Revista ELECTRO, Vol. 47, 2025, pp. 13 - 16
https://itchihuahua.mx/revista_electro
ISSN Electrónico: 3061-774X

ESTEGANOGRAFÍA LINGÜÍSTICA EN TEXTO EN ESPAÑOL.

María del Rocío Hernández Flores[1], Bárbara Emma Sánchez Rinza[2], Mario Rossainz Lopez[1]
Benemérita Universidad Autónoma de Puebla
Faculta de Ciencias de la Computación[1], Faculta de Ciencias Físico Matemáticas[2]
Ciudad Universitaria
hf223470495@alm.buap.mx, barbara.sanchez@correo.buap.mx, mario.rossainz@correo.buap.mx

RESUMEN

La esteganografía lingüística permite ocultar mensajes dentro de textos naturales, disimulando su existencia. Este trabajo presenta una herramienta en Python que inserta periódicamente palabras de un mensaje secreto en textos generados automáticamente, usando corpus lingüísticos para mantener la naturalidad del texto portador. La aplicación incluye una interfaz gráfica y opera en español e inglés. Los resultados muestran que el sistema es determinista y reversible, recuperando mensajes correctamente bajo parámetros conocidos. Sin embargo, la técnica presenta limitaciones: es vulnerable a análisis estadísticos y lingüísticos, su evaluación se basó en ejemplos ilustrativos sin métricas objetivas, y depende de un corpus específico que limita la diversidad de textos. Como trabajo futuro se propone integrar modelos generativos avanzados (GPT, BERT), evaluar la resistencia frente a detección automatizada y comparar con otras técnicas de esteganografía textual. Este estudio constituye una contribución inicial y demostrativa en esteganografía lingüística en español.
Palabras Clave: esteganografía, criptografía, corpus lingüísticos.

ABSTRACT

Linguistic steganography allows hiding messages within natural texts, concealing their existence. This work presents a Python tool that periodically inserts words from a secret message into automatically generated texts, using linguistic corpora to preserve the naturalness of the carrier text. The application includes a graphical interface and supports both Spanish and English. Results show that the system is deterministic and reversible, correctly recovering messages under known parameters. However, the technique has limitations: it is vulnerable to statistical and linguistic analysis, its evaluation relied on illustrative examples without objective metrics, and it depends on a specific corpus, which limits text diversity. Future work includes integrating advanced generative models (GPT, BERT), evaluating resistance against automated detection, and comparing with other textual steganography techniques. This study represents an initial and demonstrative contribution to linguistic steganography in Spanish.
Keywords: steganography, cryptography, linguistic corpora

1. INTRODUCCIÓN

La esteganografía es una disciplina utilizada como técnica para ocultar mensajes de manera que no sean detectados por observadores no autorizados. Su nombre proviene del griego steganos, que significa "cubierto" u "oculto", y grapho, que significa "escribir" [1]. Esta práctica ha sido empleada desde tiempos antiguos como una forma de asegurar la confidencialidad de las comunicaciones, mucho antes del desarrollo de los sistemas criptográficos modernos [2].

A diferencia de la criptografía, que se enfoca en cifrar el contenido de un mensaje para hacerlo ilegible a quienes no poseen la clave de descifrado, la esteganografía tiene como objetivo principal ocultar la existencia misma del mensaje [3]. Esto se logra insertando información en medios como imágenes, archivos de audio o video, de manera que el mensaje permanezca imperceptible para el observador casual y no altere perceptiblemente el contenido original del medio portador [4].

El proceso de esteganografía puede llevarse a cabo mediante diversas técnicas, dependiendo del tipo de archivo que se utilice como portador. En imágenes, por ejemplo, es posible modificar los valores de los píxeles de forma imperceptible al ojo humano, permitiendo almacenar información binaria [5]. De manera similar, en archivos de audio y video, se pueden alterar ciertos bits de la señal sin afectar de forma perceptible la calidad del archivo [6].

