Esteganografía lingüística en redes sociales. Perspectiva de futuro en lengua española

Alfonso Muñoz Muñoz
Escuela de I.T de Telecomunicación
Universidad Politécnica de Madrid
Email: amunoz@diatel.upm.es

Justo Carracedo Gallardo
Escuela de I.T de Telecomunicación
Universidad Politécnica de Madrid
Email: carracedo@diatel.upm.es

Jorge Ramió Aguirre
Escuela Universitaria de Informática
Universidad Politécnica de Madrid
Email: jramio@eui.upm.es

***Abstract*—El presente artículo analiza las posibilidades de la distribución de estegotextos en redes sociales mediante los últimos avances desarrollados en esteganografía lingüística en lengua española. Existen muchas limitaciones a considerar si se desea ocultar información en textos en lenguaje natural de manera no trivial. En la práctica, la decisión de un procedimiento de ocultación u otro condicionará, desde un punto de vista práctico, el canal (en este caso red social) donde será más fácil transmitir una información oculta; esta característica podría facilitar el trabajo a un potencial estegoanalista. A modo de ejemplo, se analizan algunas características de la red social Twitter.**

I. CONCEPTOS PREVIOS. ESTEGANOGRAFÍA LINGÜÍSTICA

La ciencia de la esteganografía puede ser definida como la ciencia y el arte de ocultar una información dentro de otra que haría la función de tapadera [1], con la intención que la existencia de dicha información no sea percibida. En teoría, sólo quienes conozcan cierta información acerca de esa ocultación (un secreto) estarían en condiciones de descubrirla. Cuando la cubierta es un texto en lenguaje natural ello implica un tipo específico de esteganografía, esteganografía lingüística, y el texto que oculta dicha información es llamado estegotexto.

La idea de ocultar información en textos en lenguaje natural no es ni mucho menos nueva. En los últimos siglos diferentes procedimientos clasificables en *open codes* (cues, null ciphers, jargon code y grilles) y *semagrams* han sido documentados [2]. Algunos ejemplos famosos son *newspaper code* en la época victoriana o la verja de cardano en el siglo XVI [3].

En la actualidad, la esteganografía lingüística intenta aunar los principios de la ciencia de la esteganografía y la lingüística computacional (análisis automático del contenido textual, el análisis morfosintáctico, generación textual, la lexicografía computacional, descripciones ontológicas, etc.) para crear procedimientos no triviales no basados en la oscuridad. En ese esfuerzo dos amplias líneas de investigación son abordadas [4]: a) la modificación de textos existentes y b) la generación automática de estegotextos.

II. ESTEGANOGRAFÍA EN REDES SOCIALES. ANTECEDENTES

En 2008 [5] generalizando el uso de la aplicación de la herramienta de estegoanálisis StegSecret [6] a Internet, se analizó el potencial de las características de las redes sociales con fines esteganográficos. En este trabajo [5] se analizó en primer caso los procedimientos esteganográficos más útiles para los estegomedios más probables en estas redes, como son el contenido multimedia (imagen, audio, video) y las tecnologías web (html, xml, http). Del mismo modo, se analizó la dificultad por parte de un atacante de acceder a posibles datos esteganografiados si éstos aprovechaban al máximo los mecanismos antibot/captchas y los grupos cerrados de usuario para evitar herramientas automatizadas de detección, así como se destacó la utilidad de distribuir contenido de forma multiportador y multiproveedor para evitar análisis por parte de proveedores con otros fines que los redactados en sus políticas de privacidad.

Es en este contexto donde se decidió profundizar en el potencial de la ocultación de información en lenguaje natural de manera no trivial. La información textual está en todas partes, de forma masiva y con características lingüísticas muy diversas, esto la convierte en lo suficientemente interesante de manera individual y en su distribución en las nuevas redes sociales.

III. POTENCIAL DE LA ESTEGANOGRAFÍA LINGÜISTICA EN ESPAÑOL

En los últimos 5 años las publicaciones sobre esteganografía lingüística en lenguas tan diversas como el inglés, ruso o mandarín van en aumento, con utilidad en la ocultación de información en texto en lenguaje natural y con utilidad en el marcado digital de textos. Por desgracia, las publicaciones en lengua española son escasas. Debido a su potencial, en el último año y medio hemos realizado un esfuerzo en acotar la utilidad de diferentes procedimientos con aplicación en esteganografía lingüística en español.

