SCRAMBLING DE MENSAJES DE VOZ BASADO EN TÉCNICAS DE INTELIGENCIA ARTIFICIAL

Autor: Alejandro Sandoval Camacho

UNIVERSIDAD MILITAR NUEVA GRANADA

FACULTAD DE INGENIERÍA

INGENIERÍA EN TELECOMUNICACIONES

BOGOTÁ

2016

SCRAMBLING DE MENSAJES DE VOZ BASADO EN TÉCNICAS DE INTELIGENCIA ARTIFICIAL

Autor: Alejandro Sandoval Camacho

Trabajo de grado para optar por el título de Ingeniero en Telecomunicaciones

Tutor: Ing. Diego Renza Torres, PhD

UNIVERSIDAD MILITAR NUEVA GRANADA

FACULTAD DE INGENIERÍA

INGENIERÍA EN TELECOMUNICACIONES

BOGOTÁ

2016

ÍNDICE

1. INTRODUCCIÓN 6
2. JUSTIFICACIÓN 7
3. OBJETIVOS 8
   3.1 General 8
   3.2 Específicos 8
4. ESTADO DEL ARTE 9
5. METODOLOGIA 11
6. MÉTODO PROPUESTO 13
   6.1 Etapa 1: Aleatorización 14
   6.2 Etapa 2: Recuperación del contenido secreto 16
7. VALIDACIÓN DEL ESQUEMA PROPUESTO 17
   7.1 Scrambling Degree (SD) 17
   7.2 Squared Pearson's Correlation Coefficient (SPCC) y Nivel de desorden (Ds) 21
8. ANÁLISIS DE ROBUSTEZ Y SEGURIDAD 28
9. COMPARACIÓN CON OTRAS TÉCNICAS DE SCRAMBLING DE VOZ 29
   9.1 Comparación de inteligibilidad residual 30
   9.2 Comparación de la calidad del mensaje secreto recuperado 30
   9.3 Comparación de la seguridad 31
10. CONCLUSIONES 32
11. REFERENCIAS 33

ÍNDICE DE FIGURAS

Figura 1 Fases del proyecto. 11
Figura 2 Modelo propuesto basado en Algoritmos Genéticos. 13
Figura 3 Ejemplo de la creación de cromosomas y genes. 14
Figura 4 Ejemplo de permutación asexual y creación de la clave. 15
Figura 5 MOS promedio de la señal aleatorizada de acuerdo al número de genes por cromosoma. 19
Figura 6 Gráfica de un mensaje secreto y su respectiva señal aleatorizada con el SD como función objetivo. 20
Figura 7 Número de iteraciones realizadas por el modelo de acuerdo al número de genes por cromosoma. 21
Figura 8 MOS promedio de la señal aleatorizada de acuerdo al número de genes por cromosoma. 22
Figura 9 Número de iteraciones realizadas por el modelo de acuerdo al número de genes por cromosoma. 23
Figura 10 Nivel de desorden promedio de la señal aleatorizada. 23
Figura 11 SPCC promedio de la señal aleatorizada. 24
Figura 12 Ejemplo de las señales resultantes al implementar el modelo final. Ds de la señal original (0,2280), Ds de la señal aleatorizada (0,3435) y SPCC (0,0099). 25
Figura 13 Señales resultantes en cada etapa del modelo desarrollado, implementando la ecuación (8) como función objetivo. 26

ÍNDICE DE TABLAS

Tabla 1 Mean Opinion Score 19

1. INTRODUCCIÓN

Las tecnologías de la información y la comunicación juegan un papel muy importante en los diferentes campos de la vida humana, sus avances van de la mano con la evolución del hombre y a través de los años se ha formado un vínculo muy estrecho entre las dos partes. En un mundo globalizado gracias a los múltiples medios de comunicación, es fundamental garantizar la seguridad de la información y utilizar sistemas confiables que permitan satisfacer de manera óptima las necesidades que surgen en los diversos panoramas de la cotidianidad. Uno de los medios de transmisión de información más utilizado por los seres humanos es la comunicación por voz, lo que lo hace también uno de los más vulnerables a la hora de mantener la privacidad de los mensajes que son transmitidos. Así, encontrar una herramienta que garantice la seguridad de la información sin aumentar los costos en los procesos de transmisión, ha sido uno de los principales objetivos de los estudios más recientes en este campo.

El *scrambling* (aleatorización) de mensajes de voz ha sido una de las herramientas más utilizadas para mejorar la seguridad en este tipo de señales. Gracias a la variedad de técnicas que utiliza, ofrece soluciones a las principales debilidades que presenta la transmisión de mensajes de voz a nivel de privacidad y ocultamiento de la información. Este método ha demostrado ser una de las formas más eficientes cuando se desea ocultar información o simplemente suplir la necesidad básica de garantizar la privacidad en un canal de comunicación. Pero, ¿logra el *scrambling* por si solo cumplir todos los objetivos que se plantean cuando se habla de seguridad de la información?

En este proyecto se presenta un modelo de aleatorización de mensajes de voz inspirado en algoritmos genéticos, a partir de un mecanismo de evolución basado en permutación asexual. El primer paso consistió en determinar la función objetivo adecuada para el algoritmo genético, validando posteriormente su funcionamiento mediante el software de modelamiento MATLAB; en este caso se realizaron pruebas de *scrambling* sobre diferentes mensajes de voz, evaluando el nivel de desorden de la señal aleatorizada.

2. JUSTIFICACIÓN

El ser humano ha utilizado diferentes técnicas de comunicación encubierta a través de la historia, haciendo uso de ciencias como la criptografía y la esteganografía. La codificación de mensajes de voz ha sido un reto muy grande para la ingeniería y es un tema esencial para la seguridad de la información en la sociedad moderna. Incluso desde inicios del siglo pasado, cuando los principales sistemas de comunicación comenzaron a surgir, se pensaba en la comunicación por voz de manera encubierta; así se podría manipular una señal de tal forma que fuera codificada e interpretada sólo por la persona autorizada para ello.

Gracias a las amplias aplicaciones de la comunicación por voz en los campos más importantes de la sociedad humana, como lo son la economía, la investigación científica, los asuntos sociales y de defensa nacional, es común escuchar en los diferentes medios, temas como delitos informáticos, pérdida de información, vigilancia y escuchas ilegales, entre otros. Por esta razón la seguridad de la información es uno de los factores más importantes a tener en cuenta cuando se desarrollan nuevos sistemas de comunicación. Así mismo, la codificación de voz ha tenido una considerable aceptación como un medio eficaz para mejorar la protección en aplicaciones sociales y políticas.

Hoy en día se utiliza la codificación en todo tipo de medios, por ejemplo, en la protección de derechos de autor, seguridad en la transferencia de información entre satélites y estaciones terrestres y en las comunicaciones militares. El *Scrambling* o aleatorización de mensajes de voz es uno de los métodos más utilizados actualmente y que gracias a sus diferentes técnicas de implementación brinda gran seguridad en la transmisión de la información garantizando la baja inteligibilidad residual y alta robustez frente a los ataques hostiles.

A pesar de los avances logrados hasta el momento, algunas investigaciones arrojan indicios sobre la posibilidad de atacar las diferentes debilidades de métodos de *scrambling* clásicos, incorporando nuevas áreas del conocimiento. En este sentido, debido a la importancia del *scrambling* en el sector de las telecomunicaciones y al reciente interés en el área de la inteligencia artificial, el presente trabajo está orientado al planteamiento de un nuevo esquema de *scrambling* basado en algoritmos genéticos.

3. OBJETIVOS

3.1 General

Aplicar un método de aleatorización de mensajes de voz inspirado en algoritmos genéticos, con el fin de garantizar la privacidad en la transmisión de la información.

3.2 Específicos

- Desarrollar un programa en MATLAB que permita aleatorizar mensajes de voz utilizando algoritmos genéticos.

- Validar la eficiencia del algoritmo con base en su robustez, seguridad y en la inteligibilidad residual del mensaje codificado.

- Comparar el esquema desarrollado con otras técnicas de *scrambling* de voz.

4. ESTADO DEL ARTE

El *scrambling* ha sido objeto de estudio de un gran número de investigaciones en el área de las comunicaciones. Hace algunos años, la codificación del audio era la única manera de ocultar la información que se quería transmitir y se basó principalmente en la aleatorización de la señal en el dominio del tiempo, en el dominio de la frecuencia o ambos. Otra alternativa consistió en utilizar una secuencia o matriz específica, tal como una secuencia pseudo aleatoria, secuencias de Fibonacci, o una matriz Hadamard, entre otros. La desventaja principal que tienen en común estos métodos es que la clave de descifrado es fija y debido al acelerado avance de la tecnología, estas técnicas se han vuelto inseguras.

En el 2012 Zeng et al., introducen una técnica de aleatorización basada en el muestreo comprimido; en este se integra la compresión de la señal con el cifrado, con el fin de disminuir la carga al proceso de transmisión. Con esto se obtiene un alto grado de seguridad debido a la gran longitud disponible para el espacio de la clave. Este método se diferencia de los existentes por aleatorizar las muestras del audio comprimidas en lugar de la señal original [1].

Madain et al. aplicaron autómatas celulares en el campo de la aleatorización de audio debido al potencial que posee para romper la correlación entre las muestras de audio de manera efectiva. Dichos autores probaron diferentes tipos de autómatas celulares con el fin de evaluar cuál conduce a un grado más alto de codificación. Después de realizar pruebas con archivos de música y voz de diferentes tamaños, se encontró que el tipo de autómata celular con el que se logra un mayor grado de aleatorización es el de 2 dimensiones con comportamiento complejo [2].

