Criptoanálisis de un cifrador caótico realizado con redes neuronales celulares

AB. Orue[1], V. Fernandez[2], G. Pastor[2], M. Romera[2], G. Alvarez[2] y F. Montoya[2]

***Resumen*—Este artículo analiza la seguridad de un criptosistema caótico basado en una realización del circuito de Chua con redes neuronales de estado controlado. Se demuestra que el texto claro se puede recuperar por filtrado paso banda del texto cifrado o usando una decodificación imperfecta con parámetros del receptor de valores inexactos. Además, se evidencia que el espacio de claves puede reducirse notablemente y que la resolución de los parámetros que conducen a la recuperación de un texto claro inteligible es tan grosera como el 5%, posibilitando un ataque por fuerza bruta. Finalmente, se demuestra que los valores de los parámetros del sistema pueden determinarse con gran precisión analizando el error de decodificación producido por el desajuste entre los valores de los parámetros de receptor y transmisor.**

***Palabras clave*—Caos, criptoanálisis, criptografía, redes neuronales celulares, circuito de Chua.**

I. Introducción

La publicación en 1990 del articulo seminal de Pécora y Carroll [19], despertó nuevamente el interés por los criptosistemas analógicos, ahora con una aproximación completamente nueva. La idea consiste en que un criptosistema basado en caos en el dominio continuo no solo garantiza una comunicación segura sino que puede conseguirse sin sincronización externa y con gestión de claves simplificada. El principio de este sistema se asemeja al sistema de cifrado de Vernam [20], pero en este caso el ruido aditivo enmascarador proviene de un generador de señal analógica caótico. La señal portadora de información se encubre con pseudo ruido caótico de gran amplitud, de manera que la forma de onda obtenida posee una relación señal a ruido (S/N) muy pequeña, por lo tanto enmascarada para un interceptor. En el extremo receptor, esta señal se emplea para excitar un circuito esclavo del primero, sintonizado con el generador caótico, cuyos parámetros son desconocidos para el interceptor [15]. El circuito esclavo autosincronizante produce una réplica bastante aceptable del ruido caótico que se resta a continuación de la señal recibida, recuperando así la señal original.

La aplicación de las redes neuronales en este ámbito consiste en la reconstrucción dinámica de sistemas caóticos; es decir, la modelación dinámica de la serie temporal producida por un sistema físico, en este caso caótico [12]. El criptosistema desarrollado con redes neuronales consiste en dos redes neuronales, una en el transmisor y la otra en el receptor. La red del transmisor genera la serie temporal de un sistema caótico, por ejemplo el sistema de Lorenz o el sistema de Chua [4], [13]. Una vez generada la serie se añade el mensaje en claro a transmitir a una de las señales de la serie temporal generada, obteniéndose el mensaje cifrado. En recepción, utilizando la sincronización caótica, se sincroniza el receptor con el transmisor [5] y se resta la señal caótica reconstruida del mensaje cifrado. La Fig. 1 muestra un esquema general de este procedimiento.

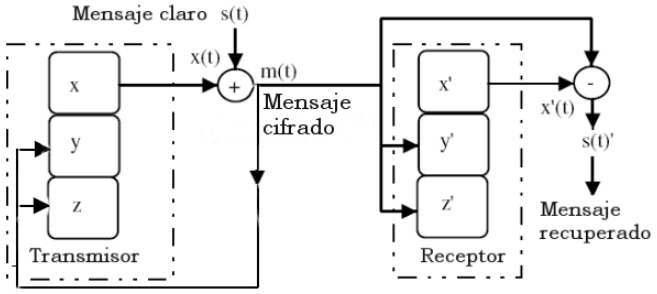

Figura 1. Sistema de enmascaramiento caótico realimentado.

II. Redes neuronales y criptografía caótica

En esta sección analizaremos un criptosistema sobre el enmascaramiento caótico basado en redes neuronales celulares (RNC), que se utiliza para generar la serie temporal del circuito caótico de Chua.

A. Modelo de las redes neuronales celulares

Las redes neuronales celulares (RNC) se introducen en 1988 por L. O. Chua y L. Yang [6]. Una RNC es una red neuronal compuesta por un un arreglo multidimensional de unidades de procesadores analógicos denominados *celdas* o células cuyos elementos interactuan directamente dentro de un entorno local o vecindad finita, de manera que capturan las propiedades geométricas, no lineales y el tipo de demora en la interacción de sus pesos [6].

Existen muchos modelos de redes neuronales celulares, la característica principal común a todos proviene de la idea original que define la celda como unidades de procesadores analógicos con señales continuas, donde las interacciones entre las celdas se realizan localmente y dentro de un radio finito.

La definición según Chua y Roska [9] de una RNC es:

Definición 2.1: La RNC es un arreglo de dimensión 2, 3 o n, de sistemas dinámicos idénticos, denominados celdas que satisfacen dos propiedades:

[1]Área de Cultura Científica del CSIC, Serrano 117, Madrid, España. Correo-e b.orue@orgc.csic.es.

[2]Instituto de Física Aplicada del CSIC, Serrano 144, Madrid, España.

- La mayoría de las interacciones son locales y están localizadas dentro de un radio finito r. Todas las variables de estado son continuas.
- La interacción entre cada celda con todas sus vecinas en términos de las variables de entrada, estado y salida se especifica en una plantilla, denominada plantilla de clonación.

B. Modelo generalizado de redes neuronales celulares

El modelo generalizado de redes neuronales celulares, se describe en [4] y está representado por las ecuaciones adimensionales siguientes,

$$\dot{x}_j = -x_j + a_j y_j + G_0 + G_s + i_j, \tag{1}$$

donde j indica el índice de la celda, a_j es un parámetro constante, i_j es un valor umbral; x_j es la variable de estado y y_j es la salida de la celda dada por,

$$y_j = \frac{1}{2}\Big[|x_j + 1| - |x_j - 1|\Big]. \tag{2}$$

En la Ec.(1) el término G_0 representa las combinaciones lineales de la salida y G_s las combinaciones lineales de las variables de estado, de las celdas conectadas que serán consideradas. Este modelo generalizado se ajusta perfectamente a la definición de Chua-Roska [9] y difiere del modelo Chua-Yang presentado en [6] en el parámetro G_s.

En [4] se propuso un modelo generalizado de las RNC denominado red neuronal de estado controlado (RCN-EC), que consiste en una RCN de tres capas, teniendo en cuenta la ecuación de estado Ec.(1), está representado por las ecuaciones siguientes,

$$\begin{aligned}
\dot{x}_1 &= -x_1 + a_1 y_1 + a_{12} y_2 + a_{13} y_3 + \sum_{k=1}^{3} s_{1k} x_k + i_1,\\
\dot{x}_2 &= -x_2 + a_{21} y_1 + a_2 y_2 + a_{23} y_3 + \sum_{k=1}^{3} s_{2k} x_k + i_2,\\
\dot{x}_3 &= -x_3 + a_{31} y_1 + a_{32} y_2 + a_3 y_3 + \sum_{k=1}^{3} s_{3k} x_k + i_3,
\end{aligned} \tag{3}$$

donde x_1, x_2 y x_3 son las variables de estado y y_1, y_2 y y_3 son las salidas correspondientes. Si se asume que,

$$\begin{aligned}
&a_{12} = a_{13} = a_2 = a_{23} = a_{32} = a_3 = a_{21} = a_{31} = 0,\\
&s_{13} = s_{31} = s_{22} = 0, \quad i_1 = i_2 = i_3 = 0.
\end{aligned} \tag{4}$$