La utilización del lenguaje natural con fines esteganográficos de manera segura, considerando criterios lingüísticos (léxicos, sintácticos, semánticos, de cohesión, de coherencia, etc.) y esteganográfico-estadísticos (análisis de entropía, análisis de frecuencia de caracteres-palabras, ataques basados en conocimiento de cubierta original y cubierta modificada, etc), no es nada sencillo.

La modificación de textos a menudo suele introducir errores detectables por un lector humano. Este inconveniente inicial

se convierte de utilidad para el productor de estos estegotextos ya que permite cuantificar en una primera instancia cómo de bueno es un procedimiento esteganográfico concreto de manera sencilla. La detección de anomalías por parte de un lector humano puede que sea mayor a las detectadas por algoritmos clasificadores automáticos.

En la práctica, el lenguaje natural, como estegomedio, es muy poco redundante (ruidoso) en comparación con otros estegomedios como son las imágenes o los videos, lo cual hace más complicado la creación de algoritmos robustos de ocultación de información en lenguaje natural (información textual). Esta condición hace que las técnicas de ocultación requieran una gran cantidad de información textual para ocultar una cantidad de información no muy abultada. Este hecho puede hacer que estas técnicas sólo sean interesantes en escenarios limitados. Por ejemplo, en Internet ciertos canales serían más aptos para ocultar información que otros.

En el objetivo de analizar qué técnicas esteganográficas en lenguaje natural para lengua española son más interesantes de detectar se analizaron los siguientes procedimientos englobados en las siguientes dos grandes líneas de investigación.

a) Generación automática de estegotextos.

Una buena alternativa para la ocultación de información en textos en lenguaje natural en español sería la generación automática de estegotextos. Su ventaja fundamental reside en la posibilidad de crear estegotextos únicos para cada comunicación, de forma que se dificulte ataques estadísticos y ataques basados en comparaciones texto original-estegotexto, así como tener un mayor control en la modelización estadística del estegotexto creado.

En la última década los esfuerzos en este sentido se centran, principalmente, en la generación de estegotextos que imiten la gramática (sintaxis) y la estadística de un texto típico en una lengua concreta. Aunque, entre las opciones interesantes detectadas, se pueden establecer propuestas de generación automática de estegotextos basadas en construcciones CFGs (Context-Free-Grammar), derivadas de los principios de la teoría de la gramática generativa propuesta por el lingüística A.Noam Chomsky en la década de los 60, en primera instancia no se abordan estos estudios dado que tienen numerosas limitaciones a sortear, como se analizó en [7], por ejemplo, el problema de la privacidad/compartición/generación de la gramática utilizada, ataques basados en estudio de terminales (información última de cada regla), limitación del contexto de utilización y léxico utilizado, etc.

En su lugar es más productivo en primera aproximación el uso de propuestas basadas en modelos estadísticos para minimizar en mejor medida ataques estadísticos a propuestas de esteganografía lingüística.

Con estas características es interesante el uso de un modelo estadístico N-GRAM con fines esteganográficos. Es decir, dado uno o más textos de entrenamiento es posible anotar las co-ocurrencias de las palabras presentes y medir las frecuencias de repetición de cada una (modelo N-GRAM), es decir, se anota qué palabra viene detrás de otra y con qué probabilidad. El algoritmo de generación de estegotextos elige en cada ocasión una de las palabras disponibles en función de la información a ocultar. Si la selección es aleatoria es más probable que salgan las palabras más probables que las que son menos, y por tanto se imite la estadística de la fuente de entrenamiento. Si se imita la estadística de la fuente es más probable que el estegotexto resultante que está basado en los textos de entrenamiento, tenga validez léxica y sintáctica (validez léxica y sintáctica que deben tener los textos de entrenamiento). Imitar el orden de las palabras tiene utilidad esteganográfica y como demostró Greenberg [8] el orden de las palabras determina hasta tal punto la hechura de una lengua, tanto que no admite trivialidades en su manipulación.