Algunos estudios recientes proponen el uso de inteligencia artificial como una herramienta fundamental para la optimización de las funciones de los procedimientos utilizados por los algoritmos tradicionales. En este campo se tienen alternativas como las redes de neuronas artificiales, la computación evolutiva o los algoritmos genéticos (AG). Aunque este último concepto nació a mediados del siglo pasado, recientemente ha sido foco de atención de los investigadores en diversos campos de la ingeniería [3]. Según Saif Hasan et al., los algoritmos genéticos se pueden definir de manera general como modelos computacionales que utilizan técnicas inspiradas en la evolución natural, tales como la herencia, mutación, selección o cruce [4]. Uno de los estudios más importantes sobre este tema fue realizado por John Holland en 1975 quien publicó el libro “La adaptación en Sistema natural y artificial” [5]. A partir de investigaciones como esta, se han utilizado los principios biológicos de la teoría de la evolución para el beneficio de los seres humanos.

De manera general, un algoritmo genético es un método de inteligencia artificial que es iterativo y cuya solución no obedece a una fórmula determinística. Un algoritmo

genético se caracteriza por: una población inicial, una función objetivo (función *fitness*) y un mecanismo de evolución [6].

La población inicial es un conjunto de padres del posible espacio de soluciones, que contienen genes o elementos base. Los padres se mezclan para obtener nuevas generaciones utilizando mecanismos de evolución [7]. Entre los mecanismos de evolución se tiene la alternativa de reproducción sexual o asexual, siendo su diferencia básica la participación de dos padres (sexual) o sólo uno de ellos (asexual). Los mecanismos de reproducción más utilizados corresponden a la mutación y la combinación. En la mutación, algunos genes del padre o de la madre se modifican, mientras que en la combinación el hijo se crea a partir de “trozos” de genes del padre y de la madre. Este proceso iterativo de creación de hijos se detiene cuando se satisface una de dos condiciones: el número de iteraciones llega a un máximo pre-definido por el sistema, o el hijo obtenido cumple con la función objetivo definida previamente [8].

Aunque la reproducción sexual ha sido la más estudiada en las diferentes aplicaciones de AG, en la última década la reproducción asexual ha aparecido como una solución alternativa [9]. La reproducción asexual se caracteriza por las siguientes condiciones:

- La población inicial consiste en un solo padre.
- El hijo tiene solamente material genético de un único padre, obtenido por alguno de los mecanismos de reproducción asexual.
- El mecanismo de combinación de reproducción sexual se transforma en el mecanismo de permutación en la reproducción asexual.

Las características de reproducción citadas anteriormente, unidas al concepto de aleatorización, definen las características del presente trabajo y serán expuestas en la siguiente sección.

5. METODOLOGIA

El proyecto fue desarrollado en 5 fases, la cuales se aprecian en la Figura 1.

Figura 1 Fases del proyecto.

Fase 1. Revisión del estado del arte: Durante esta etapa, se llevó a cabo un estudio de las investigaciones realizadas anteriormente en el campo de la aleatorización de mensajes de voz y sus diferentes métodos de implementación, con el fin de comprender y determinar los modelos más destacados en dicha área.

Fase 2. Planteamiento de hipótesis y definición de objetivos: Una vez realizado el estudio de los diferentes métodos utilizados en investigaciones anteriores, se decidió trabajar con los algoritmos genéticos, una rama de la inteligencia artificial y se planteó como hipótesis la implementación de un modelo basado en la reproducción asexual como algoritmo para la aleatorización de los mensajes de voz.

Fase 3. Desarrollo y validación experimental de la hipótesis: Después de obtener un programa funcional en MATLAB, se realizó la respectiva validación por medio de un número significativo de simulaciones, en las que se utilizaron mensajes de voz con diversas características como la duración, el idioma y género del audio.

Fase 4. Análisis de los resultados y conclusiones: Una vez obtenidos los resultados de las pruebas del modelo definitivo, se realizó un análisis estadístico y se obtuvieron las conclusiones del proyecto. Así mismo se compararon las características del método propuesto, con algunas alternativas existentes en la literatura.

Fase 5. Documentación: En esta fase se realizó un artículo que fue sometido en la revista Dyna de la universidad Nacional de Colombia, al igual que el presente documento como productos del proyecto.

6. MÉTODO PROPUESTO

En este apartado se describe la composición y el funcionamiento del algoritmo que se obtuvo como resultado de la investigación y el trabajo propuesto.

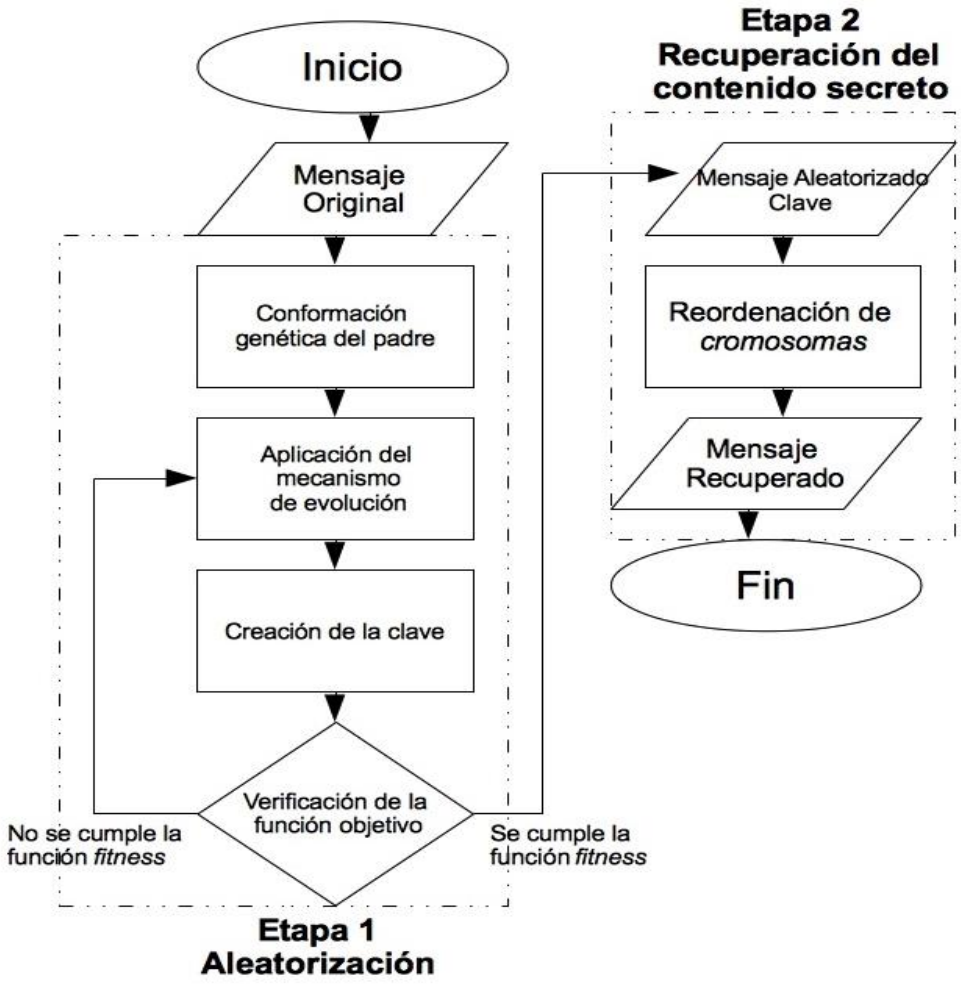


Figura 2 Modelo propuesto basado en Algoritmos Genéticos.

El objetivo principal del algoritmo desarrollado es destruir la inteligibilidad de un mensaje de voz secreto, con el fin de que, en caso de ser interceptado, su contenido no sea revelado. Si el destinatario conoce la clave correcta, puede reversar el proceso de aleatorización y recuperar el mensaje secreto original. En esta propuesta la aleatorización se realiza utilizando un proceso iterativo basado en reproducción asexual. El esquema contiene dos etapas, en el transmisor se realiza la aleatorización y en el receptor la recuperación del contenido secreto (Figura 2). A continuación se detallan cada una de ellas.

6.1 Etapa 1: Aleatorización

El proceso de aleatorización se basa en un algoritmo genético con reproducción asexual, el cual se encarga de desordenar la señal de voz para que la inteligibilidad de su contenido se destruya. Dentro del algoritmo genético se define: conformación genética del padre, mecanismo de evolución y función objetivo. Los pasos de esta fase son:

A. **Conformación genética del padre:** En la primera generación, el padre es la señal de voz original discreta constituida por M muestras. Para la conformación genética del padre, la señal se divide en N grupos, que en este caso serán los cromosomas. Cada cromosoma a su vez estará conformado por L=M/N muestras, que corresponden a los genes. Por ejemplo, una señal de voz de 80K muestras con 20 genes por cromosoma tendrá un N igual a 4K. En la Figura 3 se presenta un ejemplo de conformación del padre.

