XỬ LÝ TÍN HIỆU SỐ BẰNG DÃY GIẢ NGẪU NHIÊN

Nguyễn Văn Sơn*, Nguyễn Hoài Giang, Đặng Hoàng Anh
Trường Đại học Mở Hà Nội

TÓM TẮT

Xáo trộn (scrambling) là một phương pháp xử lý tín hiệu giúp làm tăng tính ngẫu nhiên hay là làm trắng phổ cho chuỗi tín hiệu truyền đi. Bộ xáo trộn được sử dụng rộng rãi trong các ứng dụng như: nhận dạng hệ thống, đồng bộ, đo lường từ xa, đoán nhận kênh, và nhất là bảo vệ tin và CDMA....Có hai loại xáo trộn (scrambler): Xáo trộn đồng bộ(synchronized scrambler) và tự đồng bộ(self-synchronized scrambler). Loại đồng bộ dùng mạch ghi dịch cộng (additive LINEAR FEEDBACK SHIFTREGISTER_LFSR), loại tự đồng bộ và thiết bị nhúng tự thử (build-in selftest system BITS) dùng mạch ghi dịch nhân(chia): multiplicative LFSR. Bài báo này khảo sát hướng ứng dụng của bộ trộn đồng bộ. Bộ trộn tự đồng bộ và Thiết bị nhúng tự thử sẽ được khảo sát trong bài báo tiếp theo. Công cụ thích hợp được lựa chọn là trường Galois và biến đổi D. Ta có thể thấy bộ xáo trộn đã làm cải thiện đáng kể đặc tính ngẫu nhiên tín hiệu (làm trắng phổ).

Từ khóa: *Xáo trộn, mạch ghi dịch, ngẫu nhiên, nhận dạng hệ thống, đồng bộ, đo lường từ xa, đoán nhận kênh, cân bằng nhiễu*

Ngày nhận bài: 10/01/2019; Ngày hoàn thiện: 22/01/2019; Ngày duyệt đăng: 28/02/2019

DIGITAL SIGNAL PROCESSING JUTE PSEUDORANDOM

Nguyen Van Son*, Nguyen Hoai Giang, Dang Hoang Anh
Hanoi Open University

ABTRACT

In this paper, a randomization effect of the scrambling process on the digital signal is presented. Due to the scrambling process, the transmitted signal becomes noise-like or in other words: whitened. The random properties make the signal more suitable not only for transmission media but for other specific applications like system recognition, synchronization, CDMA and cryptography... also. There are two kinds of scramblers: synchronized scrambler and self-synchronized scrambler. While for the synchronized scrambler the additive linear feedback register (LFSR) is used, the multiplicative LFSR is used for the self-synchronized one. In this paper, the general analyzing method for both kinds of scramblers based on D-transform and trace function in Galois field theory is presented. It has been shown that the statistic properties of the scrambled signal such as state distribution, runs autocorrelation function are almost noise-like.

Keywords: *Scrambler, LFSR, randomization, system recognition, synchronization*

Received: 10/01/2019; Revised: 22/01/2019 ; Approved: 28/02/2019

* Corresponding author: *Tel: 0913 048207 ; Email: sonnv@hou.edu.vn*

GIỚI THIỆU

Để nghiên cứu và áp dụng bộ xáo trộn trong các đường truyền dẫn số, nhất thiết phải tìm hiểu và phân tích các dãy nhị phân giả ngẫu nhiên PN (Pseudorandom Noise). Công cụ toán học hữu hiệu để mô tả các dãy PN là lý thuyết trường hữu hạn và biến đổi D. Sau đây sẽ đề cập đến một số tính chất thống kê của dãy PN, đồng thời trình bày các khái niệm cơ bản về trường hữu hạn và phương pháp biểu diễn, phân tích dãy PN trên trường hữu hạn và biến đổi D.

DÃY NHỊ PHÂN NGẪU NHIÊN VÀ DÃY NHỊ PHÂN GIẢ NGẪU NHIÊN

Dãy ngẫu nhiên

Dãy ngẫu nhiên có một số tính chất như sau:

+ Tính cân bằng: Tần suất xuất hiện của '0' và '1' trong dãy ngẫu nhiên là 1/2.

+ Tính chạy: Một bước chạy được định nghĩa là một dãy con liên tiếp các ký hiệu giống nhau trong dãy ngẫu nhiên. Theo tính chạy, một nửa số bước chạy có chiều dài là 1, một phần tư số bước chạy có chiều dài là 2, một phần tám số bước chạy có chiều dài là 3,... $1/2^n$ tổng số bước chạy có chiều dài n.