La Ec. (3) puede escribirse como,

$$\begin{aligned}
\dot{x}_1 &= -x_1 + a_1 y_1 + s_{11} x_1 + s_{12} x_2,\\
\dot{x}_2 &= -x_2 + s_{21} x_1 + s_{23} x_3,\\
\dot{x}_3 &= -x_3 + s_{32} x_2 + s_{33} x_3,
\end{aligned} \tag{5}$$

C. El circuito de Chua

El circuito de Chua [10] es un circuito autónomo de tercer orden, es capaz de exhibir una rica variedad de comportamientos dinámicos, entre ellos bifurcación y caos [8] y por ello se considera generalmente como el paradigma del caos [16]. El circuito contiene 4 elementos lineales, dos condensadores una bobina y una resistencia, el elemento no lineal es una resistencia no lineal denominada *diodo de Chua*. La forma adimensional de las ecuaciones de este circuito es:

$$\begin{aligned}
\dot{x} &= \alpha\,[y - h(x)]\,,\\
\dot{y} &= x - y + z,\\
\dot{z} &= -\beta y - \gamma z,
\end{aligned} \tag{6}$$

siendo $h(x) = m_1 x + 0.5(m_1 - m_0)(|x+1| - |x-1|)$, donde x, y y z son las variables del sistema; $\dot{x}$, $\dot{y}$ y $\dot{z}$, son las derivadas de las variables con respecto al tiempo τ y α, β, γ, m_0 y m_1 son los parámetros del sistema.

En [4] se implementa el circuito de Chua utilizando redes neuronales celulares de estado controlado RNC-EC. Dicha implementación está compuesta por la interconexión de tres redes neuronales generalizadas. Las Ecs. (5) son las ecuaciones adimensionales que lo definen, siendo, $y_1 = 0.5(|x_1 + 1| - |x_1 - 1|)$.

Se puede observar que las ecuaciones del circuito de Chua, Ecs. (6), se obtienen fácilmente a partir de las Ecs. (5) con $x_1 = x$, $x_2 = y$, y $x_3 = z$, siempre que se cumplan las condiciones siguientes: $a_1 = \alpha(m_1 - m_0)$; $s_{11} = 1 - \alpha\, m_1$; $s_{12} = \alpha$; $s_{21} = s_{23} = 1$; $s_{32} = -\beta$; $s_{33} = 1 - \gamma$.

La ventaja de este modelo de RNC radica en que la materialización del circuito se lleva a cabo utilizando componentes electrónicos muy simples tales como resistencias, condensadores y amplificadores operacionales, a diferencia del circuito original de Chua que contiene un *diodo de Chua*, con una resistencia negativa no lineal.

Recientemente se ha propuesto un nuevo criptosistema caótico [14] que utiliza el circuito de Chua construido con las redes celulares RNC-EC. El criptosistema resultante es del tipo de enmascaramiento caótico realimentado. Esta estructura fue inicialmente propuesta por [18] con el objetivo de obtener una sincronización robusta entre el transmisor y el receptor de un sistema de comunicaciones que utilizaba el sistema caótico de Lorenz modificado. La simulación de este circuito utilizando PSpice [21] fue presentada en [14] y, más tarde, en [11] se realizó una versión en hardware. Las ecuaciones que definen el transmisor del criptosistema (también adimensionales) están dadas por:

$$\dot{x}_1 = -x_1 + s_{11} x_1 + s_{12} x_2 + a_1 y_1, \tag{7}$$

$$\dot{x}_2 = -x_2 + s_{21} m(\tau) + s_{23} x_3, \tag{8}$$

$$\dot{x}_3 = -x_3 + s_{32} x_2 + s_{33} x_3, \tag{9}$$

donde el texto cifrado es $m(\tau) = x_1(\tau) + s(\tau)$, y el texto claro es $s(\tau)$. Puede verse que el texto cifrado $m(\tau)$ es realimentado en la segunda ecuación del criptosistema.

Las ecuaciones que definen el receptor del criptosistema están dadas por:

$$\dot{x}'_1 = -x'_1 + s_{11} x'_1 + s_{12} x'_2 + a_1 y'_1, \tag{10}$$

$$\dot{x}'_2 = -x'_2 + s_{21} m(\tau) + s_{23} x'_3, \tag{11}$$

$$\dot{x}'_3 = -x'_3 + s_{32} x'_2 + s_{33} x'_3, \tag{12}$$

donde $y'_1 = 0.5(|x'_1 + 1| - |x'_1 - 1|)$.

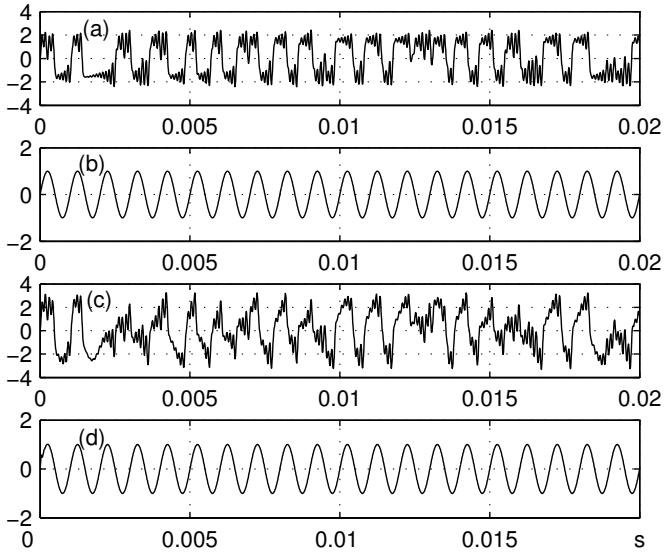

Figura 2. Forma de onda del ejemplo utilizado en [14] y [11]: (a) $x_1(t)$ variable del transmisor ; (b) texto claro $s(t) = \sin(2\,\pi\,1000\,t)$; (c) texto cifrado $m(t) = x_1(t) + s(t)$; (d) texto claro recobrado en el receptor $s'(t)$.

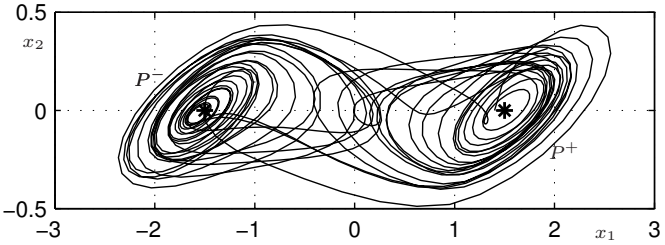

Figura 3. Proyección de la trayectoria del atractor de Chua sobre el plano (x_2, x_1).

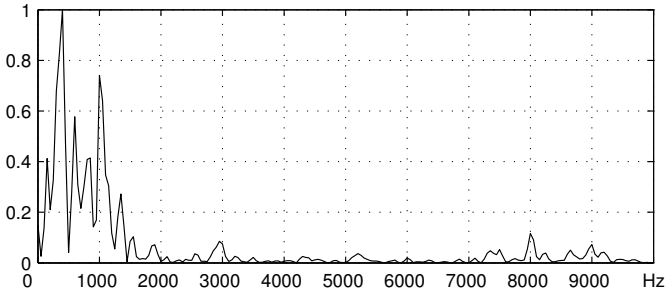

Figura 4. Espectro de potencia relativo de la variable del transmisor $x_1(t)$.

El texto claro recobrado $s'(\tau)$ en el extremo receptor se calcula mediante la ecuación: $s'(\tau) = m(\tau) - x'_1(\tau)$.