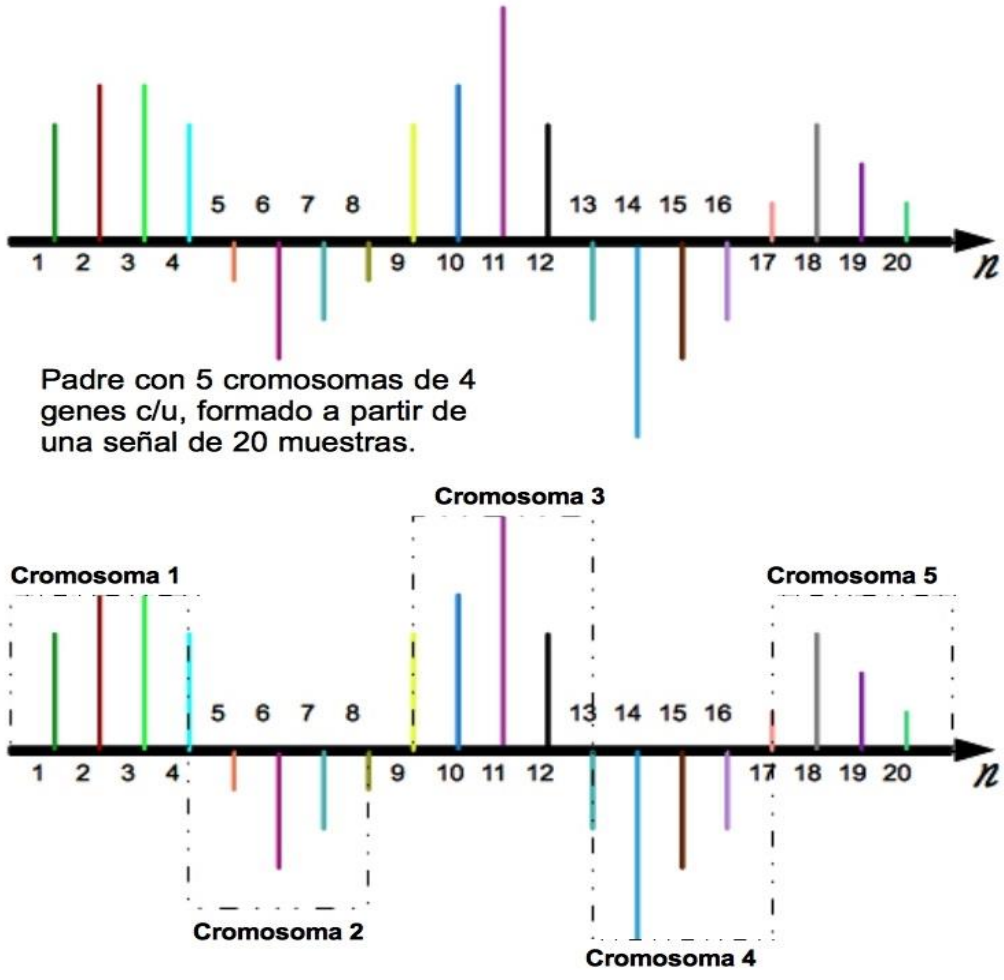


Figura 3 Ejemplo de la creación de cromosomas y genes.

B. **Aplicación del mecanismo de evolución:** El mecanismo de evolución utilizado es el de permutación asexual, en el cual dos cromosomas se intercambian de posición para formar el nuevo hijo. La selección de los cromosomas es de forma aleatoria. Con el propósito de disminuir al máximo el número de iteraciones necesarias para destruir la inteligibilidad del mensaje secreto, se incluye una restricción dentro del algoritmo de tal forma que si un cromosoma ya ha cambiado de posición, no se vuelve a permutar. Es decir, solamente se permite un cambio de posición por cromosoma.

C. **Creación de la clave:** La clave es una matriz de dos columnas en la cual se guardan los números de los cromosomas seleccionados para la permutación. Por ejemplo, suponga que se tienen 1000 cromosomas para permutar, cada uno con 80 genes. Al aplicar la permutación asexual se han seleccionado los cromosomas 102 y 512, entonces ellos se intercambian entre si y la primera fila de la clave contendrá las posiciones 102 y 512. El total de filas de la clave es igual al total de iteraciones (o permutaciones) realizadas. Se presenta un ejemplo en la Figura 4.

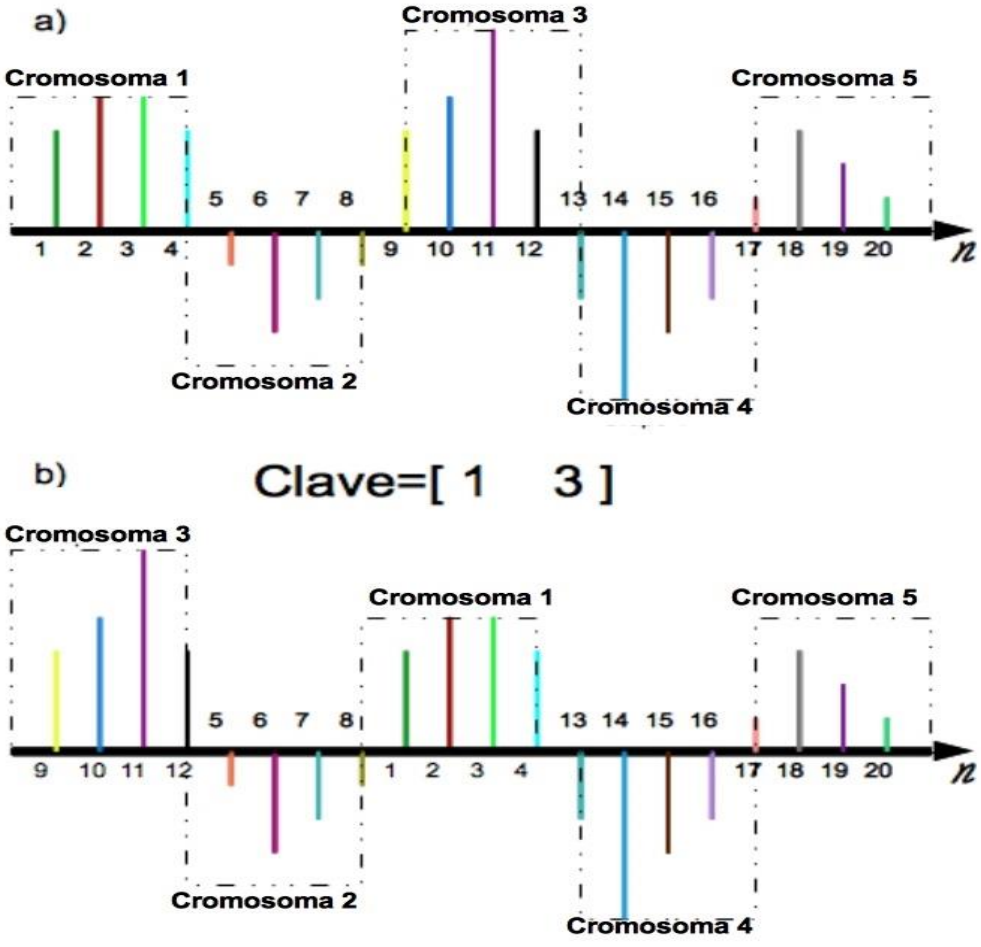


Figura 4 Ejemplo de permutación asexual y creación de la clave.

D. **Verificación de la función objetivo:** La función objetivo está basada en la similitud entre el mensaje de voz original y el nuevo hijo. Se realizaron diferentes pruebas para seleccionar los parámetros matemáticos que midieran con mayor exactitud dicha similitud, que debería ser lo más baja posible. En la etapa de aleatorización el mecanismo de evolución se ejecuta hasta que el nuevo hijo cumpla con los valores establecidos para la función objetivo o hasta alcanzar el número máximo de iteraciones permitido por el modelo.

6.2 Etapa 2: Recuperación del contenido secreto

A partir de la señal de voz aleatorizada y la clave, se recupera el mensaje secreto original. Debido a que la clave contiene las posiciones de los cromosomas permutados, el proceso consiste en devolver cada cromosoma a su posición original. Así mismo es importante considerar aquí que el proceso es totalmente reversible. Las etapas para recuperar la inteligibilidad del mensaje se detallan a continuación:

A. **Mensaje aleatorizado y clave:** Una vez el mensaje aleatorizado llega a su destino, comienza la etapa de recuperación. Esta vez el objetivo es lograr la máxima similitud entre el padre de la primera generación y el hijo de la segunda generación. El hijo obtenido en el proceso de aleatorización contiene todos los cromosomas de su padre y la clave contiene la ubicación en la que se encontraba cada uno de ellos, por lo cual es posible reubicarlos con el fin de recuperar la inteligibilidad original.

B. **Reordenación de cromosomas:** El total de permutaciones para reversar el proceso de aleatorización es igual al total de filas en la clave. Cada fila contiene 2 posiciones con los respectivos cromosomas que han sido permutados. Por ejemplo si la fila de la clave contiene los valores 345 y 24, quiere decir que el cromosoma que se encuentra actualmente en la posición 24 corresponde a la posición 345 y viceversa.

C. **Obtención del mensaje recuperado:** Después de retornar cada cromosoma a su posición original, la composición genética del hijo de esta segunda generación es idéntica a la del padre de la primera generación, por lo tanto son iguales. Al reproducir la señal resultante se percibe que la inteligibilidad se ha recuperado en su totalidad. La similitud entre los mensajes recuperados y los secretos fueron medidos con el parámetro matemático SPCC, donde un SPCC de 1 indica que las dos señales son idénticas.

7. VALIDACIÓN DEL ESQUEMA PROPUESTO

En esta sección se realiza un análisis de la eficiencia del esquema propuesto de acuerdo a la función objetivo.