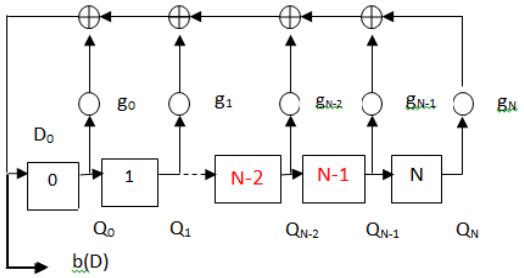
+ Tính dịch - cộng: Cộng hay dịch các dãy ngẫu nhiên sẽ tạo ra dãy ngẫu nhiên khác.

Thực tế rất khó tạo ra được một dãy số hoàn toàn ngẫu nhiên, vì vậy trong kỹ thuật người ta sử dụng các dãy nhị phân giả ngẫu nhiên (PRBS: Pseudo-Random Binary Sequence) có một chu kỳ lặp lại nào đó, hầu như thỏa mãn các yêu cầu đề ra. Chu kỳ lặp lại của PRBS được gọi là độ dài của dãy. Các PRBS như thế có thể được tạo bởi một mạch ghi dịch hồi tiếp tuyến tính LFSR (LFSR: Linear Feedback Shift Register).

+ Hàm tự tương quan nhọn: Dạng hàm delta Dirac (xem mục ***Hàm tự tương quan***)

Bộ tạo dãy nhị phân giả ngẫu nhiên (PRBS:Pseudo-Random Binary Sequence) *LFSR* [2], [5], [6]

Bộ tạo mã PN được xây dựng dựa trên mạch ghi dịch hồi tiếp tuyến tính LFSR (Linear Feedback Shift Register) có sơ đồ tổng quát minh hoạ trong hình 1.



Hình 1. *Sơ đồ tổng quát LFSR*

Như vậy, một mạch ghi dịch hồi tiếp tuyến tính gồm N trigơ D, mắc nối tiếp. Mạch hồi tiếp gồm các cổng XOR và đa thức đặc trưng g(x). N càng lớn thì chu kỳ lặp lại của dãy tín hiệu đầu ra mạch ghi dịch càng lớn.). Khi đa thức g(x) là nguyên thuỷ, dãy có chiều dài cực đại $L = (2^N - 1)$ và được gọi là dãy m.

TRƯỜNG HỮU HẠN

Trường hữu hạn GF(p)(Galois field)

Cho p là một số nguyên tố. Vành các số nguyên mod p tạo nên một trường gọi là trường Galois, ký hiệu là GF(p). Các phần tử của GF(p) có thể được ký hiệu bằng một tập các số nguyên: 0,1,2,...,p-1. Các thuật toán +, - , * , / , được thực hiện theo mod p.[6].

Các đa thức trên trường F [1], [2], [6], [7]

Biểu diễn các đa thức trên trường F

Một đa thức f(d) bậc m trên trường F có thể được biểu diễn như sau:

$f(d) = c_0 + c_1 d + ... + c_m d^m$

trong đó: ci lấy các giá trị trên trường F. Trường hữu hạn hai phần tử đóng một vai trò quan trọng trong các ứng dụng liên quan đến dãy nhị phân lấy các giá trị 0, 1 hoặc +1, -1,

Bậc của f(d), ký hiệu deg[f(d)], là số nguyên i lớn nhất, sao cho: ci≠ 0 (đa thức f(d) = 0 có bậc là 0).

Đa thức tối giản

Đa thức f(d) trên trường được gọi là tối giản nếu nó không thể phân tích thành dạng thừa số của các đa thức bậc thấp hơn trên cùng một trường.

Đa thức nguyên thuỷ

Với mọi m và p (nguyên tố), tồn tại ít nhất một đa thức tối giản bậc m và luỹ thừa $T = p^m - 1$. Đa thức đó gọi là đa thức nguyên thủy.

Trường mở rộng bậc m của GF(p)

Nếu y(d) là một đa thức tối giản trên GF(p), vành các đa thức trên GF(p) module y(d), nghĩa là cộng và nhân theo module y(d), sẽ tạo ra một trường.

Nếu bậc của y(d) là m, trường này sẽ được đại diện bởi p^m đa thức chứa d có bậc bé hơn hoặc bằng m-1, trường được gọi là trường Galois GF(cấp p^m) hoặc là trường mở rộng bậc m của GF(p) và ký hiệu là GF(p^m). [5], [6], [7].