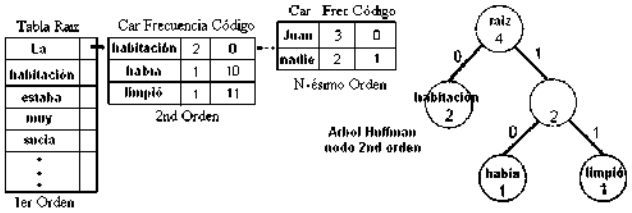

Fig. 1. Algoritmo de generación automática de estegotextos basado en modelo N-GRAM

Esta posibilidad se comprobó en lengua española mediante la implementación de la herramienta Stelin [7]. Aunque depende del texto de entrenamiento, para órdenes de n mayor de 7 y textos de referencia mayores de decenas de KB se obtiene textos con validez léxica y gramatical. El tamaño final de estegotexto generado dependerá de la información a ocultar, de los textos de entrenamiento y del orden n. La capacidad mínima de información podría aproximarse entre 0,5 a 1 bit de información oculta por palabra generada. No obstante, este tipo de algoritmos todavía podrían generar estegotextos con errores gramaticales (depende en mayor medida del texto de entrenamiento). Estos errores podrían minimizarse basándose en los principios de la ley de Zipf (pocas palabras se repiten muchas veces) y en la idea que es posible añadir palabras a medida entre las palabras añadidas en la generación automática de estegotextos sin la necesidad que el receptor las conozca de manera adicional [9]. Por ejemplo, en la Fig1, si el estegotexto creado fuera *La habitación* sería posible añadir palabras entre medias no incluidas en el nivel de la palabra *habitación* (añadir una o más palabras diferentes a *había* y *limpió*). De nuestra experiencia se observa que es posible conseguir estegotextos que oculten menos de 1.000 bits para una relación tiempo-esfuerzo razonable (si se usa edición manual) con una calidad lingüística elevada, que podría dificultar incluso la detección a lectores humanos. Este procedimiento sería de enorme utilidad para la distribución de claves criptográficas, urls, mensajes breves, coordenadas gps, etc. El tamaño del texto resultante dependería de la capacidad editora del emisor, del orden n y de los textos de entrenamiento.

Siguiendo esta idea todavía sería posible dotar al emisor

de mayor libertad en la creación de estegotextos, creando automáticamente estegotextos de peor calidad pero con mayor facilidad de edición (herramienta ryfle) [10]. Un ejemplo de ello consistiría en dividir un texto fuente en grupos de palabras de tamaño W y basar la ocultación en la selección de una de cada grupo (por ejemplo asignamos el mismo peso a cada palabra). Así por ejemplo si la información a ocultar (en bits) fuera I, el número de palabras generadas sería Nmin y el tamaño mínimo necesario (en palabras) de la fuente de entrenamiento sería Source_{min} :

$$\text{N}_{min} = I/log_2 W$$
$$\text{Source}_{min} = N_{min} * 2log_2 W$$

Los estegotextos generados no tendrían validez (ni siquiera gramatical), pero para tamaños de grupo (W=4, 8 ó 16) [10] sería fácil añadir palabras antes de cada palabra seleccionada y que no estuvieran contenidas en cada grupo. Esta sencilla idea, el algoritmo es público, proporcionaría una alta calidad lingüística dificultando la tarea de detección a un lector humano (y por tanto a un algoritmo automatizado), si bien a un mayor coste de tiempo-esfuerzo por parte del emisor. Hoy día, se analizan variantes basadas en categorías lingüísticas y co-ocurrencias para facilitar y disminuir el tiempo necesitado en la edición manual [10].

Un procedimiento de este tipo sería ideal para la distribución de una manera altamente segura de claves criptográficas o urls (128-256 bits). Aunque el tamaño del estegotexto resultante dependería de la capacidad editora del emisor, este valor podría aproximarse a 3 ó 4 veces más que el tamaño Nmin, es decir, lo que supondría añadir de 3 a 4 palabras más por palabra generada (con 3 palabras se puede construir en español una oración básica: nombre+verbo+complemento). En [10] se adjunta un ejemplo de estegotexto de 281 palabras, 0,44bits/palabra, que oculta 126 bits de información oculta.

b) Modificación de textos existentes.

La modificación de textos con utilidad esteganográfica con ciertos criterios de seguridad limita su actuación a una serie de procedimientos de ocultación relacionado con el léxico y la semántica, la gramática y la co-ocurrencia (orden de las palabras). Algunas de las propuestas más interesantes que se han documentado son [4]: sustituciones léxico-semánticas, sustituciones sintácticas, aprovechar el ruido de traducir un texto a diferentes idiomas, ocultación basada en formato y errores tipográficos/ortográficos (abreviaturas, acrónimos y otros). En este contexto se analizó las técnicas que parecen ser más productivas en otros idiomas, como por ejemplo los procedimientos en lengua inglesa.