7.1 Scrambling Degree (SD)

El *scrambling degree* es un parámetro matemático que mide el grado de aleatorización en el contenido del mensaje secreto, a partir de la relación entre las diferencias de los vectores de muestras de la señal original y aleatorizada. La forma para calcular el valor del SD se muestra a continuación:

En primer lugar, se calcula la diferencia de la señal D, por medio de la ecuación (1).

$$D(i) = \frac{1}{4}\sum_{i=3}^{m-2}\{4 * S(i) \ -(S(i-1) \ + S(i-2) \ + S(i+1) + \ S(i+2)\} \qquad \textbf{(1)}$$

Donde S(i) corresponde a cada una de las muestras de la señal de audio, m es el número total de muestras y D es un vector de m-4 valores.

El segundo paso es calcular A y B, que corresponden a la suma y la resta de la diferencia de la señal original y la señal aleatorizada. Estos valores se obtienen por medio de las ecuaciones (2) y (3).

$$B = D_2 - \ D_1 \qquad \textbf{(2)}$$

$$A = \ D_2 + \ D_1 \qquad \textbf{(3)}$$

Donde D_2 es la diferencia de la señal aleatorizada y D_1 es la diferencia de la señal original, calculadas mediante la ecuación (1).

Por último el SD se obtiene por medio de la ecuación (4).

$$SD = \ {}^{B}/_{A} \qquad \textbf{(4)}$$

Los valores para el SD se encuentran entre 0 y 1, donde 1 significa que la inteligibilidad del contenido secreto ha sido alterado por completo en el proceso de aleatorización, esto quiere decir que la inteligibilidad residual es 0.

Las primeras pruebas del modelo propuesto se realizaron estableciendo el SD como único parámetro en la función objetivo. Adicionalmente, se incluyó un número máximo de iteraciones como medida de control del proceso iterativo. El parámetro SD se mide entre la señal de voz original y la señal de voz aleatorizada, es decir entre el padre de la primera generación y el hijo de la generación actual. El SD indica el grado de diferencia entre dos señales, teniendo en cuenta la relación existente entre los vecinos de cada una de las muestras, tanto en los genes del padre como del hijo. En este caso, si la inteligibilidad del mensaje secreto ha sido destruida, los vecinos de los genes de la señal aleatorizada deben ser diferentes a los vecinos de los genes del padre. Para este caso se establecieron los dos siguientes criterios para la función objetivo:

$$SD > 0.3$$

$$OR \qquad \textbf{(5)}$$

$$No.\ Iteraciones = 10.000$$

Es decir que el algoritmo realizará las permutaciones hasta que el valor del SD llegue a ser mayor que 0.3 o hasta que el número de iteraciones sea igual a 10.000. El modelo incluye una restricción con el fin de disminuir el número de iteraciones, que consiste en etiquetar los cromosomas para que solo puedan ser permutados una sola vez. Se observó que a medida que se agota el número de cromosomas disponibles para permutar, el incremento en el SD disminuye considerablemente hasta el punto en que no es posible alcanzar el valor determinado en la función objetivo. Por esta razón, se ve afectado el tiempo de ejecución del algoritmo ya que la búsqueda de cromosomas sin permutar se hace más larga, sin mencionar los gastos computacionales incluidos en el proceso de comprobación de cromosomas etiquetados por la restricción.

Para evaluar el funcionamiento de la función objetivo, se aplicó el algoritmo a 10 señales de voz de 10 segundos de duración. De igual forma se trabajó con diferentes valores de genes por cromosoma, así: 5, 10, 20 y 25 genes. En total se realizaron 200 pruebas puesto que se analizaron los resultados cuando el SD alcanza valores de 0.05, 0.1, 0.2, 0.25 y 0.3.

En este caso, para la validación se utilizó la función objetivo de la ecuación (5) y se verificó de forma perceptual si la inteligibilidad del mensaje secreto es destruida. Para esto, se utilizó el Mean Opinion Score (MOS), con la siguiente asignación de puntaje:

Tabla 1 Mean Opinion Score

4	La señal aleatorizada no tiene rastros del mensaje original (suena como ruido gaussiano).
3	La señal aleatorizada contiene algunos rastros (vocales, silabas) del mensaje original.
2	Se logra entender el mensaje por partes o por zonas.
1	Se logra escuchar el mensaje original.

El balance de estos resultados se presenta en la Figura 5, en la cual se muestra el valor MOS promedio obtenido para cada uno de los tamaños de cromosoma evaluados.

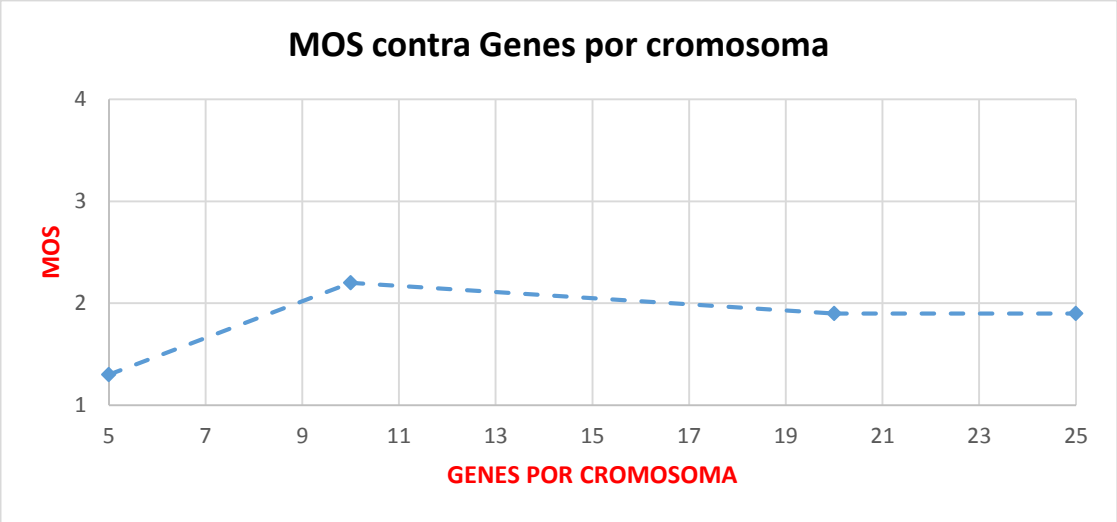


Figura 5 MOS promedio de la señal aleatorizada de acuerdo al número de genes por cromosoma.

En la Figura 6 se presenta una comparación, entre las gráficas del mensaje secreto y el mensaje aleatorizado obtenido con la función objetivo (5). A pesar de que visualmente el contenido del mensaje secreto parece haber sido destruido, perceptualmente la inteligibilidad residual del mensaje aleatorizado es alta, es decir en la señal resultante se perciben rastros del mensaje original.

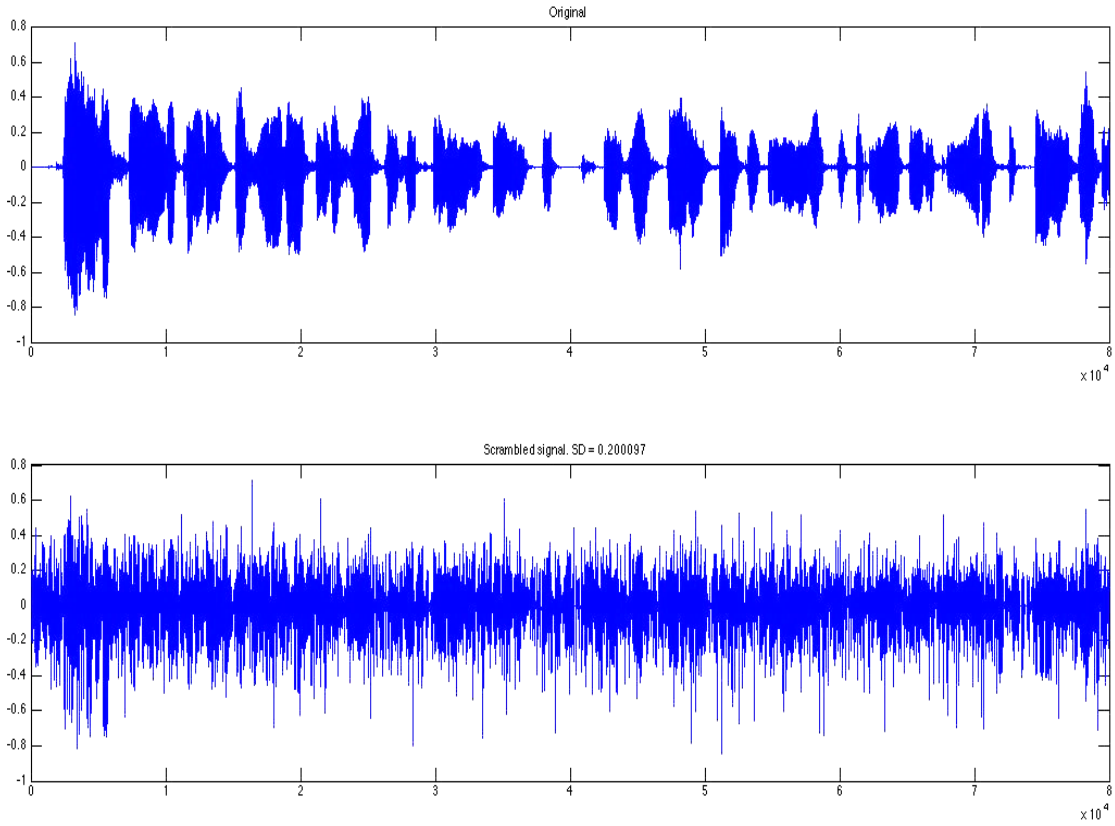


Figura 6 Gráfica de un mensaje secreto y su respectiva señal aleatorizada con el SD como función objetivo.

