Teoría del Caos en la Protección de Información

M. en C. María Aurora Molina Vilchis ♦
M. en C. Eduardo Vega Alvarado
Dr. Ramón Silva Ortigoza ♣
Profesores del CIDETEC_IPN

En las últimas décadas ha surgido como una nueva corriente para la protección de información la Teoría del Caos. Partiendo del principio de Kerckhoff, en donde se establece que la robustez de un sistema de seguridad criptográfica reside en la dificultad para obtener la clave, entonces los métodos de generación de claves deben poseer ciertas propiedades de aleatoriedad. Para propósitos prácticos se han experimentado diversos métodos que proporcionan secuencias pseudo-aleatorias apropiadas; entre éstos se ubican los mapeos[1] caóticos. En este artículo se presenta el estado del arte de la Teoría del Caos en la generación de dichas secuencias, obtenidas a partir de los mapeos caóticos y sus aplicaciones en los diferentes tipos de protección de información multimedia y en las comunicaciones, especialmente las de espectro disperso.

Introducción

Con el origen de la Teoría del Caos, hace apenas unas tres décadas y con el descubrimiento de la sincronización del caos, se han formulado una gran cantidad de cifradores que de una u otra manera, y en general con poco rigor criptoanalítico, explotan las buenas propiedades de confusión y difusión de las aplicaciones caóticas [1]. Por ello, las señales resultantes de sistemas caóticos son impredecibles en la práctica a mediano y largo plazo, debido a la sensibilidad de las condiciones iniciales y a sus propiedades estadísticas aleatorias a pesar de ser generadas por algoritmos deterministas.

En los sistemas de seguridad de información resulta útil contar con secuencias binarias que puedan aplicarse a procesos que otorguen confidencialidad, privacidad o autenticidad. Para efectos prácticos, dichas secuencias deben cumplir con dos requisitos fundamentales: aleatoriedad e impredictibilidad. Entre las aplicaciones de estas secuencias resaltan las criptográficas, esteganográficas y canales subliminales [1]; su uso ha tenido buena aceptación por las propiedades estadísticas que poseen y su comportamiento *caótico*[2]. Las secuencias pseudo-aleatorias o aproximadamente aleatorias pueden generarse mediante la aplicación recursiva de *mapeos caóticos*.

En este artículo se presenta el estado del arte de las aplicaciones que tiene la Teoría del Caos en la protección de información, específicamente en criptografía, esteganografía y canales subliminales. Este documento consta de cinco secciones. La sección 2 trata de los conceptos fundamentales de la Teoría del Caos. En la sección 3 se detallan los trabajos más representativos para la generación de secuencias pseudo-aleatorias utilizando los principios de la Teoría del Caos. En la sección 4 se tratan los trabajos más relevantes de las aplicaciones de mapeos caóticos en la protección de información. Finalmente, en la sección 5 se presentan las conclusiones.

Teoría del Caos

El primer trabajo relacionado con los sistemas dinámicos fue presentado por el matemático francés Poincaré [3], quien con el estudio del movimiento planetario inventó nuevos métodos topológicos para estudiar soluciones para ciertas ecuaciones diferenciales; en lugar de usar los métodos basados en series, desarrolló el concepto de funciones automórficas, para resolver ecuaciones diferenciales lineales de segundo orden con coeficientes algebraicos, y de esta manera descubrió que las ecuaciones diferenciales admitían soluciones de una complejidad inimaginable hasta ese momento. Con este trabajo se reveló por primera vez que las ecuaciones diferenciales que describen fenómenos naturales

1 Dados dos conjuntos X y Y, el término mapeo significa que, para cada $x \in X$ le corresponde uno y solo un elemento $y \in Y$.

2 El caos en los sistemas se define como "el comportamiento inestable y no periódico de las soluciones de los sistemas dinámicos no lineales y deterministas".

pueden tener soluciones que se comportan en forma "caótica". Sin embargo, éste trabajo tuvo poco interés; a pesar de ello podemos decir que representa el origen de una nueva era para el estudio de las ecuaciones diferenciales.