Phương pháp biểu diễn dãy PN trên trường GF(2)

***Biểu diễn bằng biến đổi d* [1], [7], [8]**

Biến đổi d:

Có thể biểu diễn một cách thuận tiện dãy nhị phân: u_0, u_1, ... u_n bằng biến đổi d của nó, được định nghĩa như sau:

$$u(d) = u_0 + u_1 d + ... + u_n d^n \quad (1)$$

hoặc có thể viết:

$$D[u] = u(d) = \sum_{i=0}^{n} u_i d^i \quad (2)$$

trong đó D[u] là biến đổi d của u.

Biến đổi d của một day nhị phân tuần hoàn có dạng:

$$u(d) = \frac{r(d)}{q(d)} \quad (3)$$

với điều kiện q(d) không chia hết cho d và bậc của r(d) nhỏ hơn bậc của q(d).

- Nếu r(d) và q(d) nguyên tố cùng nhau thì chu kỳ của u là luỹ thừa của q(d), nghĩa là chu kỳ của u là số T nhỏ nhất sao cho q(d) chia hết $(1+d^T)$.

- Các dãy do một LFSR m tầng có đa thức đặc trưng h(d) tạo nên và ký hiệu là tập u có thể biểu diễn trong không gian d như sau:

$$u(d) = \left\{ \frac{s(d)}{h(d)},\ \deg[\, s(d)] \ < \ m \right\} \quad (4)$$

Đa thức s(d) đặc trưng cho trạng thái ban đầu của LFSR và có thể được xác định từ nội dung nhớ nhị phân ban đầu của các phần tử của LFSR qua một hệ thống phương trình tuyến tính.

- Nếu s(d) và h(d) là nguyên tố cùng nhau và h(d) không chia hết cho d thì h(d) là đa thức sinh của LFSR ngắn nhất tạo ra dãy u có biến đổi d dạng:

$$u(d) = \frac{s(d)}{h(d)} \quad (5)$$

- Nếu h(d) đa thức sinh của một LFSR là nguyên tố thì u(d) là biến đổi d của dãy m có chu kỳ: $T=2^m-1$ với m là bậc của h(d).

- Tồn tại (2^m-1) pha của một dãy m với (2^m-1) đa thức s(d) bậc nhỏ hơn (m-1).

- Gọi $D^j u$ là dãy dịch pha j nhịp so với u ta có:

$$\begin{aligned} D^j u &= u(d).d^j \ (\text{mod } h(d)) \\ &= \frac{s(d)}{h(d)}.d^j \ (\text{mod } h(d)) \end{aligned} \quad (6)$$

Phương pháp này có một số ưu điểm:

- Có thể mô tả dãy lồng ghép có độ dài bất kỳ $\left(L \neq 2^n - 1\right)$.

- Có thể sử dụng để mô tả tín hiệu (dãy) và cả phần cứng (hàm truyền đạt, hưởng ứng tự do, hưởng ứng cưỡng bức của mạch ghi dịch), vì biến đổi d là một phép biến đổi có thể dễ dàng suy từ biểu thức toán học sang dạng nhị phân và ngược lại. Do đó phù hợp với các yêu cầu ứng dụng trong kỹ thuật phân thời gian (Time multiplexing). Tóm lại công cụ toán học hữu hiệu để mô tả các dãy PN là biến đổi d.

Biểu diễn bằng hàm vết [8,9]

Hàm vết là một ánh xạ tuyến tình từ trường mở rộng GF(p^m) xuống trường con GF(p) và được định nghĩa một cách tổng quát như sau:

$$Tr_p^{p^m}(\alpha) = \sum_{i=0}^{m-1} \alpha^{p^i} \quad (7)$$

α là phần tử của GF(p^m).

Trong trường nhị phân GF(2), công thức trên trở thành:

$$\mathrm{Tr}_1^m(\alpha)=\sum_{i=0}^{m-1}\alpha^{2^i},\alpha\in GF(2^m) \quad (8)$$

Trong đó, thay $\mathrm{Tr}_2^{2^m}$ bằng Tr_1^m để biểu diễn đơn giản hơn.