Teniendo en cuenta los resultados anteriores, se concluyó que el parámetro SD no cumple con los objetivos del proyecto ya que las señales aleatorizadas presentan una alta inteligibilidad residual, es decir poseen rastros del mensaje original. De la misma manera no se evidencia una mejora significativa en la destrucción de la inteligibilidad de la señal aleatorizada a medida que el número de genes por cromosoma aumenta, ya que el MOS mantiene niveles similares para todos los tamaños de cromosoma, tal como se observa en la Figura 5.

En cuanto al costo computacional, en la Figura 7 se observa un decremento en el número de iteraciones realizadas a medida que se aumenta el número de genes por cromosoma. Esto es de esperarse, ya que la cantidad de permutaciones depende del tamaño del cromosoma. Sin embargo, durante las simulaciones se observó que el modelo emplea demasiado tiempo en aproximarse a los valores seleccionados para la función objetivo, en este caso un SD de 0,3.

Los resultados obtenidos en este caso no permitieron establecer rangos funcionales para los valores de los parámetros involucrados en el proceso de aleatorización. Por esta razón, se consideró otra alternativa para la función objetivo.

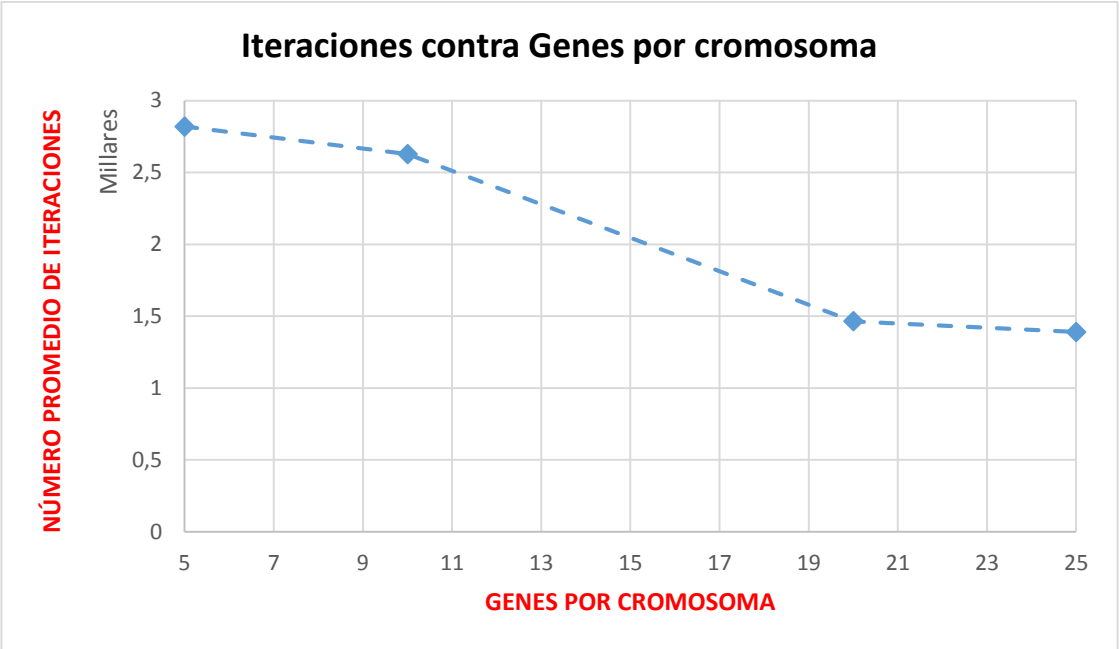


Figura 7 Número de iteraciones realizadas por el modelo de acuerdo al número de genes por cromosoma.

7.2 Squared Pearson's Correlation Coefficient (SPCC) y Nivel de desorden (Ds)

El coeficiente de correlación cuadrático de Pearson mide la relación entre dos señales U y V, la cual está definida por la ecuación (6).

$$SPCC = \frac{cov\,(U,V)}{\sigma U\ \sigma V} \tag{6}$$

Donde σ es la desviación y *cov* es la covarianza. Si el SPCC es cercano a cero indica que las dos señales son poco similares o no correlacionadas y por lo tanto, que el contenido del mensaje secreto se ha destruido.

Por otro lado el promedio de nivel de desorden de las muestras de una señal se calcula de acuerdo a la ecuación (7).

$$\overline{Ds} = \frac{\sum_{i=1}^{m-2} \sqrt{|S(i+1) - S(i)| + |S(i+1) - S(i+2)|}}{m-2} \tag{7}$$

Donde S es la señal de voz aleatorizada y m es el número total de muestras de la misma. A mayor destrucción de la inteligibilidad del mensaje secreto, mayor será el valor del Ds.

Teniendo en cuenta los resultados obtenidos con el SD**,** se decidió utilizar dos parámetros para evaluar la similitud entre las dos señales: el SPCC y el nivel de desorden de la señal aleatorizada. También se optimizó la medida de control del proceso iterativo, haciéndola dependiente del número de cromosomas de cada generación de hijos. El SPCC se mide al igual que el *Scrambling Degree*, entre el padre de la primera generación y el hijo de la generación actual, mientras que el Ds se mide directamente en el hijo. Se estableció la función objetivo así:

$$SPCC < 0.01 \text{ AND } Ds > 0.38$$

$$OR \qquad \textbf{(8)}$$

$$No.\ Iteraciones = No.\ Cromosomas / 2$$

Para comprobar la funcionalidad de estos parámetros como función objetivo, se realizaron simulaciones del modelo con señales de voz de 10 segundos de duración, utilizando 20 mensajes secretos diferentes, la mitad de ellos en inglés. Se trabajó con diferentes valores de genes por cromosoma, así: 2, 5, 10, 16, 20, 25, 40, 50, 80 y 100 genes.

Al igual que en el caso anterior, la primera validación para la nueva función objetivo que se realizó fue con el MOS (Tabla 1). Los resultados de las simulaciones se muestran en la Figura 8.

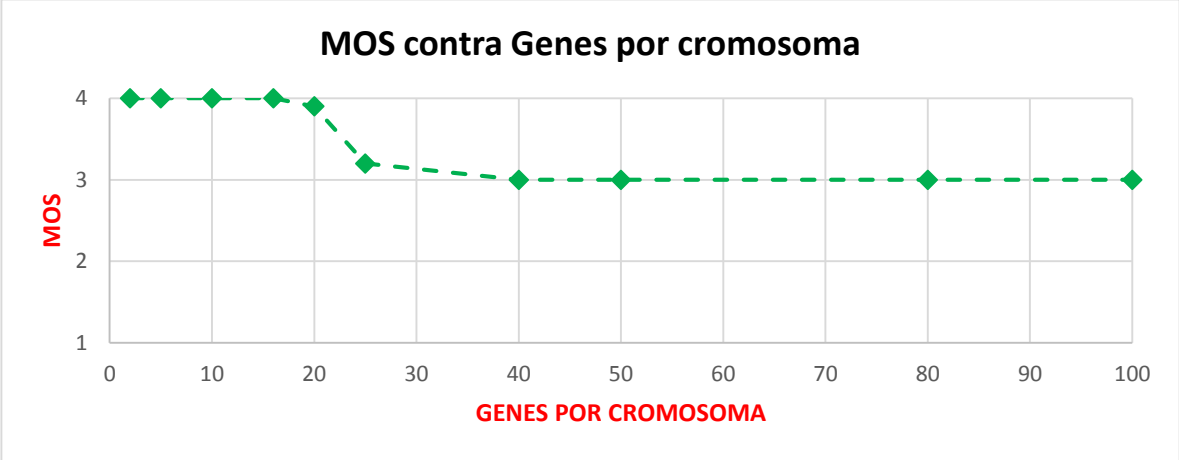


Figura 8 MOS promedio de la señal aleatorizada de acuerdo al número de genes por cromosoma.

De acuerdo a los resultados, a medida que disminuye el número de cromosomas y aumenta el total de genes por cromosoma, la calidad de la aleatorización disminuye, y viceversa. Es decir que para lograr cumplir con el objetivo de destruir la inteligibilidad del contenido secreto, el número de genes por cromosoma debe ser bajo.

En relación al total de iteraciones necesarias para cumplir con la función objetivo, se tienen los resultados de la Figura 9. En estos términos, se deduce que al aumentar la cantidad de genes por cromosoma disminuye el número de iteraciones necesarias para satisfacer el objetivo de la aleatorización. Sin embargo, si el total de genes es muy grande, aun cuando se cumpla con la función objetivo, no se destruye la inteligibilidad del contenido secreto.

