000525

Ghobadi, A., Boroujerdizadeh, A., Yaribakht, A. H., & Karimi, R. (2013, January). Blind audio watermarking for tamper detection based on LSB. In 2013 15th International Conference on Advanced Communications Technology (ICACT) (pp. 1077-1082). IEEE.

Jeruchim, M. (1984). Techniques for estimating the bit error rate in the simulation of digital communication systems. IEEE Journal on selected areas in communications, 2(1), 153-170. DOI: 10.1109/JSAC.1984.1146031

Kabal, P. (2002). An examination and interpretation of ITU-R BS. 1387: Perceptual evaluation of audio quality. TSP Lab Technical Report, Dept. Electrical & Computer Engineering, McGill University, 1-89.

Kundur, D., & Hatzinakos, D. (1999). Digital watermarking for telltale tamper proofing and authentication. Proceedings of the IEEE, 87(7), 1167-1180. DOI: 10.1109/5.771070

Li, S., Song, Z., Lu, W., Sun, D., & Wei, J. (2017). Parameterization of LSB in self-recovery speech watermarking framework in big data mining. Security and Communication Networks, 2017. DOI:10.1155/2017/3847092

Naskar, R., & Chakraborty, R. S., (2014). Reversible digital watermarking: theory and practices. Synthesis Lectures on Information Security, Privacy, & Trust, 5(1), 1-130. DOI:10.1007/978-3-031-02342-2

Pak, C., & Huang, L. (2017). A new color image encryption using combination of the 1D chaotic map. Signal Processing, 138, 129-137. DOI: 10.1016/j.sigpro.2017.03.011

Sathishkumar, G. A., & Sriraam, D. N. (2011). Image encryption based on diffusion and multiple chaotic maps. arXiv preprint arXiv:1103.3792.

Vargas, L. M., Vera. E., & Di Gianantonio, A. (2016). Marcas de agua: una contribución a la seguridad de archivos digitales. Revista de la Facultad de Ciencias Exactas, Físicas y Naturales, 3(1), 49-54.

Wang, D., & Zhang, X. (2015). Thchs-30: A free chinese speech corpus. Available: http://arxiv.org/abs/1512.01882

Wang, J., & He, J. (2017). A speech content authentication algorithm based on a novel watermarking method. Multimedia Tools and Applications, 76, 14799-14814. DOI:10.1007/s11042-016-4027-5