En [14] se utilizaron los siguientes valores: $\alpha = 9$, $\beta = 14 + \frac{2}{7}$, $\gamma = 0$, $m_0 = -\frac{1}{7}$, $m_1 = \frac{2}{7}$, $s_{21} = s_{23} = 1$, $s_{33} = 1 - \gamma = 1$. Obsérvese que estos parámetros corresponden a las ecuaciones adimensionales. En las implementaciones del circuito real, la respuesta temporal (y el espectro) del circuito se puede ajustar cambiando el valor del condensador en cada célula. De acuerdo con el esquema general del circuito del criptosistema de enmascaramiento caótico basado en RNC-EC mostrado en la Fig. 6 de [14], el factor de escala temporal es $t/\tau = R_{24}C_{21}$, donde t es el tiempo asociado con la construcción del circuito real.

En una de las simulaciones con PSpice que aparecen en [14], el valor de la resistencia es $R_{24} = 100\text{K}\Omega$ y el del condensador es $C_{21} = 51\text{nF}$, lo que equivale a que el factor de escala temporal sea $t/\tau = R_{24}C_{21} = 51 \times 10^{-6}$ (o $\tau/t = 10^6/51 \approx 19608$). Esta configuración fue utilizada también en la realización via hardware en [11] y fue la configuración que utilizamos en nuestro experimento. Nótese que los resultados del criptoanálisis descrito en este trabajo es aplicable a diferentes configuraciones de circuitos que respondan a las mismas ecuaciones.

En la Fig. 2 se ilustra la forma de onda de la señal x_1, el texto claro $s(t) = \sin(2\,\pi\,1000\,t)$, el texto cifrado y el texto recuperado en uno de nuestros experimentos.

La Fig. 3 muestra el atractor de doble ovillo de Chua que resulta de la proyección en el espacio de fase de una porción de la trayectoria que se extiende durante 0.2 s sobre el plano (x_2, x_1). La trayectoria del atractor de Chua dibuja dos ovillos tridimensionales situados en la vecindad de los puntos de equilibrio P^+ y P^-. Estas trayectorias tienen la forma de una espiral, la cual crece regularmente en amplitud y salta de un punto de equilibrio a otro a intervalos regulares y en forma aparentemente aleatoria. La trayectoria puede pasar arbitrariamente cerca de los puntos de equilibrio, pero nunca los alcanza, mientras esté en regimen caótico. Los dos asteriscos representan los puntos de equilibrio del atractor, cuyas coordenadas son $x_{1P^\pm} = \pm(1 - \frac{m_0}{m_1})$, $x_{2P^\pm} = 0$, $x_{3P^\pm} = \mp(1 - \frac{m_0}{m_1})$ [7].

En la Fig. 4 se muestra el espectro de potencia de la variable del transmisor $x_1(t)$ donde puede verse que la mayoría de la energía está localizada en una banda por debajo de 2 kHz. Esta energía corresponde a la mayor amplitud y a la oscilación más lenta de $x_1(t)$, asociada con el salto entre los dos ovillos. Existe un pico notable a la frecuencia del texto claro $f = 1000$ Hz, la cual se debe a la presencia de $m(\tau)$ en la Ec. (8). Pueden verse también algunos componentes de alta frecuencia cerca de los 8 kHz, como consecuencia del rizado de pequeña amplitud de $x_1(t)$, asociado a las vueltas alrededor de los puntos de equilibrio.

El trabajo continúa con la siguiente organización: en la sección Sec. III se analizan varias debilidades del sistema criptográfico. En las Secs. IV y V se muestran varias formas de romper este criptosistema, tales como filtrado y el ataque por fuerza bruta. Finalmente en la sección Sec. VI se presentan las conclusiones finales y algunas observaciones.

Es importante mencionar que la realización material con redes neuronales RNC-EC del circuito de Chua debe considerarse solamente como un ejemplo típico de su construcción, ya que el criptoanálisis mostrado en este trabajo es efectivo no solo en este tipo de realización del circuito de Chua, sino en cualquier realización del mismo basado en las Ecs. (6).

III. Problemas en la definición de Criptosistema

A pesar de que los autores de [14] y [11] basan la seguridad de su criptosistema en el comportamiento caótico asociado a la salida del circuito de Chua, no realizan ningún análisis de seguridad que así lo avale. No dan ningún detalle acerca de la selección de la clave ni de las frecuencias permitidas

para el texto claro, ni su amplitud. Tampoco mencionan las condiciones iniciales que fueron utilizadas.

A. Ausencia de especificación de la clave

El primer aspecto a considerar en un criptosistema es la clave secreta. Un criptosistema no puede existir sin una clave. Cuando se realiza el criptoanálisis de un sistema se asume que el criptoanalista conoce exactamente el diseño y funcionamiento del criptosistema bajo estudio, es decir, que conoce cada detalle acerca del algoritmo de cifrado pero no posee ninguna información sobre la clave secreta utilizada. Este es un requerimiento primordial en los sistemas de comunicaciones seguras de hoy día, usualmente conocido como principio de Kerckhoffs [3]. En [14] y [11] ninguno de los siguientes aspectos fueron considerados: la necesidad de una clave en el sistema propuesto, qué parámetros la componen, el espacio de claves disponible (es decir, cuántas claves distintas existen en el sistema), la precisión que debe tener y cómo puede ser creada y manipulada. Ninguno de estos elementos deben descuidarse cuando se diseña un sistema de comunicaciones seguro [2], [3].

La mayoría de los diseñadores de criptosistemas caóticos asumen que la clave del criptosistema está compuesta por los parámetros del sistema caótico [1]. Asumiremos esta premisa en este trabajo.

B. Condiciones iniciales peligrosas y Regiones de Trabajo Prohibidas

Es bien conocido que para los valores de los parámetros del ejemplo dado en [14] y [11], existen muchas órbitas periódicas inestables embebidas en el atractor de doble ovillo de Chua [16, Cuadro B1]. Si por alguna razón, durante la operación del sistema, se alcanzan determinados puntos, o éstos se encuentran incluidos en las condiciones iniciales, el sistema se convierte en inestable, mostrando un crecimiento continuo de la amplitud de las variables del mismo. Tales puntos deben considerarse prohibidos durante el funcionamiento normal del sistema.

El sistema propuesto en [14] y [11] difiere del circuito de Chua tradicional, en el que se conocían estos puntos prohibidos, en que este esquema está fuertemente realimentado. Por lo tanto, habría que preguntarse si estos puntos inestables han desaparecido o no.

Se estudió la estabilidad de varios de estos puntos conflictivos, encontrándose que algunos de ellos no permanecen inestables, aunque otros sí, como por ejemplo $\{x_1(0), x_2(0), x_3(0)\} = \{1.83487,\ 0,\ 2.53784\}$, y lo que es peor, se observó la existencia de una región prohibida de órbitas del atractor y/o de condiciones iniciales correspondientes a los valores $x_2 \geq 1.08$, para cualquier valor de x_1 y x_3.

Este problema no está relacionado con una amenaza a la seguridad del sistema, sino que degrada la fiabilidad del sistema de comunicaciones. Por lo tanto, se debe prestar una atención especial a su detección durante la operación del sistema a fin de aplicar las medidas correctoras apropiadas.

C. Texto claro peligroso

En [14] se afirma que la señal del texto claro de amplitud en el margen de 1 V a 2 V no perturba la sincronización caótica; ilustrándose un ejemplo de la señal de texto claro empleando dos señales con forma de onda sinusoidal y forma de onda triangular de frecuencia de 1 kHz y una amplitud de 1 V.