En 1963 Lorenz [4] publicó un modelo matemático para describir la forma en la que el aire se mueve en la atmósfera terrestre. Observó, que si ocurrían alteraciones mínimas en los valores de las variables iniciales de su modelo primitivo, resultaban patrones climatológicos ampliamente divergentes. Esta sensible dependencia de las condiciones iniciales fue conocida después como el *Efecto Mariposa*[3]; aún cuando las fluctuaciones en las condiciones iniciales o intermedias de un proceso sean muy pequeñas, casi imperceptibles, pueden ser lo suficientemente poderosas como para provocar grandes alteraciones de los estados finales. Con este hecho se inicia el estudio moderno de la dinámica caótica. También dedujo algunas de las propiedades generales del caos como son: la dependencia y sensibilidad a las condiciones iniciales, la existencia de órbitas periódicas densas, los puntos periódicos, el acotamiento de trayectorias y la no linealidad, entre otras. Sin embargo, para los mapeos denominados caóticos, se consideran tres propiedades importantes:

a) Dependencia sensitiva.
b) Transitividad (órbita densa).
c) Conjunto denso de puntos periódicos.

Las propiedades ***a*** y ***b*** corresponden a un comportamiento irregular e impredecible, de aquí que concuerdan con las ideas intuitivas del caos; la propiedad ***c***, sin embargo, parece discrepar de las dos propiedades anteriores, al interpretarse como un comportamiento periódico, regular y predecible. No obstante, estos puntos periódicos son inestables, lo que significa que tan pronto como una órbita se acerca a un punto periódico, esta se empujará lejos en alguna otra parte. Por esta razón, en términos prácticos, el conjunto denso de órbitas periódicas no implica un comportamiento ordenado. Basándose en estas tres propiedades Devaney [2], planteó en 1989 la definición de Caos como sigue:

Sea V un intervalo. Decimos que f: V → V es caótica en V si:

a) *f tiene dependencia sensitiva en las condiciones iniciales,*
b) *f es transitiva y*
c) *los puntos periódicos son densos en V.*

Las tres propiedades se demuestran cuando se presentan una gran cantidad de iteraciones, dando origen al siguiente Teorema:

Si se tiene una función f, tal que:

f: [0,1] → [0,1], para muchas iteraciones *f* es caótica en [0,1].

Secuencias pseudo-aleatorias

En criptografía resulta muy útil contar con secuencias binarias que permitan aplicar procesos de transformación que otorguen confidencialidad, privacidad o autenticidad a la información que se transmite o se almacena. En este sentido, un trabajo muy referido es el realizado por Rueppel [4][5], quien en 1986 mostró la manera de analizar y diseñar secuencias para aplicaciones de cifrado en flujo. En su trabajo, utiliza el cifrador de Vernam como único cifrador incondicionalmente seguro; para que esto sea verdadero, es necesario que la secuencia de la clave sea finita y aleatoria, siendo necesario un algoritmo determinístico para su generación. Las secuencias generadas por este tipo de algoritmos se han catalogado como pseudo aleatorias, las que para usos criptográficos deben guardar ciertas propiedades que las hagan parecer verdaderamente aleatorias. Estas propiedades están contenidas en los llamados postulados de aleatoriedad que formuló Golomb [6] en 1967, los cuales se asocian con la aleatoriedad de secuencias finitas. No obstante, el cumplimiento de estos postulados no implica que una secuencia binaria necesariamente sea impredecible (condición asociada también a las secuencias verdaderamente aleatorias), pero al menos se asegura un comportamiento aleatorio adecuado.

Se define a las secuencias pseudo-aleatorias o ***PN*** (***Pseudo Noise***) como un conjunto de señales binarias periódicas de cierta longitud; para cada periodo la señal puede considerarse que se aproxima a una señal aleatoria (no siéndolo). Existen muchas y muy variadas aplicaciones de las secuencias en campos finitos cuyos términos dependen de forma sencilla de sus predecesores. Una ventaja computacional de estas secuencias reside en su facilidad para ser generadas mediante procedimientos recursivos.

En aplicaciones de comunicaciones un trabajo representativo es [7], donde se describen sistemas en los cuales la energía de la señal es dispersada sobre un gran ancho de banda; de esta manera se consiguen propiedades particulares. Esto proporciona inmunidad en el canal a interferencias intencionales y sin intención. Además de un mayor rechazo al ruido, existe la propiedad

3 Es un concepto que hace referencia a la noción de sensibilidad a las condiciones iniciales de la Teoría del Caos. Su nombre proviene de un antiguo proverbio chino: "el aleteo de las alas de una mariposa se puede sentir al otro lado del mundo".

de minimizar el efecto de propagación multitrayecto y, por supuesto, la posibilidad de comunicaciones con cierto grado de privacidad y de acceso aleatorio. Tanto el sistema de transmisión como el de recepción comparten un código idéntico para conseguir la demodulación y la recuperación de la señal útil. Esta tecnología tiene aplicaciones en telefonía celular, transmisión inalámbrica de datos, redes de comunicaciones personales, redes de computadoras inalámbricas, comunicaciones por satélite, sistemas de prueba y calibración, etc.