Một số tính chất quan trọng sau đây của hàm vết có thể giúp cho việc tính toán và khảo sát các dãy PN một cách linh hoạt, rõ ràng hơn:

$\forall\alpha,\beta\in GF(p)$, ta có:

$$\mathrm{Tr}(\alpha)\in GF(p^m) \quad (9)$$

$$\mathrm{Tr}(\alpha+\beta)=\mathrm{Tr}(\alpha)+\mathrm{Tr}(\beta) \quad (10)$$

$$\mathrm{Tr}(\lambda\alpha)=\lambda\mathrm{Tr}(\alpha),\quad \lambda\in GF(2^m) \quad (11)$$

$$\mathrm{Tr}(\alpha^p)=\mathrm{Tr}(\alpha) \quad (12)$$

Dãy nhị phân cực đại (dãy m) có chu kỳ 2^m-1 có thể được biểu diễn bằng hàm vết như sau:

$$\{a_n\}=(a_0,a_1,...,a_{2^m-2})=\left(\mathrm{Tr}_1^m(\alpha^0),\mathrm{Tr}_1^m(\alpha),...,\mathrm{Tr}_1^m(\alpha^{2^m-2})\right)$$

Phương pháp biểu diễn bằng hàm vết còn được gọi là biểu diễn bằng phần tử nguyên thủy(α). Biểu diễn bằng hàm vết có ưu điểm là công thức biểu diễn ngắn gọn. Tuy nhiên, sử dụng hàm vết có nhược điểm là

+ Hàm vết chỉ được định nghĩa cho các dãy có độ dài $(L=2^n-1)$, trong khi đó các dãy có độ dài $(L\neq 2^n-1)$ thì hàm vết không thể biểu diễn được. Lúc đó phải sử dụng một công cụ khác.

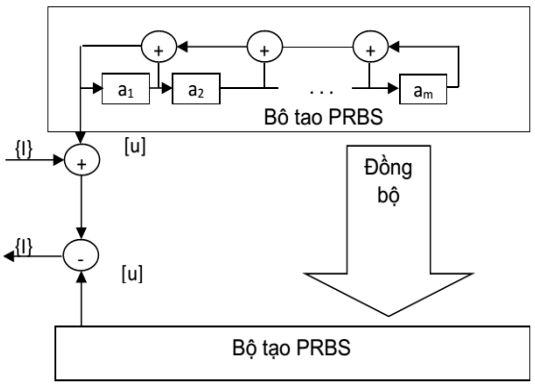
+ Hàm vết không thể biểu diễn được hưởng ứng cưỡng bức máy trình tự tuyến tính

BỘ XÁO TRỘN TÍN HIỆU (SCRAMBLER)

Khái niệm bộ xáo trộn

Bộ xáo trộn số là một thiết bị dùng để tạo nên một sự thay đổi cần thiết trong dòng thông tin nhị phân bất kỳ. Những tính chất mong muốn của dãy ra, khi đưa vào đầu vào bộ xáo trộn một dòng nhị phân bất kỳ, là: sự cân bằng giữa các bit '1' và '0' và hàm tự tương quan nhọn, hay nói khác đi: phân bố lại các bit '1' và '0' để đạt được một xác suất trạng thái và một giá trị hàm tự tương quan (ACF) đã cho. Do đó các bộ xáo trộn số còn được sử dụng rộng rãi trong thủy văn, mật mã, dấu tin (steganography)....[12-14]. Có hai loại scrambler synchronized scrambler (Xáo trộn đồng bộ) và self-synchronized scrambler (Tự đồng bộ). Bộ xáo trộn đồng bộ (thiết lập lại Reset scrambler) được mô tả ở hình 2 ở phần phát, việc cộng mô-đun p theo từng symbol giữa dãy số liệu {I}(vào) với dãy giả ngẫu nhiên {u} tạo thành dãy số liệu đã được xáo trộn {O}(ra).

$$\{O\}=\{I\}+\{u\}\bmod p \quad (13)$$



Hình 2. *Bộ xáo trộn đồng bộ*

Thay cho việc truyền đi dãy số liệu nguyên thủy {I}, phần phát truyền di dãy đã xáo trộn {O}. Tại máy thu, chúng ta có một bộ tạo PRBS hoàn toàn giống và đồng bộ với bộ tạo PRBS phần phát. Trong trường hợp nhị phân (*p = 2*), chúng ta có dãy bít số liệu tách được nhờ theo mô-đun 2 từng bít của dãy nhận được với từng bít của PRBS tạo được ở phần thu:

$$\{I\}=\{O\}\oplus\{u\} \quad (14)$$

Hiển nhiên, các bộ tạo PRBS phần phát và thu nhất thiết phải đồng bộ với nhau. Nhược điểm căn bản của bộ xáo trộn và giải xáo trộn "thiết lập lại" là cần phải có các thiết bị đồng bộ. Bù lại, lợi thế của chúng là các lỗi truyền dẫn không gây nên các bội lỗi tại phần thu.