Como el texto claro se introduce en la Ec. (8) del transmisor, se provoca una perturbación del comportamiento normal del circuito de Chua. A mayor amplitud de la señal $s(t)$, más seria es la perturbación que se ocasiona en el comportamiento del circuito. Se encontró que todas las variables del atractor permanecieron sincronizadas con la frecuencia del texto claro cuando este tenía una amplitud de 1 V y frecuencias comprendidas dentro del margen de 4700 Hz a 4970 Hz. Esta es una situación muy peligrosa ya que el texto cifrado revela el texto claro.

Para frecuencias comprendidas entre 4970 Hz y 12500 Hz, se alcanza una órbita inestable alrededor de la frecuencia de 9500 Hz, lo que hace que el sistema no sea operable para textos claros de frecuencias desde los 4700 Hz a los 12500 Hz con una amplitud de 1 V. Se encontró además que para amplitudes de 2 V esta banda de frecuencia inutilizable se extiende desde 3200 Hz a 14300 Hz. Para asegurar que las órbitas del circuito permanezcan como parte del doble ovillo para cualquier frecuencia del texto claro, la amplitud debe permanecer menor que 0.24 V.

La señal de voz tiene un espectro cuya amplitud máxima se alcanza aproximadamente a los 1000 Hz, decayendo muy rápidamente con el incremento de la frecuencia, siendo despreciable la densidad de potencia para valores mayores de 3200 Hz. Por esta razón podemos concluir que el sistema que utilice los valores dados en el ejemplo de [14], es adecuado para el cifrado de la señal de voz, pero no para otras señales que tengan un espectro de gran amplitud a frecuencias mayores que 3200 Hz.

IV. RUPTURA DEL SISTEMA POR FILTRADO PASO BANDA

El problema principal asociado con el criptosistema bajo estudio consiste en que el mecanismo de sincronización entre el transmisor y el receptor es excesivamente robusto. Como consecuencia de esto, dado un conjunto de parámetros del transmisor, se puede alcanzar una sincronización casi perfecta para un gran número de combinaciones de parámetros inexactos del receptor.

Una condición necesaria para que cualquier criptosistema sea seguro es que los parámetros del sistema que sirven de clave, sean lo suficientemente sensibles para garantizar que el texto claro cifrado con una clave dada k_1 difiera dramáticamente del descifrado con una clave errónea k_2, de manera que desaparezca cualquier información acerca del texto claro [3], por pequeña que sea la diferencia entre las claves k_1 y k_2. En otras palabras, los coeficientes de la correlación cruzada normalizada entre el texto claro y el texto recobrado utilizando cualquiera de las posibles claves erróneas deberían tener valor cero o muy cercano a cero.

Desafortunadamente, el criptosistema propuesto no satisface estos requerimientos. Por el contrario, dado un texto cifrado,

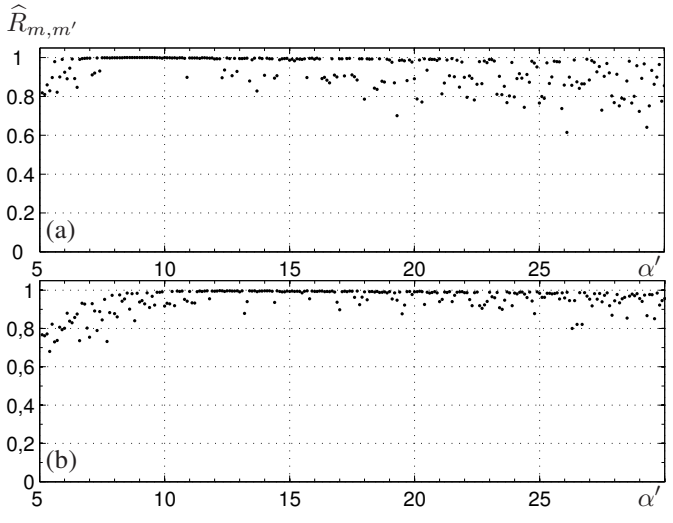

Figura 5. Coeficiente máximo de correlación cruzada $\widehat{R}_{m,m'}$ entre el texto claro m y el texto filtrado recobrado m' para varios conjuntos de claves erróneas de descifrado: (a) Parámetros del transmisor del ejemplo citado en [14] $\alpha = 9$, $\beta = 14 + \frac{2}{7}$, $m_0 = -\frac{1}{7}$, $m_1 = \frac{2}{7}$; (b) Parámetros arbitrarios del transmisor $\alpha = 12$, $\beta = 18$, $m_0 = -\frac{3}{25}$ y $m_1 = \frac{1}{4}$.

obtenido utilizando un conjunto específico de parámetros k_1, es posible encontrar una fórmula empírica que nos permita encontrar muchos conjuntos distintos de parámetros para el receptor, completamente diferentes de k_1, que posibiliten el descifrado y la recuperación casi exacta del texto claro, utilizando un filtro paso banda.

La Fig. 5 muestra el valor del coeficiente de correlación cruzada máximo entre el texto claro original y el texto claro recuperado, mediante filtrado paso banda, descifrado utilizando varios valores de $\alpha' \neq \alpha$. Se utilizó un filtro paso banda digital de respuesta impulsiva finita, de 200 etapas, cuya banda pasante iba desde 300 Hz a 3400 Hz, que es el ancho de banda típico de los enlaces telefónicos. El texto claro utilizado fue un tono puro de 1000 Hz. Los valores de los parámetros del receptor fueron tomados de acuerdo con la siguiente fórmula empírica: α' fue variada entre el valor 5 y el 30, mientras el resto de los parámetros se eligieron en función de α', tales que $\beta' = \alpha' + 5.3$, $m'_0 = \frac{\pi}{100} - \frac{\pi}{2\alpha'}$ y $m'_1 = \frac{\pi}{\alpha'} - \frac{\pi}{50}$. En la Fig. 5 se presentan dos casos, el caso (a) corresponde a los valores de los parámetros del transmisor del ejemplo citado en [14], es decir $\alpha = 9$, $\beta = 14 + \frac{2}{7}$, $m_0 = -\frac{1}{7}$ y $m_1 = \frac{2}{7}$. El caso (b) corresponde a otro posible conjunto de parámetros del transmisor, elegidos de forma totalmente arbitraria dentro del rango válido y sin relación con los del caso (b): $\alpha = 12$, $\beta = 18$, $m_0 = -\frac{3}{25}$ y $m_1 = \frac{1}{4}$.

Se puede observar que para la mayoría de los 250 ensayos del parámetro α', el coeficiente máximo de la correlación cruzada entre el texto claro y el texto filtrado recobrado con claves erróneas tiene un valor muy cercano a la unidad: es decir, el texto claro es recuperado sin ruido ni distorsión. Para las pruebas restantes el valor máximo del coeficiente de covarianza cruzada tiene un valor superior a 0.6, lo que significa que esta es una buena aproximación, aunque no perfecta, del texto claro. Por consiguiente, cualquier información que se pretenda ocultar con el criptosistema propuesto podrá estar seriamente comprometida.