Una aplicación de secuencias pseudo-aleatorias para las marcas de agua se presenta en [8], donde se discute la viabilidad de codificar una marca de agua digital indetectable en una imagen estándar en escala de grises de ocho bits. La marca de agua es capaz de contener códigos de autenticación o autorización, o una leyenda esencial para la interpretación de la imagen. Esta capacidad se afronta para encontrar aplicación en una imagen etiquetada, para propósitos de autoría de derechos, protección falsificada, y control de acceso; sin una degradación apreciable en la imagen protegida. También se discuten dos métodos de implementación; el primero basado en la manipulación plana del bit menos significativo (***LSB, Less Significative Bit***), que ofrece decodificación fácil y rápida, y el segundo, que utiliza la suma lineal de la marca de agua a los datos de la imagen, y es más difícil de descifrar, ofreciendo seguridad inherente. En esta misma dirección en [9] se propone la construcción de marcas de agua a través de bloques arbitrarios de imagen al filtrar una secuencia ***PN***, por medio de un esquema de marca de agua para esconder información de derechos de autor en imágenes, empleando para ello un enmascarado visual para garantizar que la marca de agua incluida sea invisible y para aumentar al máximo la robustez de los datos ocultos. La marca de agua se construye para bloques arbitrarios de la imagen filtrándose una secuencia ***PN*** como un número de identificación del autor con un filtro que aproxima la frecuencia que enmascara características del sistema visual. Los resultados experimentales obtenidos en este caso, muestran que la marca de agua es robusta a varias distorsiones que incluyen ruidos blancos y de color; codificación JPEG a calidades diferentes y la segmentación de imágenes.

En el contexto de las comunicaciones inalámbricas, en [10] se hace referencia al uso de distintas capas de abstracción en un modelo de seguridad, una de las cuales considera la utilización de secuencias cifrantes ***PN*** en sistemas de espectro disperso, presentándose como una etapa en el mejoramiento de la protección de información durante su transmisión por el canal.

En los últimos años, se ha trabajado con el uso de secuencias cifrantes, como se describe en [11], en donde se proponen nuevos métodos de generación de verdaderas secuencias aleatorias a través de circuitos lógicos en los FPGA (*Field Programmable Gate Arrays*) y en ASIC (*Application Specific Integrated Circuits*), usando anillos osciladores de Fibonacci y de Galois, las secuencias binarias aleatorias generadas tienen inherentemente una alta velocidad y un índice de entropía muy alto y robusto en comparación con las propuestas anteriores para los generadores digitales de números aleatorios. Así también, se introduce un nuevo método para el post-proceso digital de datos aleatorios, basado en los circuitos lógicos síncronos no-autónomos con retroalimentación, y se presenta una técnica de reloj que controla el registro de desplazamiento de la retroalimentación lineal. El proceso posterior puede proporcionar extracción de aleatoriedad y aumento de la velocidad computacional segura de datos aleatorios de entrada. En la misma tendencia, pero con fines criptoanalíticos [12] propone dos métodos, uno para generar una secuencia pseudo aleatoria y otro para la predicción de la salida de un circuito generador de secuencias a partir de un hardware generador de números aleatorios. También se analiza este método para generar aleatoriedad, y como consecuencia de este análisis, se describe un método de ataque a este generador. Sin embargo, los diseños de hardware para la generación de números aleatorios y su evaluación no han sido tratados frecuentemente en las publicaciones. Esto significa un serio contraste, ya que por un lado está la importancia del hardware de generación de números aleatorios y su evaluación para aplicaciones de seguridad, y por otro lado, se manifiesta el poco interés que este tema ha encontrado en la literatura publicada.