En primer lugar es interesante analizar el potencial de utilizar el orden de las palabras dentro de una frase con utilidad esteganográfica. En [11] se analizó, investigación que sigue en curso, el potencial de las modificaciones sintácticas en lengua española con utilidad esteganográfica y utilidad en natural language watermarking. Basándonos en mediciones en el corpus LEXESP [11] se demuestra cómo en español estructuras, con utilidad en otras lenguas como el inglés, como es la transformación activa-pasiva no tiene utilidad esteganográfica, al no ser la pasiva una estructura lingüística frecuente en español. Este trabajo [11] analiza además la posibilidad de mover ciertos elementos etiquetados previamente (Part of Speech) en una oración. De esta forma a nivel de frase, los resultados indican mayor utilidad esteganográfica en el movimiento de adverbios, especialmente si éste se sitúa al principio o final de frase, y el movimiento de adjetivos con respecto al nombre que modifica (anteponiéndolo o posponiéndolo), más concretamente 1.537 parejas (nombre+adj o adj+nombre) muestran la posibilidad de mover el adjetivo delante o detrás de un nombre y por tanto de ocultar un bit por pareja.

Aunque de momento estas investigaciones están en curso, en media (con suerte) 1 ó 2 bits por frase podrían ser ocultados mediante estos procedimientos. Si bien se podrían generar estegotextos indistinguibles de textos originales, es cierto que sería necesario una cantidad razonable de texto para ocultar información, por ejemplo, 128 frases para 128 bits (entiéndase que un texto podría ser fragmentado en frases considerando puntos, partículas que subordinen oraciones, etc.).

Por otro lado, la ocultación basada en modificaciones léxico-semánticas sería posible. En [12] se analizó la posibilidad de ocultar información mediante la implementación de una herramienta avanzada de sustitución de sinónimos (23.918 sinónimos únicos y soporte para variantes verbales, de género y número). Considerando procedimientos estadísticos para disminuir el impacto de una sustitución de una palabra en el contexto cercano de palabras vecinas se pueden realizar diferentes aproximaciones, basadas en reglas de Word Sense Disambiguation. Entre las propuestas documentadas en [12] una puede basarse en la construcción de tablas estadísticas que alimenten una función de ponderación que cuantifique cómo de bueno es un sinónimo para sustituirlo en un contexto dado. Siguiendo esta idea se podría conseguir una capacidad de ocultación próxima al 29,71% de las palabras totales de un texto. La ocultación mínima en esta situación considerando 1 bit por palabra útil, sería en media de 0,2971 bits/palabra. No obstante, aunque se construya un tabla estadística significativa, nuestros experimentos están basados en un corpus de 120 millones de palabras, y sólo se realizarán modificaciones considerando los dos sinónimos más probables de cada palabra a sustituir, en función de su contexto, todavía sería posible implementar un ataque haciendo uso de un clasificador SVM que detectara al menos el 50% de los documentos con información oculta, documentos de 269 palabras y 52,45 bits/documento.

IV. ESTEGANOGRAFÍA LINGUISTICA EN REDES SOCIALES. MITOS Y REALIDADES

La distribución de información oculta en textos en lenguaje natural mediante procedimientos no basados en oscuridad hace que las opciones disponibles para ocultar información se vean seriamente reducidas. En la práctica estos procedimientos estarán basados en sustituciones léxico-semánticas, uso del orden de las palabras (modelos estadísticos) y el posible uso

de reglas gramaticales. Desde un punto de vista aproximativo y puramente conceptual procedimientos de ocultación de este tipo daría una capacidad mínima de ocultación del orden expresado en la siguiente figura.

Técnica	Capacidad Mínima	Procedimiento Automático
Modelo N-Gram (sin edición manual)	0,5 - 1 bit / palabra	SI
Herramienta Ryfle	0,44 bits / palabra	NO
Sustitución de palabras por sinónimos (uso tablas estadísticas)	0,2971 bits / palabra	SI
Modificaciones sintácticas basadas en el movimiento de palabras	1-2 bits / frase	SI

Fig. 2. Capacidad mínima de ocultación por palabra del estegotexto creado o modificado.