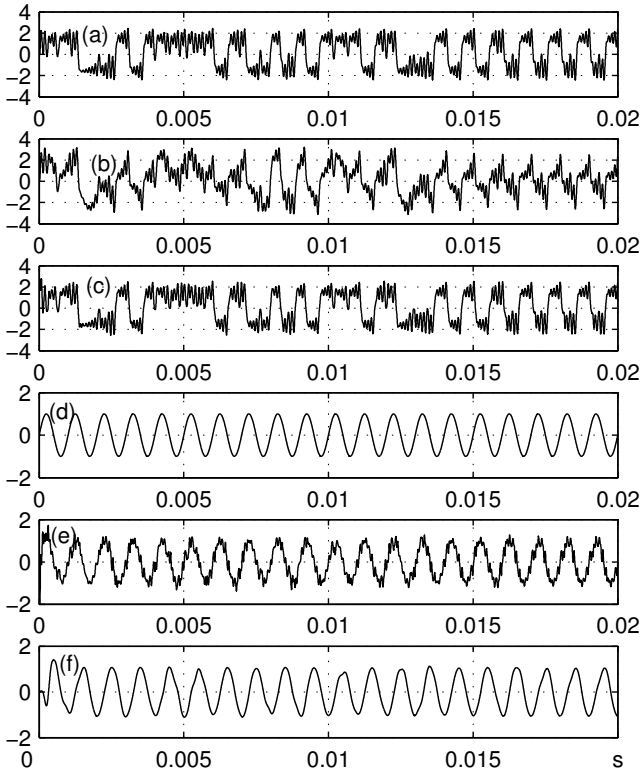

Figura 6. Texto claro recuperado con los parámetros erróneos: (a) $x_1(t)$ es la variable caótica del transmisor; (b) $m(t) = x_1(t) + s(t)$ Es el texto cifrado; (c) $x'_1(t)$ es la variable caótica del receptor; (d) El texto claro $s(t) = sin(2\,\pi\,1000\,t)$; (e) Texto recuperado $s'(t)$; (f) Texto claro recuperado mediante el filtro paso banda $s'(t)$

La figura 6 ilustra el problema paso a paso. Se presenta la señal del sistema correspondiente a los parámetros del transmisor: $\alpha = 9$, $\beta = 14 + 2/7 \approx 14.2857$, $m_0 = -1/7 \approx -0.1429$, $m_1 = 2/7 \approx 0.2857$; y a un conjunto arbitrario de valores de parámetros del receptor, elegidos suficientemente separados de los parámetros del transmisor : $\alpha' = 17$, $\beta' = 23.3$, $m'_1 = 0.1366$, $m'_0 = -m'_1/2 = -0.0683$. El texto claro usado era $s(t) = \sin(2\,\pi\,1000\,t)$.

Se puede apreciar que la forma de onda de la variable caótica del receptor $x'_1(t)$ se parece mucho a la variable caótica del transmisor $x_1(t)$. Por lo tanto el texto recuperado $s'(t)$ difiere fundamentalmente del texto claro original $s(t)$ en los componentes de alta frecuencia; es decir, los saltos entre los puntos de equilibrio son casi idénticos, pero no así la velocidad y amplitud de las vueltas alrededor de ellos, lo que ocasiona un ruido de alta frecuencia en el texto claro recuperado. Este ruido puede eliminarse fácilmente mediante filtrado baso banda. La figura 6(f) muestra el texto claro recuperado después del filtrado, utilizando el mismo filtro digital paso banda utilizado en los experimentos mostrados en la figura 5.

V. Ruptura del sistema mediante un ataque por fuerza bruta

Una de las formas posibles de romper el sistema es empleando un *ataque por fuerza bruta*, que consiste en probar todos

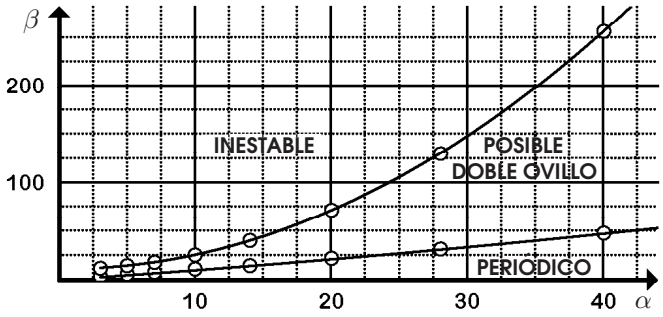

Figura 7. Región del plano (α, β) que origina el atractor de doble ovillo.

los valores posibles de sus parámetros hasta obtener un texto claro significativo y libre de ruido [3, Sec. 4.4].

Un ataque por fuerza bruta solo es factible en el caso de que el espacio de claves sea pequeño. De este modo, para evitarlo, el número de claves posibles debe ser tan grande como sea posible. Actualmente el veterano Data Encryption Standard se considera obsoleto y ha sido abandonado ya que este tiene *solamente* $2^{56} = 7.2 \times 10^{16}$ claves diferentes. Efectivamente, de acuerdo a la potencia de los ordenadores de hoy día, el tamaño del espacio de claves recomendado es como mínimo de 2^{100} para resistir este tipo de ataque.

A. Reducción del hipotético espacio de claves del circuito de Chua

El problema asociado la utilización del circuito de Chua como criptosistema es que el rango útil de los valores de los parámetros es muy reducido. En [17] y [16] se demuestra que el circuito de Chua exhibe casi todas las bifurcaciones y comportamientos caóticos descritos en la literatura. Su variedad es muy compleja y ésta es la razón por la que se le conoce como el paradigma del caos. Diferentes combinaciones de los parámetros α y β conducen a diferentes proyecciones sobre el plano (x_2, x_1) de la trayectoria. Entre ellas: el atractor extraño de doble ovillo, sumideros, órbitas periódicas asimétricas, órbitas n-periódicas, órbitas heteroclínicas similares a la espiral de Rössler, órbitas homoclínicas y focos repulsivos. El único comportamiento caótico apropiado para enmascarar el texto claro es el atractor extraño de doble ovillo, ya que los demás comportamientos dan lugar a formas de onda muy simples que no pueden ocultar el texto claro de forma eficiente. Desafortunadamente, la región del plano (α, β) que origina el atractor extraño de doble ovillo es una pequeña fracción, del orden de 4% de todas las posibles combinaciones de los valores de los parámetros, como se muestra en [17] y [16]. Por lo tanto, el hipotético espacio de claves basado en los valores de los parámetros del sistema será mucho menor que el inicialmente esperado.

Esta situación se agrava por el hecho de que algunos parámetros del circuito de Chua tienen una relación directa con las coordenadas x_1 de los puntos de equilibrio del atractor P^+ and P^-, que pueden ser delimitados aproximadamente, simplemente observando la forma de onda del texto cifrado. Esto favorece la reducción del espacio de claves, como se describe más adelante.

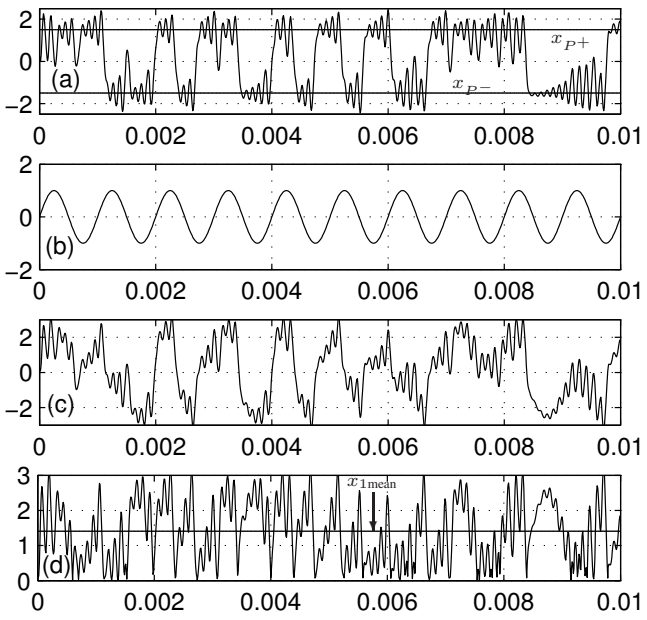

Figura 8. Estimación de los puntos de equilibrio: (a) Variable caótica del transmisor $x_1(t)$ con x_{1P^+} y x_{1P^-}; (b) Texto claro $s(t)$; (c) Texto cifrado $m(t) = x_1(t) + s(t)$; (d) Valor absoluto del texto cifrado $|m(t)|$.