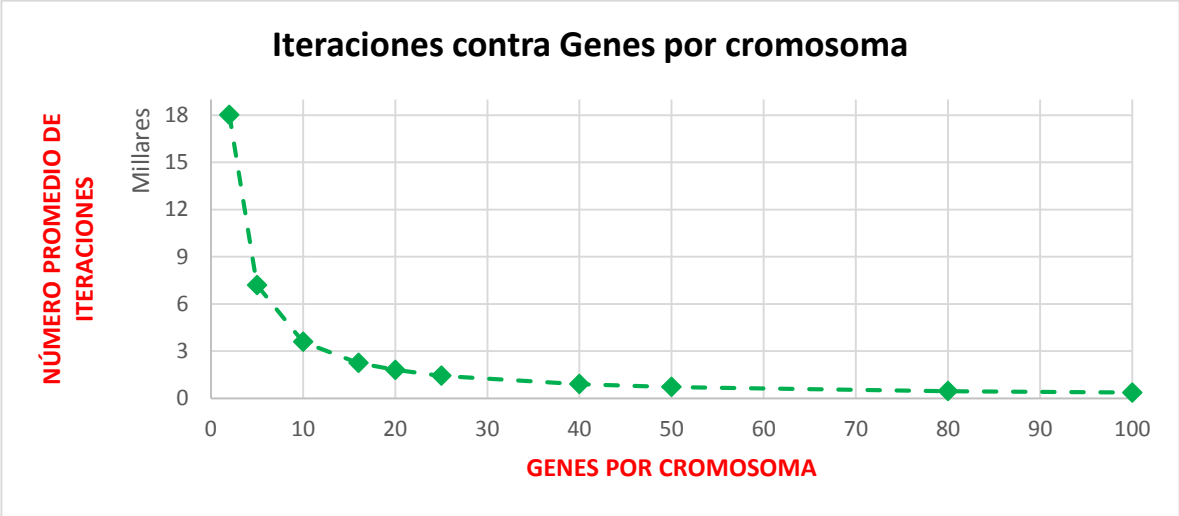


Figura 9 Número de iteraciones realizadas por el modelo de acuerdo al número de genes por cromosoma.

En cuanto al Ds, la Figura 10 evidencia que el nivel establecido en la función objetivo, se obtiene con la totalidad de los valores seleccionados como número de genes por cromosoma.

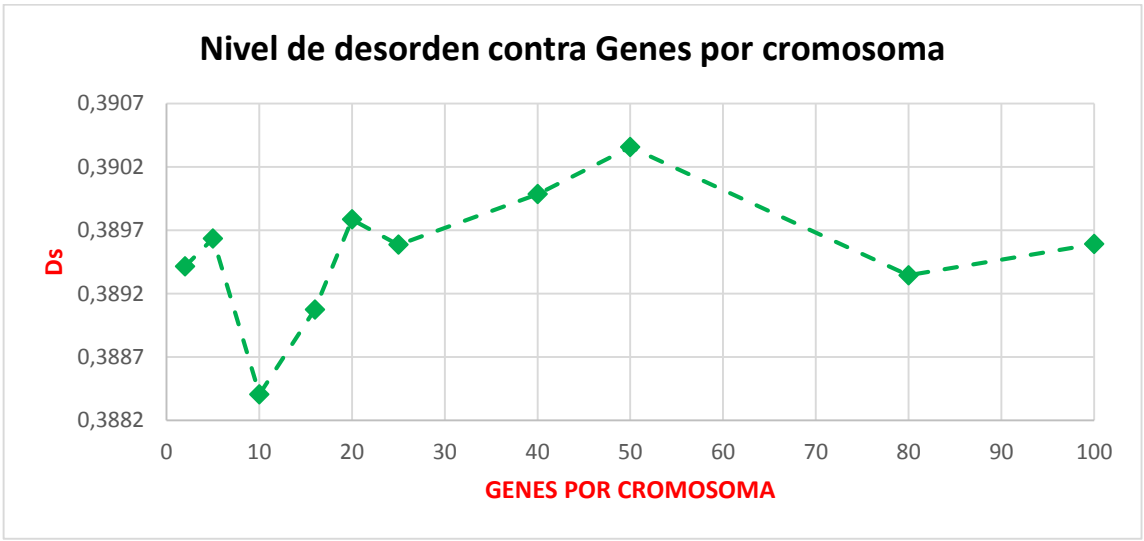


Figura 10 Nivel de desorden promedio de la señal aleatorizada.

Por su parte, el valor umbral asignado en la ecuación (8) para el parámetro SPCC, también se satisface para los diferentes valores de genes por cromosoma evaluados. En la Figura 11 se puede apreciar cómo los índices del parámetro tienden incluso a descender por debajo de 0,01 a medida que el número de genes aumenta.

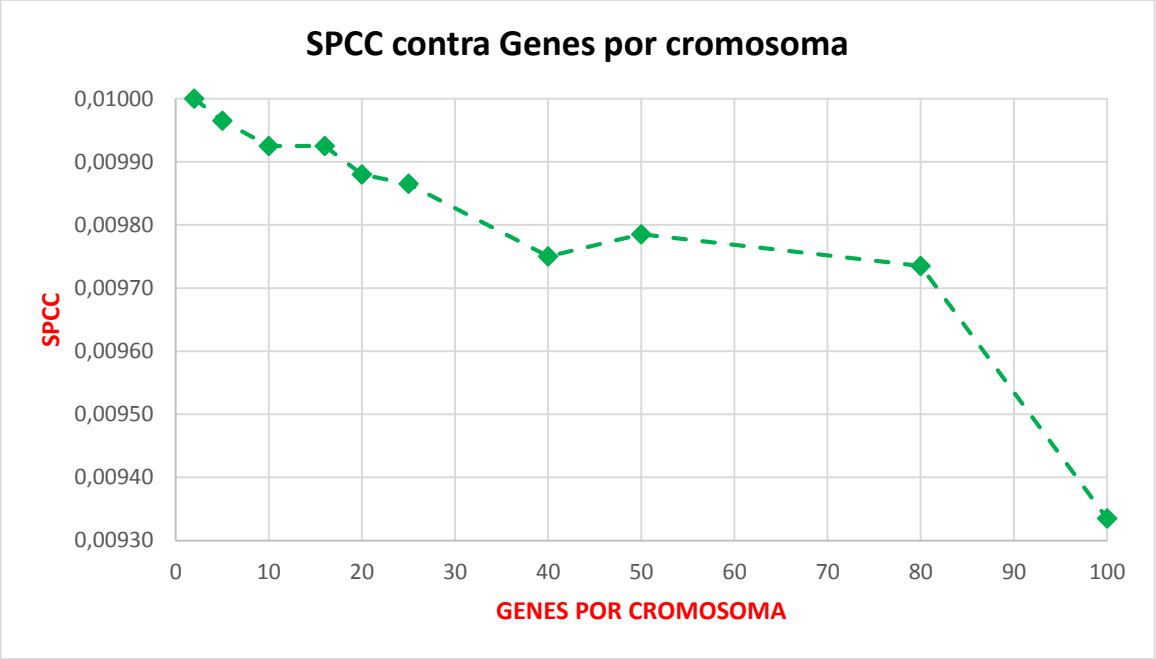


Figura 11 SPCC promedio de la señal aleatorizada.

Después de implementar el modelo con la función objetivo (8) se obtiene una señal de audio con muy baja inteligibilidad residual, que puede ser transmitida por un canal de comunicación inseguro sin revelar el contenido del mensaje secreto original. Adicionalmente, con el fin de no despertar sospechas sobre la existencia de contenido secreto en la transmisión, se recomienda implementar el modelo desarrollado con un número de 16 genes por cromosoma. Esto, debido a que se evidencia mayor similitud de la señal aleatorizada con una señal de ruido, cuando se trabaja con valores entre 10 y 20 genes por cromosoma.

En la Figura 12 se muestra a manera de ejemplo los gráficos de la señal original, aleatorizada y recuperada obtenidas al implementar la función objetivo (8).