Según estos valores, e independientemente de la seguridad de cada técnica concreta, para ocultar, por ejemplo, una mínima información útil, que podría ser de unos 128 bits, se necesitaría fragmentos de textos de al menos: 128-256 palabras (N-Gram), 290 palabras (ryfle) y 430 palabras (sinónimos). Esta capacidad de ocultación implica la necesidad de textos de cierto tamaño. Por ejemplo, en su aproximación a las redes sociales todas estas técnicas son ideales en la publicación de estegotextos en blogs, más cuestionable puede ser (sin ser troceados o distribuidos) en comentarios en foros, mensajes en redes tipo facebook, tuenti y similares.

Obtener los tamaños medios de los mensajes en este tipo de redes no es sencillo, al estar los usuarios en grupos cerrados, ya que el acceso implicaría anular sistemas de protección o autenticación. En cualquier caso y de forma aproximativa, el tamaño de los comentarios en facebook o tuenti se encuentra entre las 10 y las 25 palabras, excepciones puntuales (en los casos manualmente comprobados) excederían hasta no más de 100 palabras.

En el caso que se utilicen otros procedimientos, basados en técnicas sin documentar/publicar, la modelación del lenguaje, el uso de la estadística y el uso de expresiones lingüísticas podría facilitar la automatización del estegoanálisis de estas técnicas no documentadas. En este sentido, dado que existe un número importante de redes sociales donde el tamaño de la información textual no permitiría (de forma simple) las técnicas documentadas (para ocultar al menos cientos de bits) podría ser interesante explotar nuevas características de ocultación. Por ejemplo, son interesantes las redes sociales basadas en comunicaciones rápidas, frases cortas, con poca preocupación en la calidad lingüística, produciéndose faltas ortográficas y tipográficas en cuantía. Un entorno donde podría ser interesante el uso de procedimientos esteganográficos basados en errores ortográficos y tipográficos sería precisamente este tipo de redes sociales donde el lenguaje no está muy cuidado. Dado que la presencia de faltas de ortografía podría hacer a un lector humano delatar rápidamente anomalías, sobre todo si un texto presenta estadísticamente más faltas de las esperadas incluso en un canal ruidoso, el enfoque se debe centrar, si esta propuesta es considerada en serio, en distribuir la información dificultando la tarea de programas automáticos en clasificar textos originales y textos con información oculta. Aunque no es sencillo articular una propuesta pública basada en estos principios, en [13] se presenta una aproximación interesante y razonablemente seria. La pregunta clave a resolver es ¿cómo sabe un programa automático que existe una falta de ortografía en un texto?. Con esta idea en mente se pueden formular diferentes sistemas de ocultación:

a) Sistemas que usan malformaciones para producir palabras que no existen en un diccionario. El ataque a estas propuestas se simplifica utilizando diccionarios grandes y modelado de estadística de uso de las palabras.

b) Sistemas basados en errores tipográficos y uso de acrónimos, abreviaturas y similares. La detección de estos procedimientos queda supedita a algoritmos similares a los utilizados en la corrección de textos (por ejemplo, en productos ofimáticos). Algunas de las técnicas que se emplean en su detección son medidas basadas en la mínima distancia de edición, modelos estadísticos n-gram, etc.

c) Malformaciones que provocan que una palabra válida se convierta en otra palabra válida (vaca/baca, toro/loro, etc), separar palabras compuestas (saca puntas/sacapuntas), errores gramaticales, etc. La detección de estos procedimientos requiere de técnicas avanzadas de reconocimiento lingüístico, etiquetadores precisos, técnicas de desambigüación y comprensión semántica.

Aparte de la aportación conceptual de [13], razonan una capacidad de ocultación de 1bit/1 palabra útil, no queda lo suficientemente claro el interés práctico de este tipo de propuestas. No ha sido posible avanzar en el análisis de esta propuesta concreta al no poder acceder a la herramienta comentada en [13] y analizar la seguridad real de los estegotextos producidos respecto de un modelo de lenguaje obtenido de textos de entrenamiento, así como ver como se comportarían estos mensajes ocultos respecto de la estadística de un canal concreto donde se transmitieran.

Analizadas muchas de las variantes esteganográficas más razonables en textos en lengua española se va a analizar a continuación algunas características de una red social concreta y se va a razonar si es productiva en términos de esteganografía lingüística.