Como el sistema descrito en [14] y [11] difiere del circuito ordinario de Chua, ya que utiliza una realimentación, sería posible que tuviese un comportamiento distinto que el circuito original. Por esta razón, se investigó experimentalmente la región del plano (α, β) que origina el atractor de doble ovillo, para diferentes combinaciones de los parámetros restantes m_0 y m_1. El resultado se presenta en la Fig. 7. Dependiendo de los valores de los parámetros de m_0 y m_1 los puntos que se encuentran dentro de esta región pueden o no originar el atractor de doble ovillo. Sin embargo, los puntos que quedan fuera de esta región nunca originarán el atractor para ninguna combinación de valores de m_0 and m_1. Por consiguiente, éstos no son adecuados para ocultar información y no necesitan ser investigados cuando se realiza un ataque por fuerza bruta.

La región que debe ser examinada por fuerza bruta queda delimitada aproximadamente por las curvas $\beta = 0.0062\,\alpha^2 + 0.92\,\alpha + 0.5$ y $\beta = 0.157\,\alpha^2 - 0.16\,\alpha + 12$. De este modo, se comprueba que el espacio de claves utilizables queda notablemente reducido.

Correspondientemente, la región del plano (m_0, m_1) que origina el doble atractor se puede delimitar también a partir de la definición del circuito de Chua y del texto cifrado de la manera siguiente.

De acuerdo a la definición del circuito de Chua, los parámetros m_0 y m_1 son: $m_0 = (G_a/G) + 1$ y $m_1 = (G_b/G) + 1$, donde G es la conductancia positiva, mientras G_a y G_b son las dos conductancias negativas del circuito equivalente de la resistencia no lineal del circuito de Chua. Estas satisfacen la relación $G_a < G_b < 0$, de manera que $1 > m_1 > m_0$. Si se determinan las coordenadas de los puntos de equilibrio del atractor $x_{1P^\pm} = \pm(1 - \frac{m_0}{m_1})$ se puede establecer una relación muy cercana entre m_0 y m_1.

Si la variable caótica del transmisor $x_1(t)$ fuera accesible

sin perturbaciones, las coordenadas $x_{1P^{\pm}} = \pm(1 - \frac{m_0}{m_1})$ de los puntos de equilibrio $P^{\pm}$ se podrían determinar a partir de la forma de onda de la variable. La figura 8 (a) muestra la forma de onda de $x_1(t)$ y los valores reales de x_{1P^+} y x_{1P^-}. Como puede, verse no es difícil aproximar los valores de x_{1P^+} ó x_{1P^-} como la línea equidistante entre los máximos relativos y los mínimos relativos de la parte positiva o negativa respectivamente de la forma de onda $x_1(t)$.

Sin embargo, como el único dato accesible al criptoanalista oponente es el texto cifrado $m(t) = x_1(t) + s(t)$, mostrado en la Fig. 8 (c), la variable del transmisor $x_1(t)$ permanece enmascarada por la presencia del texto claro. Por consiguiente, sólo se puede lograr una estimación a groso modo de $x_{1P^{\pm}}$. No obstante, este valor se podrá delimitar eficazmente estableciendo dos límites fácilmente medibles. Debido a que $x_{1P^+} = -x_{1P^-}$, es preferible trabajar con el valor absoluto de $m(t)$, representado en la Fig. 8 (d).

El valor de $|x_{1P^{\pm}}|$ se puede delimitar entre los límites $x_{1\max}$ y $x_{1\text{mean}}$, siendo el primero el valor máximo de $|m(t)|$ y el siguiente, la media de $|m(t)|$. En la Fig. 8 (d) se evidencia que $|x_{1P^{\pm}}| < x_{1\max}(t)$ y se encontró experimentalmente –para una gran variedad de valores de los parámetros y de textos claros– que en todos los casos se cumple que $|x_{1P^{\pm}}| > x_{1\text{mean}}(t)$. El valor real de $\frac{m_0}{m_1}$ es -0.5. Por lo tanto $|x_{1P^{\pm}}| = 1 - \frac{m_0}{m_1} = 1.5$, lo que concuerda con los límites encontrados experimentalmente: $x_{1\max} = 3.00$ y $x_{1\text{mean}} = 1.41$.

Esta medida posibilita una nueva reducción importante del rango de búsqueda de los valores de m_0 y m_1. Como $x_{1\max} = 3.00 > \pm(1 - \frac{m_0}{m_1}) > x_{1\text{mean}} = 1.41$ y $1 > m_1 > m_0$, se desprende que:

$$1 > m_1 > 0 \quad y \quad -0.41m_1 > m_0 > -2m_1,$$

lo que supone una reducción adicional del espacio de claves.

B. Falta de precisión de la clave

El establecimiento de los requerimientos de precisión de la clave del criptosistema propuesto es un punto crítico para resistir un ataque por fuerza bruta. En un criptosistema perfecto el mensaje cifrado usando una clave específica no debe ser vulnerable frente a un intento de descifrado con una clave distinta, aún si ambas claves difieren sólo en la mínima cantidad permitida por la precisión de la máquina [3, Regla 9].

El problema de este sistema consiste en que la precisión requerida por la clave es muy baja y, consecuentemente, el número de claves efectivamente diferentes es muy pequeño.

La figura 9 ilustra este problema. En ella se muestra el texto claro recuperado para tres conjuntos de valores de los parámetros del receptor α', β', m_0', m_1', que son diferentes de los valores de los parámetros del transmisor α, β, m_0, m_1. Como se observa, el texto claro se recupera casi perfectamente para un error del 1% en la estimación de la magnitud de cada parámetro del receptor. El error inicial se debe al transitorio ocasionado por las diferentes condiciones iniciales entre el transmisor y el receptor. Además, puede verse que un error tan grande como el 5% proporciona la recuperación de un texto claro reconocible.

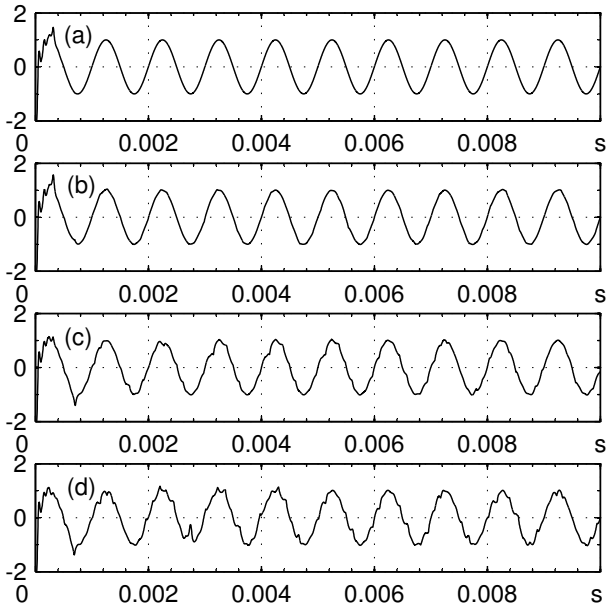

Figura 9. Texto claro recuperado con una estimación ligeramente errónea de los parámetros: (a) texto claro recuperado con $\{\alpha',\ \beta',\ m_0',\ m_1'\} = \{\alpha,\ \beta,\ m_0,\ m_1\}$; (b) texto claro recuperado con $\{\alpha',\ \beta',\ m_0',\ m_1'\} = 1.01 \times \{\alpha,\ \beta,\ m_0,\ m_1\}$; (c) texto claro recuperado con $\{\alpha',\ \beta',\ m_0',\ m_1'\} = 0.97 \times \{\alpha,\ \beta,\ m_0,\ m_1\}$; (d) texto claro recuperado con $\{\alpha',\ \beta',\ m_0',\ m_1'\} = 1.05 \times \{\alpha,\ \beta,\ m_0,\ m_1\}$.