Entre las técnicas más comunes de esteganografía se encuentran:

a) Inserción en el último bit (LSB - Least Significant Bit): Consiste en modificar el último bit de cada píxel en una imagen o el último bit de una muestra de audio, lo que permite ocultar información sin ser detectada [7].
b) Transformaciones de dominio: Implican la manipulación de la representación de los datos en un dominio matemático transformado, como la Transformada Discreta del Coseno (DCT) en imágenes [8].
c) Esteganografía en archivos de texto: Incluye métodos como la manipulación de espaciados, la sustitución de caracteres o el uso de caracteres invisibles para ocultar mensajes dentro de un texto aparentemente inofensivo [9].

A pesar de la diversidad de técnicas, muchos métodos presentan limitaciones frente a análisis automatizados o ataques estadísticos, por lo que la evaluación rigurosa y el uso de corpus representativos son esenciales para garantizar la efectividad del enfoque.

2. ESTEGANOGRAFÍA LINGUISTICA

La esteganografía es la disciplina que se encarga de ocultar información dentro de otro medio, de manera que la existencia del mensaje oculto pase inadvertida. En el caso particular de la esteganografía lingüística basada en texto, la información se

disfraza dentro de textos naturales. Este enfoque ha cobrado relevancia en los últimos años debido a la ubicuidad de los datos textuales en las comunicaciones digitales [10].

La esteganografía en texto se clasifica en distintos métodos según el nivel lingüístico en el que se opere. Estos niveles incluyen caracteres, palabras, frases y estructuras gramaticales [11]. Entre las técnicas más comunes se encuentran la esteganografía basada en características tipográficas, en la que se modifican aspectos visuales del texto —como el espaciado entre palabras o el uso de caracteres invisibles—, y la esteganografía semántica y sintáctica, que manipula las palabras y estructuras de las frases para ocultar información sin alterar el significado aparente del texto [12].

Una de las primeras aproximaciones a la esteganografía textual se centró en la manipulación de caracteres individuales o propiedades tipográficas, utilizando espacios adicionales o caracteres no imprimibles para codificar información binaria. También se ha empleado la alteración de estilos de fuente o espaciado [13]. Aunque estos métodos pueden ser eficaces, son susceptibles de ser detectados mediante análisis automatizados si el formato del texto se examina en detalle.

En la esteganografía basada en palabras, se seleccionan o reemplazan términos siguiendo patrones que codifican información, como la sustitución de sinónimos o el uso de palabras clave. Este tipo de técnica tiene la ventaja de ser menos vulnerable a sistemas de detección automatizados basados en análisis estadístico; sin embargo, puede presentar debilidades si las selecciones alteran significativamente el estilo o tono del texto [14].

Las técnicas avanzadas de esteganografía textual operan a nivel sintáctico y semántico, permitiendo ocultar mensajes mediante el reordenamiento de frases o la inserción de errores gramaticales controlados, sin afectar la gramática ni el significado general del mensaje. Estas técnicas son más sofisticadas y resistentes al análisis, aunque su implementación resulta más compleja y requiere modelos lingüísticos robustos [15].

Con el auge de la inteligencia artificial y el desarrollo de modelos de lenguaje como GPT y BERT, han surgido métodos más avanzados para generar texto esteganográfico mediante redes neuronales. Estos sistemas permiten incorporar mensajes en textos que mantienen una apariencia natural, adaptando la distribución estadística del lenguaje para evitar anomalías detectables por herramientas automatizadas [16].

Las aplicaciones de la esteganografía textual incluyen comunicaciones seguras, marcas de agua digitales y sistemas de autenticación. Se prevé que, en el futuro, la integración de estas técnicas con modelos de inteligencia artificial permita el desarrollo de métodos más robustos y adaptativos, capaces de evadir sistemas de detección avanzados [17].

3. ARQUITECTURA DEL CÓDIGO

En este trabajo implementa un sistema de esteganografía textual en idioma español mediante la inserción de palabras de un mensaje secreto en posiciones específicas dentro de un texto portador generado automáticamente. El programa se apoya en la biblioteca Natural Language Toolkit (NLTK) y en el corpus lingüístico CESS-ESP para garantizar la naturalidad del contenido generado, así como en una interfaz gráfica (Figura 1) construida con Tkinter para facilitar la interacción con el usuario. Las funciones principales del código se describen en las secciones 3.1 – 3.