V. ESTEGANOGRAFÍA LINGUISTICA EN TWITTER

La red social twitter es un buen ejemplo de red social donde el tamaño de los mensajes dificultaría ocultar centenas de bits (sin trocear o distribuir estegotextos) mediante una propuesta esteganográfica pública. En el pasado esta red ha sido utilizada de manera muy diversa, en el envío de información cifrada, la administración de botnets, etc [14][15].

Es en este entorno donde podría ser interesante analizar si también es útil en esteganografía lingüística, conclusiones que

se podrían derivar a otras redes similares.

Si se piensa por un momento en este tipo de redes, es factible la implementación de un número elevado de procedimientos esteganográficos basados en oscuridad y que en muchos casos recuerdan a procedimientos utilizados siglos atrás (utilizar ciertas letras, usar el número de palabras para codificar una información, etc). Por ejemplo, sería posible utilizar procedimientos clásicos de alterar las letras de cada palabra poniéndola en mayúscula o minúscula en función de si se codifica un bit 0 o un bit 1, de esta forma en el mejor de los casos se podría ocultar 140 bits (140 caracteres) por mensaje twitter. Como puede observarse sería fácilmente detectable, no obstante siempre podría ser utilizado como mecanismo para transmitir más información en los mensajes sin fines esteganográficos (en el mejor de los casos 140 bits/7 bits-carácter = 20 caracteres más) teniendo el receptor la capacidad todavía de comprender el mensaje publicado. Sin ser puramente lingüística otras técnicas de ocultación podrían ser factibles. Por ejemplo, usar los sistemas de codificación de urls compactadas utilizadas en Twitter, como por ejemplo el sistema web bit.ly, con fines esteganográficos. Por ejemplo, un enlace compactado del tipo bit.ly/aWvi3x podría ocultar información jugando con la codificación de los mismos (minúscula, mayúscula y números). Si se toma la molestia de que el enlace realmente exista (existen un gran número de ellos), un atacante no sólo debería comprobar que ese enlace apunta en realidad a una página web sino tener la capacidad a) de analizar la semántica de la frase donde está presente y correlarla con el contenido del enlace para ver si el enlace corresponde a lo que se indica o b) establecer algún tipo de criterio estadístico que permita deducir que de la presencia de muchos enlaces de este tipo se puede destacar la presencia de un sistema de codificación concreto. Como se describirá posteriormente la presencia de direcciones web en un mensaje Twitter se produce en un 23,23% de los casos.

En cualquier caso, para analizar mejor el potencial de la esteganografía ligüística en esta red, se decide la construcción de un corpus basado en 103 usuarios españoles de la red twitter, con la condición que tengan al menos 1.000 seguidores y como poco 3.200 mensajes publicados (número máximo accesible por la web). La intención es medir en un grupo de comunicaciones normales, que pretenden ser entendidas por un conjunto amplio de usuarios, como se comunican los emisores y si esa comunicación permite extraer alguna conclusión con utilidad esteganográfica. Se intenta evitar, por tanto, jerga propia de un conjunto reducido de usuarios y sí la expresión de mensajes que desean ser leídos por una mayoría pero que pueden tener las características o comodidades propias del canal donde se emiten. Se busca por tanto un modelo estadístico común de conducta con utilidad esteganográfica. Se implementa un crawler en lenguaje JAVA y se recopila esta información construyendo un corpus final de 319.381 frases y 4.029.257 palabras [16]. Para el etiquetado de este corpus se implementa un programa en lenguaje JAVA que utiliza el etiquetador TreeTagger [17] considerando un mensaje Twitter como una frase [16]. La primera característica interesante detectada es que sólo se consume en media 48,1214% de la capacidad total (67 caracteres por frase), con la presencia por tanto de frases cortas e información muy instantánea lo cual da muestra de la presencia mayoritaria de adverbios de lugar y tiempo. El 38,57% de las frases empiezan con @, un 6,32% con RT, acaban con emoticon un 6,21% y acaban con direcciones web un 23,23%. A la vista de este estudio, lingüísticamente más ampliado en [16], no parece factible en primera aproximación ocultación lingüística por procedimiento público, a excepción de las técnicas comentadas con anterioridad y ocultando menos información (al ser los mensajes twitter de pocas palabras).