La mejor estrategia para un ataque de fuerza bruta consiste en probar todas las claves posibles utilizando una resolución de los parámetros tan grosera como el $\pm 5\%$, comenzando por los valores más probables y subsecuentemente expandiendo el área de búsqueda si no se alcanza un resultado satisfactorio. Una vez que se ha encontrado el mejor conjunto de valores de los parámetros, se refina la precisión de dichos parámetros hasta encontrar aquellos que proporcionen el texto claro recuperado más limpio. Con una resolución reducida de $\pm 5\%$ el número de pruebas estará limitada a 24 valores de los parámetros para cubrir una variación de una década.

Inicialmente el valor del parámetro α puede ser explorado en el margen entre 4 y 40, mientras el valor de m_1 puede ser investigado entre 0.05 y 0.5. Utilizando los límites establecidos en la Sec. V-A para β y m_0, el número total de ensayos será aproximadamente de $390,000 \approx 2^{18.6}$, lo que constituye un número modesto de claves. En caso de fallo, el espacio de búsqueda deberá ensancharse progresivamente.

C. Determinación de los parámetros

Como se ilustra en la Fig. 4, la variable del transmisor $x_1(t)$, que actúa como ruido para enmascarar al texto claro, tiene dos bandas de frecuencias bien diferenciadas. La banda de menor frecuencia tiene componentes espectrales generalmente por debajo de 2 kHz, lo que corresponde a los saltos en el atractor entre los dos ovillos centrados en los puntos de equilibrio P^+ y P^-. Esta parte oculta de manera efectiva el texto claro caracterizado con la misma banda de frecuencia.

La segunda parte constituida por la banda de frecuencias más altas, localizada cerca de los 8 kHz, está asociada con la trayectoria de los ovillos del atractor alrededor de los puntos de equilibrio.

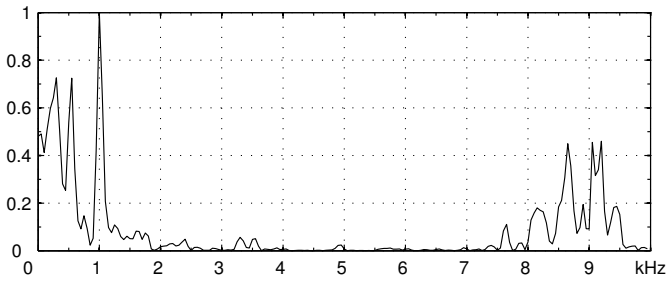

Figura 10. Espectro de potencia relativo del error de decodificación del receptor ε, con parámetros del transmisor: $\alpha = 9$, $\beta = 14+2/7 \approx 14.2857$, $m_0 = -1/7 \approx -0.1428$, $m_1 = 2/7 \approx 0.2857$; y parámetros del receptor elegidos arbitrariamente: $\alpha' = 4.5$, $\beta' = 9$, $m'_0 = -0.12$, $m'_1 = 0.21$.

Cuando el texto cifrado se descifra con un receptor no autorizado con una estimación errónea de los parámetros, puede verse que el texto claro recuperado $s'(t) = m(t) - x'_1(t) = s(t)+x_1(t)-x'_1(t)$ está compuesto por el texto claro y por el error de decodificación $\varepsilon = x_1(t)-x'_1(t)$, el cuál puede considerarse como un ruido de enmascaramiento indeseado. Si los parámetros del transmisor y receptor fueran iguales, el error de decodificación desaparecería. Consecuentemente, una estrategia para recuperar el texto claro consistirá en determinar los parámetros del receptor que minimicen el error de decodificación. Sin embargo, como el ruido y el texto claro comparten la banda de frecuencias más bajas del espectro, su separación completa no es posible. No obstante, todavía es posible aislar la banda de frecuencias más altas del error de decodificación, no contaminada por el texto claro. Por tanto, el texto claro deberá tener un espectro de frecuencia limitado a las frecuencias bajas y debe estar suficientemente separado de la banda de ruido de alta frecuencia.

La Fig. 10 ilustra el espectro de potencia relativo del error de decodificación del receptor. Los componentes de frecuencias más bajas están mezclados con el texto claro mientras que los componentes de alta frecuencia están lejos de la frecuencia del texto claro. Por tanto, el error de decodificación creado por las frecuencias altas de ε se puede extraer fácilmente del texto inexactamente descifrado mediante un filtro paso alto con una frecuencia de corte de 6.5 kHz, para eliminar las frecuencias del texto claro y retener los componentes de alta frecuencia del ruido.

La Fig. 11 ilustra la potencia logarítmica de los componentes de alta frecuencia del error de decodificación ε para diferentes conjuntos de parámetros del receptor α', m'_1 y m'_0 como una función de β'. Puede observarse que el error mínimo de decodificación se alcanza cuando los parámetros del transmisor y del receptor concuerdan. Todas las curvas muestran la misma tendencia: el error de decodificación crece con el desajuste entre los parámetros del transmisor y receptor. Por otra parte, su mínimo relativo se alcanza para valores cercanos a los del transmisor.

Se ha desarrollado un procedimiento de optimización iterativo que consiste en un número de vueltas de aproximación a fin de determinar los valores de los parámetros que dan lugar a un error de decodificación mínimo. Se variaron uno a uno los parámetros en cada vuelta, empezando por un conjunto de valores arbitrarios $\alpha' = 5$, $\beta' = 7$, $m'_1 = 0.1$, $m'_0 = 0.2$. En total se probaron 31 valores de cada parámetro dentro del margen definido en la Sec. V-A, reteniendo el valor que produjo el menor error de decodificación. El margen de variación de cada parámetro se redujo progresivamente en cada vuelta. El proceso se terminó cuando se alcanzó un valor fijo de cada parámetro, lo que sucedió al cabo de 30 vueltas y el tiempo de cálculo fue de 965 segundos, en un PC con una CPU Pentium Dual. La Fig 12 ilustra la secuencia del proceso, presentando la variación de cada parámetro en función del número de vuelta. Se determinaron los valores de los parámetros con una precisión de cinco a seis dígitos, lo que permitió la recuperación exacta del texto claro.

VI. Conclusiones

Se ha realizado el estudio del criptosistema descrito en [14] y [11], y se ha demostrado que el mismo es inseguro.

El mecanismo de sincronización utilizado es excesivamente robusto. Como consecuencia de esto, se puede alcanzar la sincronización casi exacta utilizando un número prácticamente infinito de combinaciones de parámetros del receptor, diferentes a los del transmisor.

Por tanto, se puede recuperar el texto claro filtrando paso bajo el texto decodificado obtenido con un receptor con parámetros erróneos, o incluso filtrando directamente el texto cifrado.

También se encontró que el espacio de claves del sistema se puede reducir notablemente estudiando las propiedades geométricas y las regiones caóticas del atractor de Chua. La precisión de los parámetros necesaria para recuperar un texto inteligible es tan basta como el $\pm 5\%$; por tanto, resulta económico realizar un ataque por fuerza bruta.