Mapeos Caóticos

Las secuencias pseudo-aleatorias también pueden generarse por la aplicación recursiva de mapeos caóticos. Desde mediados de la década de los 90, se han presentado trabajos relacionados con la aplicación de la teoría del caos a la generación de secuencias cifrantes. Así, en [14] se presenta un esquema de comunicaciones de espectro disperso por salto de frecuencia usando sistemas caóticos. Se hace una descripción de una nueva familia de secuencias de salto en frecuencia generada para estos sistemas; dichas secuencias dan una dispersión uniforme sobre el ancho de banda total de la frecuencia, además de tener buenas propiedades de salto-correlación.

En aplicaciones criptográficas en [16] se aborda el problema de mezclar sistemas caóticos para las comunicaciones seguras con cifradores de flujo, con la mezcla de sistemas generadores de secuencias caóticas como métodos fidedignos para garantizar la confidencialidad; estas mezclas dan como resultado una secuencia que no tiene dependencia con la información cifrada. Los estados de salida generados por éste cifrador se aproximan al máximo teórico para las medidas de complejidad y la longitud del ciclo. Además se hace un análisis de la distribución de la función de entropía para un estado estable, así como de las secuencias de entrada de texto plano.

Un sistema digital caótico de comunicaciones, basado en un esquema *One-Time-Pad*[4] se propone en [19], donde la secuencia aleatoria se sustituye con una secuencia caótica generada por el circuito de Chua. En este sistema, se utilizó una estrategia impulsiva de control para sincronizar dos sistemas caóticos idénticos incluidos en el cifrador y el descifrador. Las longitudes de intervalos impulsivos son seccionalmente constantes, mejorando la seguridad del sistema. En [20] se explora la posibilidad de usar mapeos mezclados al diseño de cajas de sustitución (S-boxes) de acuerdo con Shannon [1], teniendo como expectativa que las propiedades buenas de difusión de tales mapeos serán heredadas por sus aproximaciones, por lo menos si el índice de convergencia es apropiado y las particiones asociadas son suficientemente buenas. En esta misma dirección en [21] se propone el diseño de secuencias binarias basadas en mapeos caóticos unidimensionales, donde se describen métodos simples de diseño de secuencias binarias caóticas con propiedades de autocorrelación exponencial y estadísticas. También emplea fragmentos monótonos unidimensionales sobre los mapeos y una función de umbral simple para generar estas secuencias, mostrando algunos diseños para este propósito.

Adicionalmente, en [22] se propone aplicar y extender los conceptos del exponente de Lyapunov y entropía a transformaciones sobre conjuntos finitos, afirmando que un sistema dinámico a tiempo discreto y con estados finitos es discretamente caótico si su exponente de Lyapunov discreto tiende a un número positivo (0 a ∞) cuando $M \rightarrow \infty$. También se discuten las aplicaciones del caos en criptografía, ya que los sistemas caóticos resultan muy atractivos a investigadores teóricos y prácticos. Recuérdese que en 1949, Shannon propuso este tipo de transformaciones en [23] por lo que no es sorprendente que cuando la Teoría del Caos floreció en las décadas de los años 80 y 90 del siglo pasado, pronto se propusieron varios criptosistemas basados en la discretización de transformaciones caóticas. Así mismo, se intenta formular el concepto de pseudo-caos o caos discreto, proponiendo como primera herramienta para medirlo al exponente de Lyapunov discreto. Adicionalmente en [24] se presenta una comparación entre la criptografía convencional y la caótica con re-inyección de la información (*message embedding*), así como una propuesta para resolver el problema de sincronización en la implementación digital de criptosistemas basados en el caos.

4 Término que empleó Vernam para un sistema criptográfico simple en el que se emplea una secuencia aleatoria de la misma longitud del texto plano.

Conclusiones

A partir del nacimiento de la Teoría del Caos, muchas disciplinas en la ciencia han visto aplicaciones convenientes en diferentes campos; no sorprende que en el terreno de la seguridad de la información tenga gran aceptación, por lo que en las últimas décadas, un gran número de trabajos importantes se han publicado sobre este tema, lo que refleja el gran interés sobre el mismo.

Los mapeos caóticos son la principal aportación que hace la Teoría del Caos a la protección de la información; dichos mapeos generan secuencias pseudo-aleatorias con propiedades estadísticas. Esto los convierte en uno de los mejores métodos para propósitos criptográficos, esteganográficos y de canales subliminales, donde se requieren secuencias cifrantes con muy buenas propiedades de impredictibilidad y aleatoriedad, además de ofrecer baja probabilidad de interceptación, sobre todo en las aplicaciones de canales subliminales.