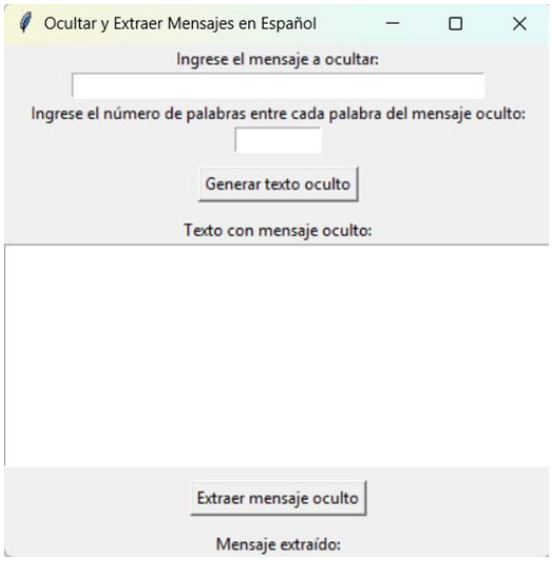

Figura 1. Interfaz Gráfica del Cifrador con Esteganografía Lingüística.

3.1. Generación del texto portador (generar_texto_base)

Esta función toma como entrada un mensaje a ocultar y una longitud total deseada para el texto. Extrae palabras del corpus CESS-ESP, seleccionándolas aleatoriamente para construir un texto que simule coherencia lingüística. La longitud del texto se ajusta para dejar espacio suficiente para insertar el mensaje oculto sin superar el límite predefinido.

3.2. Inserción del mensaje (ocultar_mensaje)

Dado un texto portador, un mensaje y un parámetro de intervalo, esta función inserta las palabras del mensaje en posiciones regulares del texto base, concretamente cada n palabras, donde n es el intervalo definido por el usuario. Esto permite dispersar el mensaje de forma controlada sin alterar significativamente la estructura superficial del texto, logrando así ocultar su presencia de manera efectiva.

3.3. Extracción del mensaje (extraer_mensaje)

A partir de un texto con mensaje oculto, un intervalo y la longitud original del mensaje, esta función recupera las palabras ocultas mediante la lectura periódica del texto en los mismos intervalos establecidos durante la ocultación. Este procedimiento es determinista y simétrico respecto al método de inserción.

Revista ELECTRO, Vol. 47, 2025, pp. 13 - 16
https://itchihuahua.mx/revista_electro
ISSN Electrónico: 3061-774X

4. RESULTADOS

Para evaluar el funcionamiento del sistema propuesto, se realizaron pruebas utilizando mensajes en español de longitud variable. En la Figura 2 se muestra un ejemplo del resultado generado por el sistema tras ocultar un mensaje utilizando un intervalo definido por el usuario. El texto resultante mantiene una estructura gramatical aparentemente natural, lo cual dificulta la detección visual del mensaje oculto.

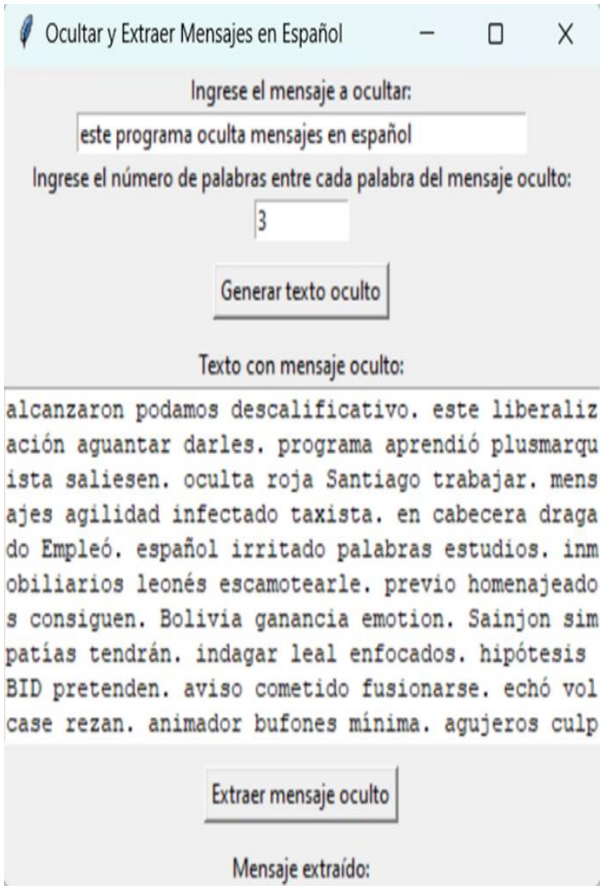

Figura 2. Texto con mensaje oculto generado mediante inserción de palabras a intervalos regulares.