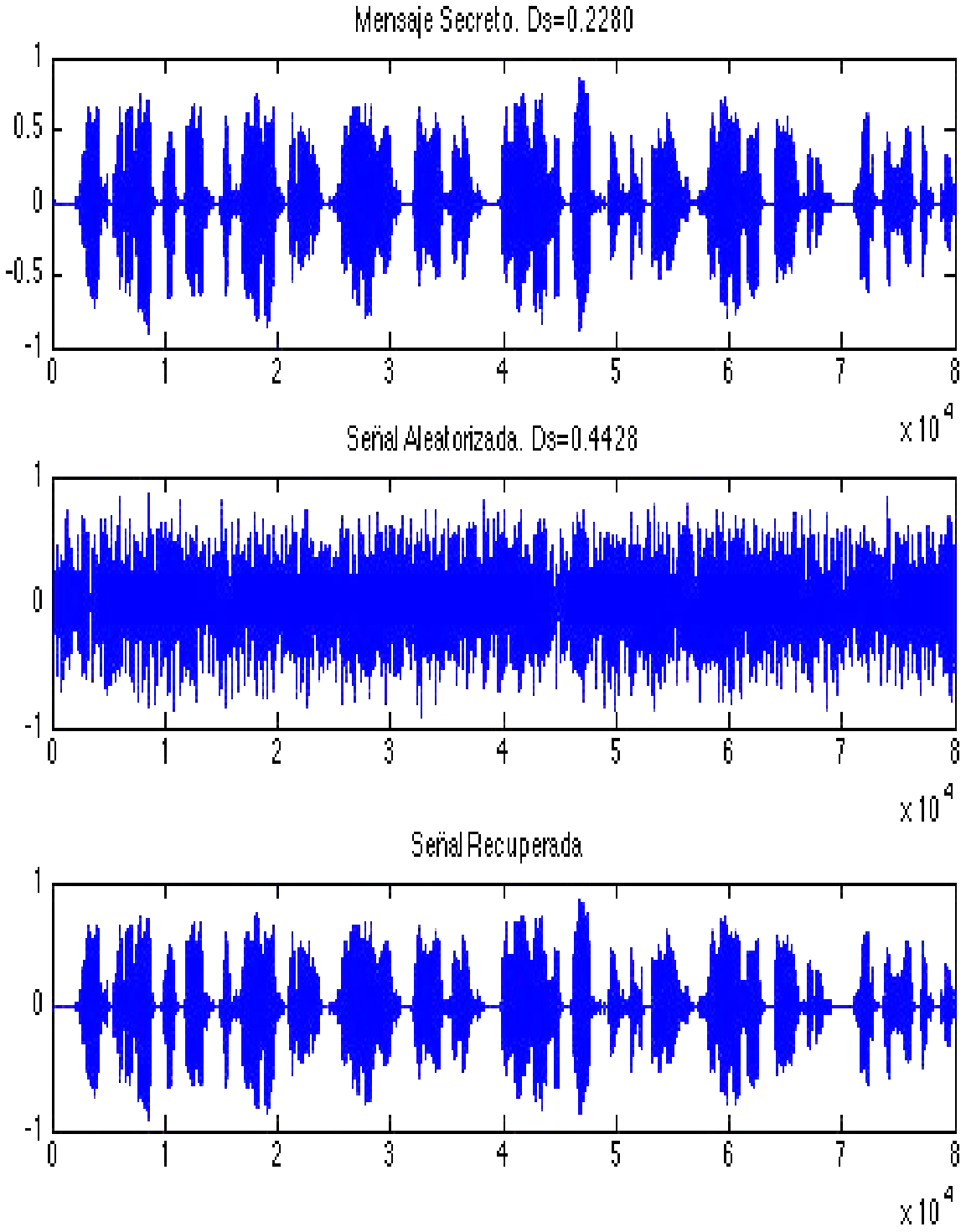
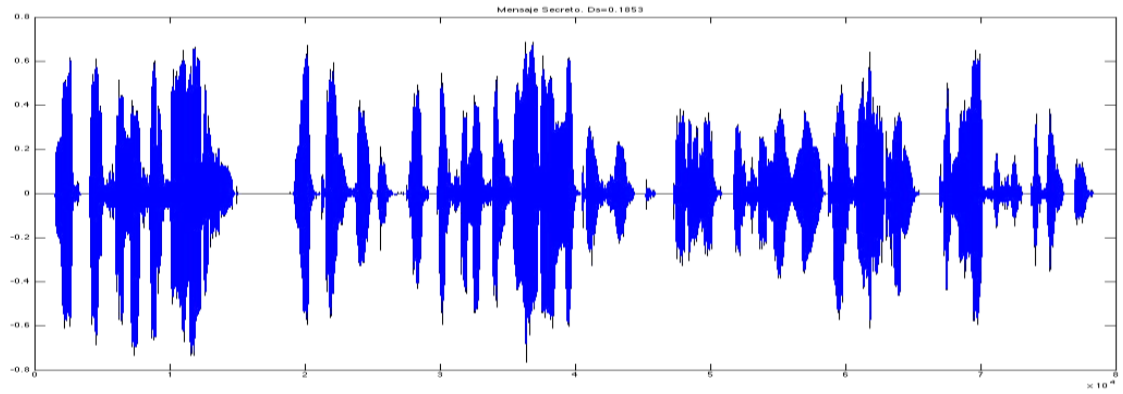
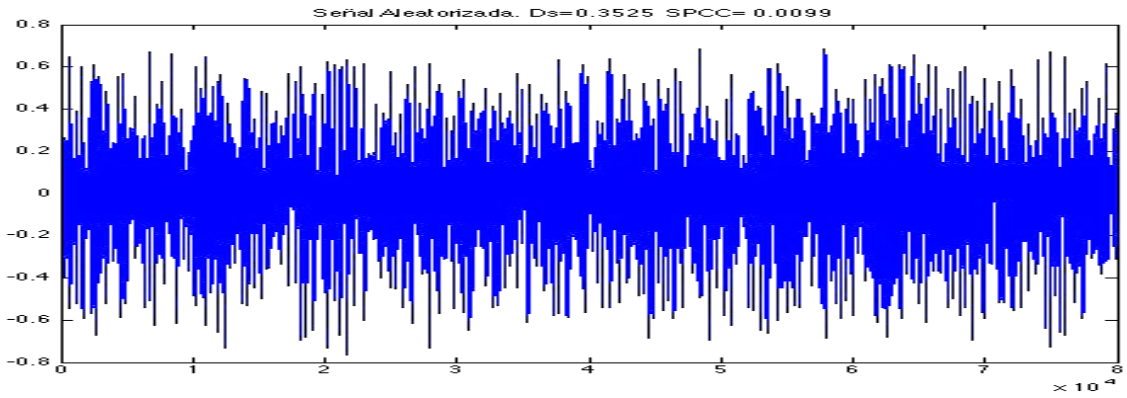


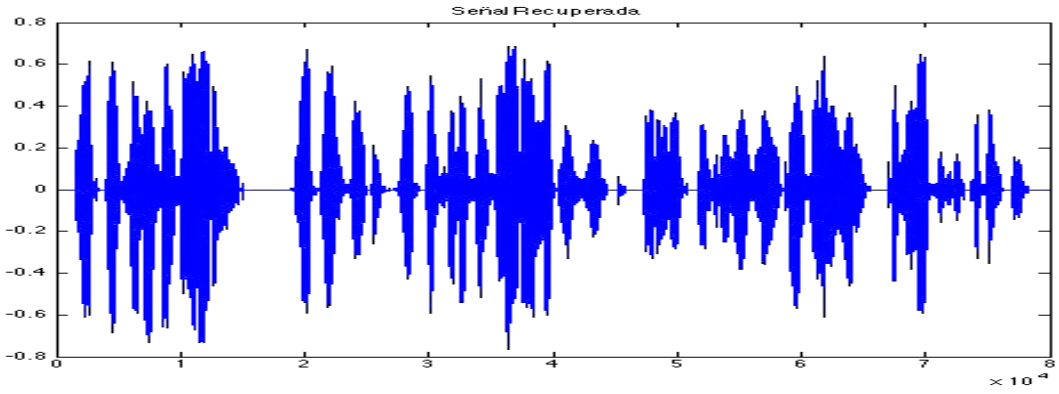
Figura 12 Ejemplo de las señales resultantes al implementar el modelo final. Ds de la señal original (0,2280), Ds de la señal aleatorizada (0,3435) y SPCC (0,0099).



a) Ejemplo de una señal original con contenido secreto y un Nivel de desorden de 0.1853



b) Ejemplo de un hijo de primera generación con un Nivel de desorden de 0.3525 y un SPCC de 0.0099 con respecto al padre.



c) Ejemplo de la señal recuperada utilizando la clave generada durante el proceso de aleatorización.

Figura 13 Señales resultantes en cada etapa del modelo desarrollado, implementando la ecuación (8) como función objetivo.

Como se mencionó anteriormente las señales originales tienen una duración de 10 segundos y se trabajaron con una frecuencia de muestreo de 8 KHz, dando como resultado un total de 80 K muestras que conforman el padre de la primera generación (Figura 13.a).

Un hijo resultante de la primera generación que cumple con los parámetros asignados en la función objetivo, no presenta rastros del mensaje secreto y presenta similitud a una señal de ruido (Figura 13.b). En la evaluación perceptual de los hijos resultantes de este caso, se encontró que la inteligibilidad residual es muy baja a diferencia de los hijos resultantes con el parámetro SD. En cuanto al Ds, se muestra un incremento aproximadamente del doble del nivel de desorden de genes del hijo, con respecto a los del padre.

Durante el proceso de aleatorización se genera una clave que permite revertir el proceso y recuperar el mensaje secreto que se tenía originalmente (Figura 13.c), las señales resultantes también fueron evaluadas con el índice SPCC, el cual se midió entre la señal recuperada y la señal original. Se encontró que las dos señales presentan bastante similitud a tal grado que el SPCC toma valor de 1. Esto comprueba que el modelo propuesto es eficiente tanto en la etapa de aleatorización como en la de recuperación del mensaje.

8. ANÁLISIS DE ROBUSTEZ Y SEGURIDAD

El modelo propuesto en este trabajo, ha demostrado ser seguro y cumplir con los objetivos para los que fue desarrollado. Una de las características que cabe resaltar, es el hecho de que la señal resultante del proceso de aleatorización, presenta similitud con una señal de ruido y por esta razón difícilmente levantaría sospechas de transmitir un contenido secreto, en caso de ser interceptada por un usuario no autorizado. Adicionalmente la señal original y la señal aleatorizada son de igual longitud, pero el grado de similitud entre las dos señales es muy bajo.

Como se mostró en apartados anteriores, al seleccionar los valores adecuados para los parámetros de la función objetivo, las señales aleatorizadas cumplen matemática y perceptualmente con características que garantizan la seguridad y robustez del modelo propuesto, como lo son la baja inteligibilidad residual (medida perceptualmente con el MOS), un alto grado del nivel de desorden de las muestras que conforman la señal aleatorizada con respecto a su original (nivel de desorden o Ds) y muy poca correlación de las mismas (SPCC).

La clave para recuperar la inteligibilidad del mensaje secreto se genera durante el proceso de aleatorización del mismo. Por esta razón la clave es variable al igual que su longitud con cada proceso de aleatorización, aún si la señal que contiene el mensaje secreto es la misma, es decir que para cada mensaje aleatorizado existe una clave de recuperación correspondiente. La selección de los cromosomas a permutar no obedece a ninguna secuencia existente o predefinida, por el contrario, es generada de forma aleatoria aumentando la robustez del sistema y ofreciendo un cierto nivel de seguridad. En relación con lo mencionado anteriormente se puede deducir que la inteligibilidad del mensaje secreto no puede ser recuperada, si no se cuenta con la matriz de la clave obtenida durante el proceso de aleatorización. El método también es altamente resistente a los ataques de fuerza bruta debido al gran tamaño de las claves generadas.