Aunque al seleccionar un conjunto de usuarios públicos no se esperaba un número notorio de faltas de ortografías (que podría tener utilidad en esteganografía). En la siguiente tabla se recopilan algunas mediciones sobre faltas de ortografía donde se observa, al menos para éstas y en el corpus recopilado, que este tipo de técnicas basada en estos fallos no sería muy productiva desde el punto de vista de la esteganografía lingüística.

Estructura Punto+espacio+Término		**N° total 102.179**
Mayus	**91494**	**89,5428%**
Minus	**1654**	**1,6187%**
Número	**847**	**0,8289%**
Palabra+o+Palabra que empieza por 'o'	**11 ocurrencias**	
Palabra+y+Palabra que empieza por 'i'	**42 ocurrencias**	
Omisión de tildes en:	**Ocurrencias palabras**	**N° palabras 4.027.746**
Palabra acabada en "cion"	**1544**	**0,0383%**
Palabra acabada en "sion"	**674**	**0,0167%**
Adverbio acabado en "mente" (que debería ir acentuado)	**148**	**0,0036%**

Fig. 3. Medición de algunas erratas y faltas de ortografías de fácil computación.

Podría ser interesante orientar la investigación en este tipo concreto de redes sociales (frases breves) al uso de gramáticas libres de contexto (CFGs) [18]. En el pasado no se consideró este tipo de estructuras desde un punto de vista práctico debido a la problemática de que para crear estegotextos con una cierta cohesión-coherencia se necesita la creación de gramáticas complejas (cuya automatización no es trivial) y diccionarios de gran volumen de palabras con categorizaciones semánticas. En el caso de twitter o similares este problema puede ser minimizado en gran medida. En twitter los mensajes son frases sueltas, sin necesidad alguna de que se mantenga una coherencia global con las frases publicadas anteriormente. De hecho en la práctica muchas de las frases no tienen nada que ver con sus antecesoras. Considerando esto sería factible profundizar en la investigación de reglas gramaticales

simples y la ocultación estar basada en la selección de palabras concretas (que se eligen de un conjunto W categorizado semánticamente y por PoS/probabilidad) para cada parte de la regla gramatical definida. En la práctica, es aunar los conocimientos en modelos N-GRAM, etiquetado lingüístico y sustitución de sinónimos según contexto para avanzar en un tipo evolucionado de CFG probabilística. Si esto fuera así, y a falta (en el momento de escribir este artículo) de una implementación real, la información a ocultar se centraría en las categorías lingüísticas más probables (nombres, verbos, adjetivos y adverbios en este orden). La capacidad de ocultación dependería del sistema de codificación elegido, si la codificación tiene el mismo peso para cada palabra a seleccionar hablaríamos de una capacidad de logW (base2) por término de la regla gramatical habilitada. Independientemente de otros criterios de aceptabilidad, un mensaje twitter podría ocultar como poco del orden de 3*logWi (base2) bits, siendo Wi el conjunto de palabras para cada elemento de la regla gramatical considerada, dado que una frase sencilla al menos podría tener un nombre, verbo y complemento.

Las gramáticas CFGs son muy efectivas esteganográficamente para este objetivo, una especie de adaptación de la herramienta NICETEXT [18] con las indicaciones resaltadas anteriormente. A falta de una implementación real, queda pendiente analizar si un detector sería capaz de inferir anomalías que detectaran la presencia de información oculta por estos procedimientos.

VI. CONCLUSIÓN

En el último año y medio se ha avanzado en el análisis de la seguridad de propuestas públicas de esteganografía lingüística en lengua española. Las cifras actuales indica que la capacidad de ocultación es baja y requiere del uso de al menos textos de centenas de palabras para ocultar una cantidad de unos 128 bits útiles para distribuir una clave criptográfica, una url, etc.

El uso masivo de las redes sociales hace interesante analizar la posibilidad de ocultación de información en texto digital en ellas. En la práctica, redes sociales como tuenti, facebook, twitter o similares por la inmediatez provoca que en media los mensajes intercambiados sean de decenas de palabras. Esta limitación hace que sin trocear o distribuir los estegotextos creados, en frases o otras unidades lingüísticas, los procedimientos esteganográficos analizados serían más apropiados en blogs y si acaso en comentarios algo extensos (centenas de palabras) en foros.

La posibilidad de utilizar esteganografía lingüítica en redes sociales con los procedimientos documentados, con una seguridad no basada en oscuridad, se complica bastante. En este orden se decide analizar cómo se expresa una comunidad de usuarios determinados en la red twitter y aunque aparecen elementos que permitirían crear algún modelo esteganográfico útil resulta difícil justificar su seguridad en una propuesta pública.