Agradecimientos

♦ MAMV agradece el apoyo recibido del programa EDD del IPN.

♣ RSO agradece el soporte económico recibido de la Secretaría de Investigación y Posgrado del IPN (SIP-IPN), a través del proyecto 20071024 y del programa EDI, así como del Sistema Nacional de Investigadores (SNI-México).

Referencias

[1] Banks J., Dragan V., Jones A., *Chaos, a Mathematical Introduction,* Cambridge University Press. 2003. Pág. 168.

[2] Devaney R.L., *An Introduction to Chaotic Dynamical Systems*, Addison-Wesley, Redwood

[3] Freedman D. H., *Chaos Theory*, Boston.1998. Pág. 55.

[4] Rueppel, R. A., *Analysis and Design of Stream Ciphers*, Springer Verlag. 1986.

[5] Rueppel R.A.. Stream Ciphers. *Contemporary Cryptology - The Science of Information Integrity*, IEEE Press. 1992.

[6] Menezes, J. A., Van Oorschot, P.C., Vanstone, S.A., *Handbook of Appled Cryptography*, CRC,EUA, 1997.

[7] Simon, M. K., Omura, J. K., Scholtz, R. A. and Levitt, B. K., *Spread Spectrum Communication.* Vol. III, Computer Science Press, Rockville. 1985.

[8] Van Schyndel R.G., Tirkel A.Z., C.F. Osborne. *A Digital Watermark.* International Conference on Image Processing. 1994.

[9] Swanson M. D., Zhu B., Tewfik A. H., *Transparent Robust Image Watermarking.* SPIE Conf. on Visual Communications and Image Proc. 1996.

[10] Schaumont P., Verbauwhede I., *Domain-Specific Codesign for Embedded Security.* IEEE Computer Society, Volume 36, Issue 4. 2003.

[11] Golic J. D., *New Paradigms for Digital Generation and Post-Processing of Random Data.* Access Network and Terminals. Telecom Italia Lab. 2004.

[12] Dichtl M., *How to Predict the Output of a Hardware Random Number Generador.* Siemens AG, Corporate Technology. 2003.

[13] Nokolaidis N., Tsekeridou S., Solachidis V., *Applications of Chaotic Signal Processing Techniques to Multimedia Watermarketing.* Aristotle University of Thessaloniki. 2002.

[14] Cong L. , Songgeng S., *Chaotic Frequency Hopping Sequences.* Nanjing Inst. of Communications. 1998.

[15] Tsekeridou S., Nikolaidis N., Tefas A., *Theoretic Investigation of the Use of Watermark Signals Derived from Bernoulli Chaotic Sequences.* Department of Informatics Aristotle University of Thessaloniki. 2003.

[16] Philip N. S., Babu J., *Chaos for Stream Cipher.* ADCOM 2000, Tata McGraw Hill. 2001.

[17] Tefas A., Nikolaidis A., Pitas I., *Statistical Analysis of Markov Chaotic Sequences for Watermarking Applications.* Department of Informatics, Aristotle University of Thessaloniki. 2001.

[18] Masuda N., Aihara K.. *Cryptosystems With Discretized Chaotic Maps.* IEEE. Fundamental Theory and Applications, Vol. 49, No. 1. 2002.

[19] Li Z., Li K., Wen Ch., and Chai S.Y., *A New Chaotic Secure Communication System.* IEEE Transactions on Communications, Vol. 51, No. 8. 2003.

[20] Szczepanski J., Michalek T., Kocarev L. *Cryptographically Secure Substitutions Based on the Approximation of Mixing Maps.* IEEE Regular Papers, Vol. 52, No. 2. 2005.

[21] Tsuneda A., *Design of Binary Sequences With Tunable Exponential Autocorrelations and Run Statistics Based on One-Dimensional Chaotic Maps.* IEEE Regular Papers, Vol. 52, No. 2. 2005.

[22] Amigó J. M., Szczepanski J., *Caos Discreto y Criptografía*. CIBSI' 05, Valparaíso-Chile. 2005.

[23] Shannon, C. E. *Communication theory of secrecy systems*. Bell Syst. Techn. J., Págs. 656 – 715. 1949.

[24] Amigó J. M., Hernández A., Millérioux G., *Criptografía caótica con reinyección de la información.* CIBSI' 05, Valparaíso-Chile. 2005.