Durante las diferentes simulaciones se comprobó que el método tiene validez para diversas señales de voz, sin importar las características que las conformen tales como, la duración, el idioma del mensaje o el género del hablante. Esto quiere decir que la seguridad y la eficiencia del modelo propuesto son independientes del mensaje de voz original.

En general las características que hacen seguro y robusto al modelo propuesto son: la señal aleatorizada posee baja inteligibilidad residual y esta no depende de la señal original. La clave y su longitud son altamente variables debido a su proceso de generación al azar y particular para cada proceso de aleatorización. El proceso es completamente reversible permitiendo la recuperación de la inteligibilidad del mensaje secreto con muy alta calidad.

9. COMPARACIÓN CON OTRAS TÉCNICAS DE SCRAMBLING DE VOZ

En esta sección se compara el método propuesto con algunas de las técnicas más recientes de aleatorización de mensajes de voz encontradas en la literatura. Los métodos comprados son los siguientes:

- **Método 1. Espacio de dimensión variable:** Este método cambia la dimensión de las coordenadas de los parámetros de cifrado y usa matrices al azar para aleatorizar las señales de voz. Esto permite incrementar la seguridad del sistema debido a la posibilidad de combinar los parámetros de cifrado y generar una gran matriz de aleatorización que mejora el rendimiento en cuanto costo computacional [12].

- **Método 2. Autómata Celular:** Se introducen los autómatas celulares en el campo de la aleatorización debido al gran potencial que tienen para romper la correlación efectivamente en las señales de audio. Se analizan, comparan y prueban diferentes tipos de autómatas celulares con el fin de determinar los valores adecuados para el parámetro Lambda y calcular el número de regla de transición [2,13].

- **Método 3. Matrices Hadamard:** Este método pretende mejorar las debilidades de los modelos de aleatorización basados en las matrices Hadamard. Incorporando una transformación lineal se incrementa el número de claves que permiten obtener señales con baja inteligibilidad residual [14].

- **Método 4. Imitación de una señal de ruido Gaussiano:** El objetivo del método es que la señal aleatorizada suene como una señal de ruido Gaussiano, tomando ventaja de la similitud tanto de las estadísticas como de la entropía, existente entre las dos señales [10].

Los diferentes métodos se compararon según las siguientes características: inteligibilidad residual en la señal aleatorizada, calidad del mensaje secreto recuperado y seguridad del método, así:

9.1 Comparación de inteligibilidad residual

Método 1	La inteligibilidad residual es muy baja. La señal aleatorizada presenta una alta tasa de error de bits y baja relación señal a ruido.
Método 2	La inteligibilidad residual es altamente dependiente de las condiciones iniciales (Reglas de transición, reglas de vecindad, número de generaciones).
Método 3	Puede ser muy baja pero depende de las características de la matriz de aleatorización. (Ej. Números diferentes de cero y cantidad de coordenadas).
Método 4	No hay rastros del contenido de la señal original. Necesita una gran base de datos de señales objetivo.
Método Propuesto	La inteligibilidad residual es muy baja. Depende de la función objetivo con que se trabaje.

9.2 Comparación de la calidad del mensaje secreto recuperado

Método 1	El proceso es completamente reversible y con un bajo costo computacional. La calidad disminuye levemente con los ataques comunes.
Método 2	El proceso es completamente reversible.
Método 3	El proceso permite recuperar una señal con características muy similares a la original.
Método 4	El proceso es completamente reversible.
Método Propuesto	El proceso es completamente reversible.

9.3 Comparación de la seguridad

Método 1	El sistema es resistente a los ataques comunes. La eficiencia y seguridad dependen de las dimensiones con que se trabaje.
Método 2	El sistema puede ser vulnerado. En [13] el espacio de la clave es fijo y corto (2^{96}) y no depende de la longitud del audio
Método 3	Presenta mayor seguridad que los modelos convencionales basados en permutación de segmentos de tiempo. El número de claves es limitado.
Método 4	Trabaja con el concepto de secreto perfecto
Método Propuesto	Seguro. Resistente a ataques de fuerza bruta debido a la gran variabilidad de la clave.

De acuerdo con las comparaciones realizadas se puede concluir que existen distintos esquemas que cumplen con el objetivo de destruir la inteligibilidad del mensaje secreto, pero que en cuanto a seguridad y robustez presentan algunas dificultades, como por ejemplo: baja resistencia a los ataques, vulnerabilidades debido a espacios cortos o fijos para las claves y dependencia de los mensajes originales. En comparación, el esquema propuesto en este trabajo posee una buena relación entre la inteligibilidad residual, la calidad del mensaje recuperado y su seguridad.

10. CONCLUSIONES

El modelo desarrollado permite aleatorizar mensajes de voz o audio de diversas características, como lo son el tiempo de duración, idioma y género, utilizando un mecanismo de reproducción asexual. Los mensajes aleatorizados presentan baja inteligibilidad residual y no contienen rastros de información oculta, adicionalmente la señal resultante es similar a una señal de ruido. El proceso es reversible y por lo tanto permite recuperar la señal con el mensaje original. Dichas características garantizan que la información secreta pueda ser transmitida de forma segura.

El SPCC no puede ser tomado como el único parámetro en la función objetivo del algoritmo, ya que al incrementar la cantidad de genes por cromosoma, se llega a un punto donde pierde validez debido al aumento de la inteligibilidad residual. Se incluye el parámetro matemático DS como complemento de la función objetivo y con el fin de lograr resultados más eficientes en la etapa de aleatorización del método propuesto.

La clave de recuperación se genera en el proceso de forma aleatoria y varía con cada simulación. Su longitud varía dependiendo del número de iteraciones realizadas por el modelo, por lo cual no es predecible y hace que el modelo sea bastante seguro.

En cuanto a costo computacional, se introdujo una restricción para que los cromosomas solo puedan ser permutados una vez. Esto disminuye el número de iteraciones realizadas por el modelo al mínimo y consecuentemente el costo computacional. No se recomienda trabajar con valores altos de genes por cromosoma ya que esto trae como consecuencia el aumento de la inteligibilidad residual del audio aleatorizado. Un número de 16 genes por cromosoma permite obtener una buena relación entre costo computacional y robustez del audio codificado.

11. REFERENCIAS

[1] L. Zeng, X. Zhang, L. Chen, Z. Fan, Y. Wang (2012). Scrambling-based speech encryption via compressed sensing. EURASIP Journal on Advances in Signal Processing, Springer International Publishing AG

[2] A. Madain, A., A.L. Dalhoum, H. Hiary, A. Ortega, M. Alfonseca (2014). Audio scrambling technique based on cellular automata. Multimedia Tools and Applications, vol 71, no. 3, pp. 1803-1822, doi: 10.1007/s11042-012-1306-7

[3] M. Marseguerra, E. Zio (2000). Genetic Algorithms: Theory and applications in the Safety Domain. Recuperado de: http://users.ictp.trieste.it/~pub_off/lectures/lns005/Number_2/Marseguerra.pdf (Octubre 17/2014)

[4] S. Hasan, S. Chordia, R. Varshneya. (2012). Genetic Algorithm. [Presentación PDF]. Recuperado de: http://www.cse.iitb.ac.in/~cs344/seminars_2012/ga.pdf (Octubre 17/2014)

[5] Holland, J. H. (1975). Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control and artificial intelligence. Ann Arbor, Míchigan, Estados Unidos: U Michigan Press.

[6] J. Cantó, et al. (2009). A simple algorithm for optimization and model fitting: AGA (asexual genetic algorithm). Astronomy & Astrophysics, vol. 501, no. 3, 2009, pp. 1259-1268.

[7] M. Amirghasemi and R. Zamani, (2015). An effective asexual genetic algorithm for solving the job shop scheduling problem. Computers & Industrial Engineering, vol. 83, no. 0, 2015, pp. 123-138.

[8] A. Simoes and E. Costa, (2000). Using genetic algorithms with sexual or asexual transposition: a comparative study. Proc. Evolutionary Computation. Proceedings of the 2000 Congress on, 2000, pp. 1196-1203 vol.1192.

[9] P. Chakroborty and A. Manual, (2005). An asexual genetic algorithm for the general single vehicle routing problem. Engineering Optimization, vol. 37, no. 1, pp. 1-27; DOI 10.1080/03052150410001721468.

[10] Ballesteros L., D.M., Renza, D., Camacho, S.: An unconditionally secure speech scrambling scheme based on an imitation process to a gaussian noise signal. Journal of Information Hiding and Multimedia Signal Processing 7(2), In press (2016).

[11] M. Weik (2000), Computer Science and Communications Dictionary, vol. II.

Springer Science & Business Media, pp. 1632-1642.

[12] H.Li, Z. Quin, L. Shao, S. Zhang, (2009). Audio Scrambling Algorithm based on Variable Dimension Space. International Conference on Industrial an Innformation Systems.

[13] S. N. George, N. Augustine, and D. P. Pattathil (2014). Audio security through compressive sampling and celular autómata. Multimedia Tools and Aplications, pp. 1-25.

[14] V. Senk, V. D. Delié, and V. S. Milosevié (1997). A New Speech Scrambling Concept Based on Hadamard Matrices. IEEE Signal Processing Letters, vol. 4, No.6.