Se deja abierto al futuro nuevos avances para mejorar las técnicas de esteganografía lingüística y su posible adaptación a nuevos canales masivos de comunicación, como las redes sociales o cualquier medio futuro que dificultara la detección de comunicaciones subrepticias a un potencial estegoanalista.

REFERENCIAS

[1] J. Carracedo, Seguridad en Redes Telemáticas. Mc-Graw Hill. España. ISBN:84-481-4157-1 (2004), pág 123-131.

[2] F. L. Bauer. Decrypted Secrets. Methods and Maxims of Cryptology. s.l. : Springer; 1 edition (October 2, 1997), 1997. ISBN-13: 978-3540604181.

[3] D. Kahn, The Codebreakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet. s.l. : Scribner; Rev Sub edition (December 5, 1996). ISBN-13: 978-0684831305.

[4] R. Bergmair, A comprehensive Bibliography of Linguistic Steganography. SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents, volume 6505, January 2007.

[5] A. Muñoz, J. Carracedo, S. Sánchez, Detection of distributed steganographic information in social networks. EATIS 2008. Euro American Conference on Telematics and Information Systems, September 10-12. Aracaju, Brazil. ACM-DL.

[6] A. Muñoz, J. Carracedo, StegSecret: Una herramienta de estegoanálisis pública. Mar del Plata : Congreso Iberoamericano de Seguridad de la Información. CIBSI 2007. http://stegsecret.sourceforge.net.

[7] A. Muñoz, J. Carracedo, Estegoanálisis aplicado a la generación automática de estegotextos en lengua española. CIBSI 2009. V Congreso Iberoamericano de Seguridad Informática. 16-18 Nov 2009. Montevideo-Uruguay.

[8] J. Greenberg, Some universals of grammar with particular reference to the order of meaningful elements. Cambridge, Mass: MIT Press. 1966

[9] A. Muñoz, I. Argüelles, J. Carracedo, Improving N-Gram linguistic steganography based on templates. International Conference on Security and Cryptography. Secrypt 2010. July 26-28 Athens, Greece.

[10] A. Muñoz, I. Argüelles, J. Carracedo, Hiding short secret messages based on linguistic steganography and manual annotation. Third IEEE International Symposium on Trust, Security and Privacy for Emerging Applications. TSP 2010. June 29-July 1, 2010 in Bradford, UK.

[11] A. Muñoz, I. Argüelles, J. Carracedo, Modificaciones sintácticas en lengua española con utilidad en esteganografía lingüística. RAEL-Revista Electrónica de Lingüística Aplicada. I.S.S.N.: 1885-9089 número 8. Fecha estimada publicación: abril 2010.

[12] A. Muñoz, I. Argüelles, J. Carracedo, Measuring the security of linguistic steganography in Spanish based on synonymous paraphrasing with WSD. Third IEEE International Symposium on Trust, Security and Privacy for Emerging Applications. TSP 2010. June 29-July 1, 2010 in Bradford, UK.

[13] M. Topkara, U. Topkara, M.J. Atallah, Information hiding through errors: a confusing approach, in: Proceedings of SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents, San Jose, CA, USA, January 29-February 1, 2007.

[14] J. Yago, Como sincronizar un comando terrorista con Twitter. 15/09/2008. http://www.securitybydefault.com/2008/09/como-sincronizar-un-comando-terrorista.html.

[15] R. Signel, Hackers Use Twitter to Control Botnet. August 13, 2009. http://www.wired.com/threatlevel/2009/08/botnet-tweets/

[16] A. Muñoz, I. Argüelles, Análisis de discurso en redes sociales. Twitter un caso bajo estudio. XXVIII Congreso Internacional de la Asociación Española de lingüística aplicada. AESLA 2010. Abril 2010. Vigo.

[17] H. Schmid, TreeTagger - a language independent part-of-speech tagger. Institute for Computational Linguistics of the University of Stuttgart. 2009. Available: http://www.ims.unistuttgart.de/projekte/corplex/TreeTagger/

[18] M. T. Chapman and G. I. Davida, Hiding the hidden: A software system for concealing ciphertext as innocuous text, in Information and Communications Security: First International Conference, Lecture Notes in Computer Science 1334, Springer, August 1997.