Posteriormente, se aplicó el proceso de extracción utilizando el mismo valor de intervalo y la longitud original del mensaje. El sistema fue capaz de recuperar el mensaje con alta fidelidad, como se observa en la Figura 3. En todas las pruebas realizadas, el mensaje fue extraído correctamente siempre que se conservaran los parámetros iniciales (intervalo y longitud del mensaje original).

Estas pruebas confirman que el método propuesto es determinista y reversible bajo condiciones controladas. Además, el uso de un corpus lingüístico garantiza que el texto portador mantenga coherencia básica, lo cual es relevante para evadir mecanismos de detección automática.

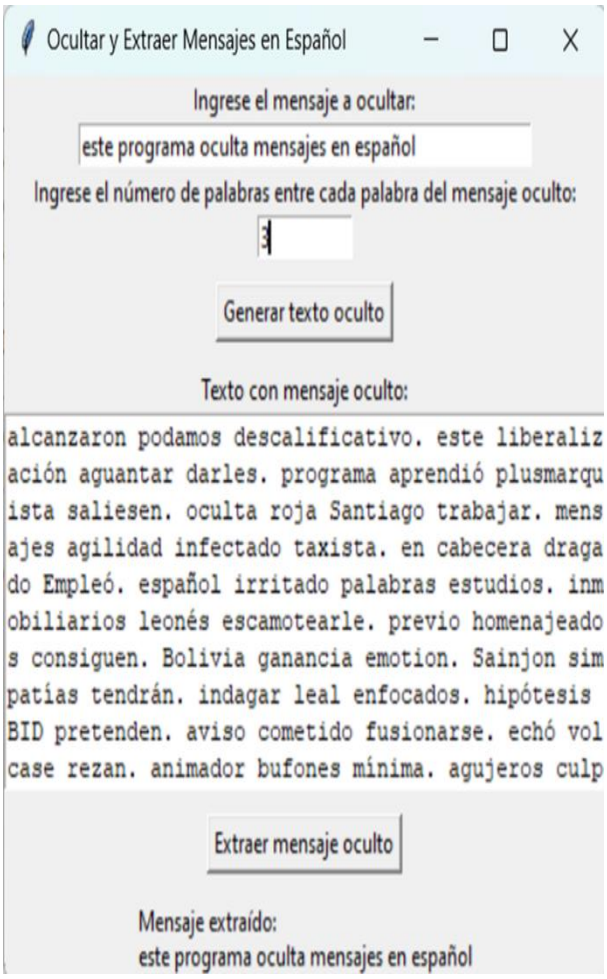

Figura 3. Recuperación del mensaje original a partir del texto con contenido oculto.

Limitaciones observadas:

- Vulnerabilidad del método de inserción periódica frente a análisis estadísticos o lingüísticos.
- Evaluación restringida a ejemplos ilustrativos, sin métricas objetivas de naturalidad o resistencia a detección.
- Dependencia del corpus CESS-ESP, que limita diversidad y naturalidad.

5. CONCLUSIONES

Se presentó una herramienta funcional para la esteganografía textual en español, integrando procesamiento de lenguaje natural, técnicas de ocultamiento basadas en intervalos y una interfaz gráfica intuitiva. El sistema permite ocultar y recuperar mensajes de manera fiable bajo condiciones controladas.

Limitaciones:

- Vulnerabilidad frente a análisis automatizados.
- Evaluaciones limitadas sin métricas cuantitativas.
- Dependencia de un corpus específico que restringe diversidad de textos.