Finalmente, se determinaron los parámetros del sistema con gran precisión analizando y minimizando el error de decodificación originado por el desajuste entre los parámetros del transmisor y el receptor.

Agradecimientos

Los autores agradecen su ayuda al Ministerio de Educación y Ciencia proyecto, TSI2007-62657 y al CDTI (Ministerio

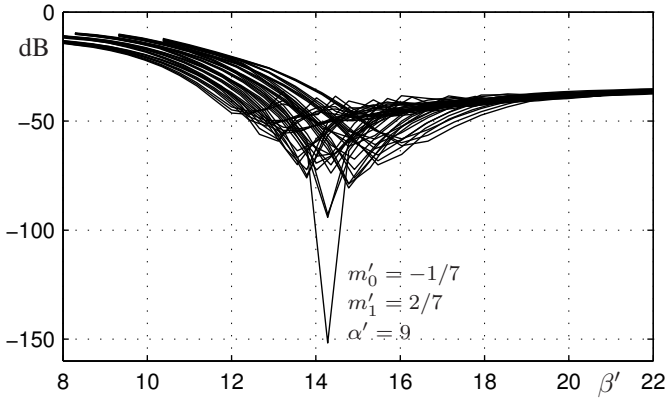

Figura 11. Potencia de los componentes de alta frecuencia del error de decodificación ε, para diferentes conjuntos de valores de los parámetros del receptor: $\alpha' = \{4, \ldots, 20\}$; $m'_1 = \{0.01, \ldots, 0.9\}$; $m'_0 = \{0.01, \ldots, 1.8\}$.

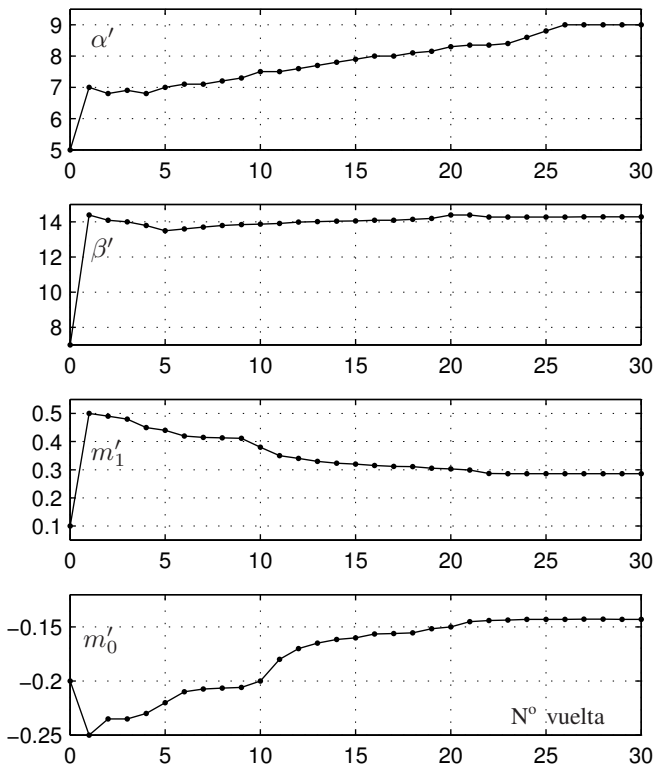

Figura 12. Historia de la aproximación seguida por los parámetros del sistema receptor hacia los valores exactos.

de Industria, Turismo y Comercio), en colaboración con Telefónica I+D, proyecto SEGUR@ (CENIT 2007-2010).

REFERENCIAS

[1] G. Alvarez, S. Li, F. Montoya, M. Romera, and G. Pastor, "Breaking projective chaos synchronization secure communication using filtering and generalized synchronization," *Chaos Soliton. Fract.*, vol. 24, no. 3, pp. 775–783, May 2005.

[2] G. Alvarez, F. Montoya, M. Romera, and G. Pastor, "Breaking a secure communication scheme based on the phase synchronization of chaotic systems," *Chaos*, vol. 14, no. 2, pp. 274–278, June 2004.

[3] G. Alvarez and S. Li, "Some basic cryptographic requirements for chaos-based cryptosystems," *International Journal of Bifurcation and Chaos*, vol. 16, no. 8, pp. 2129–2151, August 2006.

[4] P. Arena, S. Baglio, L. Fortuna, and G. Manganaro, "Chua's circuit can be generated by CNN cells," *IEEE T. CAS-I: Fundamentals theory and applications*, vol. 42, no. 2, 1995.

[5] E. Castillo and J. Gutiérrez, "Nonlinear time series modeling and prediction using functional networks. extracting information masked by chaos," *Physics Letters A*, vol. 244, pp. 71 – 84, 1998.

[6] Leon O. Chua and L. Yang. Cellular neural networks: Theory. *IEEE Transactions on Circuits and Systems*, 35, 1988.

[7] L. O. Chua, "A zoo of strange attractors from the canonical chua's circuit," *J. Circuit, Systems, and Computers*, vol. 2, 1993.

[8] L. O. Chua, C. W. Wu, A. Huang, and G. Zhong, "A universal circuit for studying and generating chaos. part I: Routes to chaos," *IEEE Transactions on Circuits and Systems: Theory and applications*, vol. 40, no. 10, pp. 147–156, 1993.

[9] Leon O. Chua and T. Roska. The CNN paradigm. *IEEE Transactions on Circuits and Systems:I*, 40, 93.

[10] L. Chua, "A zoo of strange attractors from the canonical Chua's circuits," *Proceedings of the 35th Midwest Symposium on Circuits and Systems (Cat. No.92CH3099-9) IEEE*, vol. 2, pp. 916–926, 1992.

[11] E. Günay and M. Alçi, "Experimental confirmation of SC-CNN based chaotic masking systema with feedback," *Int. J. Bifurcat. Chaos*, vol. 15, no. 12, pp. 4013–4018, 2005.

[12] S. Haykin, *Neural Networks. A comprehensive Foundation*, 2nd ed., P. Hall, Ed., 1999.

[13] R. Kilic, M. Alci, and E. Gunay, "A SC-CNN based chaotic masking system with feedback," *International Journal of Bifurcation and Chaos*, vol. 14, no. 1, pp. 245–256, 2004.

[14] R. Kiliç, M. Alçi, and E. Günay, "A SC-CNN based chaotic masking system with feedback," *Int. J. Bifurcat. Chaos*, vol. 14, no. 1, pp. 245–256, 2004.

[15] L. Kocarev, K. S. Halle, K. Eckert, U. Parlitz, and L. O. Chua, "Experimental demonstration of secure communications via chaotic synchronization," *Int. J. Bifurcat. Chaos*, vol. 2, no. 3, pp. 709–713, 1992.

[16] R. N. Madan and C. W. Wu, *Chua's Circuit: A Paradigm for Chaos*, ser. B, Nonlinear science. World Scientific, 1993, vol. 1, ch. 2, pp. 59–89.

[17] T. Matsumoto, "Chaos in electronics circuits," *Proceedings of the IEEE*, vol. 75, no. 8, pp. 1033–1057, 1987.

[18] V. Milanović and M. E. Zaghloul, "Improved masking algorithm for chaotic communications systems," *Electron. Lett.*, vol. 32, no. 1, pp. 11–12, Jan 1996.

[19] L. M. Pecora and T. L. Carroll, "Synchronization in chaotic systems," *Phys. Rev. Lett.*, vol. 64, pp. 821–824, February 1990.

[20] D. Stinson, *Cryptography: theory and practice*. CRC Press, Boca Raton, 1995.

[21] http://www.cadence.com/orcad/index.html