Trabajos futuros:

- Integrar modelos generativos avanzados (GPT, BERT) para mejorar coherencia y naturalidad.
- Evaluar resistencia frente a sistemas de detección estadísticos y de machine learning con métricas objetivas.
- Comparar experimentalmente con otras técnicas de esteganografía textual, incluyendo sinónimos, transformaciones sintácticas o estrategias híbridas.

En resumen, el artículo constituye una contribución inicial y demostrativa para la esteganografía lingüística en español, con utilidad como base didáctica o prototipo de investigación, aunque su consolidación científica requiere metodologías más robustas, evaluaciones sistemáticas y alineación con tendencias actuales en PLN e inteligencia artificial.

REFERENCIAS

[1] J. Fridrich, Steganography in Digital Media: Principles, Algorithms, and Applications. New York, NY, USA: Cambridge Univ. Press, 2009.

[2] N. F. Johnson and S. Katzenbeisser, "A survey of steganographic techniques," in Information Hiding Techniques for Steganography and Digital Watermarking, S. Katzenbeisser and F. A. P. Petitcolas, Eds. Norwood, MA, USA: Artech House, 2000, pp. 43–78.

[3] M. Kharrazi, H. T. Sencar, and N. Memon, "Image steganography: Concepts and practice," in Proc. 2nd Int. Workshop on Digital Watermarking, 2003, pp. 1–13.

[4] R. Chandramouli and N. Memon, "Analysis of LSB based image steganography techniques," in Proc. IEEE Int. Conf. on Image Processing, 2001, vol. 3, pp. 1019–1022.

[5] A. Cheddad, J. Condell, K. Curran, and P. Mc Kevitt, "Digital image steganography: Survey and analysis of current methods," Signal Processing, vol. 90, no. 3, pp. 727–752, Mar. 2010.

[6] C. Cachin, "An information-theoretic model for steganography," in Proc. 2nd Int. Workshop on Information Hiding, Portland, OR, USA, 1998, pp. 306–318.

[7] N. Provos and P. Honeyman, "Hide and seek: An introduction to steganography," IEEE Security & Privacy, vol. 1, no. 3, pp. 32–44, May–Jun. 2003.

[8] S. Katzenbeisser and F. A. P. Petitcolas, Information Hiding Techniques for Steganography and Digital Watermarking. Norwood, MA, USA: Artech House, 2000.

[9] M. Topkara, U. Topkara, and M. J. Atallah, "Words are not enough: Sentence level natural language watermarking," in Proc. ACM Workshop on Content Protection and Security, 2006, pp. 37–46.

[10] M. Topkara, M. Topkara, and M. J. Atallah, "Words are not enough: Sentence level natural language watermarking," in Proc. ACM Workshop on Content Protection and Security, Alexandria, VA, USA, 2006, pp. 37–46.

[11] H. Murakami, S. Uchida, and H. Sakoe, "A framework for text steganography using natural language processing," in Proc. Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing, Pasadena, CA, USA, 2006, pp. 400–403.

[12] F. Meng, X. Zhang, H. Wang, and Y. Liu, "Text steganography using linguistic structure," in Proc. IEEE Int. Conf. on Computer Science and Automation Engineering, Shanghai, China, 2012, pp. 564–568.

[13] R. M. Low, R. D. T. Sagar, and T. D. Nguyen, "Text steganography: An overview," in Proc. 2016 3rd Int. Conf. on Computing for Sustainable Global Development, New Delhi, India, 2016, pp. 3486–3490.

[14] A. Desoky, "Linguistic steganography: Survey, analysis, and robustness concerns for hiding information in text," IEEE Trans. on Information Forensics and Security, vol. 6, no. 3, pp. 491–509, Sep. 2011.

[15] A. Bennett and P. R. Palanisamy, "A syntactic approach to natural language text steganography," in Proc. IEEE Int. Conf. on Computational Intelligence and Computing Research, 2010, pp. 1–4.

[16] Y. Yang, Y. Li, and H. Zhang, "Hiding messages in natural language text via linguistic steganography with pre-trained language models," in Proc. 29th ACM Conf. on Computer and Communications Security, 2022, pp. 1413–1427.

[17] R. Z. Liu and C. Wang, "Towards robust steganographic text generation with pre-trained language models," IEEE Access, vol. 10, pp. 134528–134541, 2022.