Hiệu quả xử lý của scrambler

Các dãy bít thông tin có thể có những độ dài tương đối ngắn và những loạt dài các bít không chuyển đổi cực tính. Điều này dẫn đến những đặc điểm bất lợi:

- Phổ của tín hiệu truyền đi phụ thuộc vào mẫu của dãy bít được truyền.

- Các vạch phổ của tín hiệu khá thưa trên thang tần số và đồ thị phổ khá cao ở những tần số thấp.

Mục đích cơ bản của thuật toán xáo trộn là loại bỏ các chu kỳ ngắn trong dãy tín hiệu lối vào và khử bỏ các loạt dài, không phụ thuộc mẫu dãy bít lối vào như thế nào.

Hưởng ứng của các bộ xáo trộn tín hiệu

Bộ xáo trộn có thể được mô hình hóa như một máy trình tự tuyến tính. Như vậy, dãy ra có thể được chia thành hai thành phần độc lập: hưởng ứng tự do và hưởng ứng cưỡng bức. Hưởng ứng tự do là do trạng thái ban đầu của bộ trộn quyết định, còn hưởng ứng cưỡng bức là do dãy vào quyết định. Đặc tính vào – ra của bộ trộn có thể được mô tả một cách đơn giản qua hàm truyền đạt H(d) trong không gian d như trong biểu thức:

$$Y(d) = X(d).H(d) \quad (15)$$

Trong đó X(d) và Y(d) lần lượt là biến đổi d của dãy vào x(n) và dãy ra y(n). Hàm truyền đạt của bộ giải xáo trộn sẽ là 1/H(d).

Phân bố xác suất

Gọi {O} là dãy đầu ra của bộ trộn số, {I} là dãy đầu vào (hay hưởng ứng cưỡng bức của bộ xáo trộn) và {u} là dãy tạo bởi LFSR(tín hiệu hay hưởng ứng tự do của bộ xáo trộn). Xác suất bit '0' và bit '1' của dãy đầu ra tương ứng là $P_O(0)$ và $P_O(1)$. Xác suất bit '0' và bit '1' của dãy đầu vào tương ứng là $P_I(0)$ và $P_I(1)$. Xác suất bit '0' và bit '1' của dãy tạo bởi LFSR tương ứng là $P_u(0)$ và $P_u(1)$. Vì bộ xáo trộn là một hệ thống tuyến tính xếp chồng nên dãy đầu ra được tính theo công thức

$$\{O\} = \{I\} \oplus \{u\} \quad (16)$$

Vậy, xác suất bit '1' trong dãy ra {O} được tính:

$$P_O(1) = P_I(1).P_u(0) + P_I(0).P_u(1) \quad (17)$$

Dãy tạo bởi LFSR {u} thỏa mãn tính cân bằng, nghĩa là: xác suất bít '0' của dãy PRBS do LFSR tạo ra và xác xuất bit '1' thỏa mãn:

$$P_u(1) \approx P_u(0) \approx \frac{1}{2} \quad (18)$$

Vậy, ta có:

$$P_O(1) = P_I(1).\frac{1}{2} + P_I(0).\frac{1}{2} = \frac{1}{2}\left[P_I(1) + P_I(0)\right] \approx \frac{1}{2}$$

Vậy, với dãy vào {I} và dãy {u} tạo bởi LFSR là độc lập thống kê, ta có dãy ra {O} sẽ có phân bố xác suất:

$$P_O(1) \approx P_O(0) \approx \frac{1}{2} \quad (19)$$

Các dãy dài n bít không chuyển đổi mức ở lối ra cũng xảy ra với xác suất rất thấp, có xác suất như đối với dãy {u}, tức là= $1/2^n$.

Hàm tự tương quan

Hàm tự tương quan của dãy ra {O} được định nghĩa như sau:

$$R(k) = \frac{A - D}{A + D} \quad (20)$$

Trong đó: A là số bit giống nhau giữa dãy ban đầu và dãy được dịch đi k bit (hay dịch đi một khoảng thời gian τ), D là số bit khác nhau giữa hai dãy.

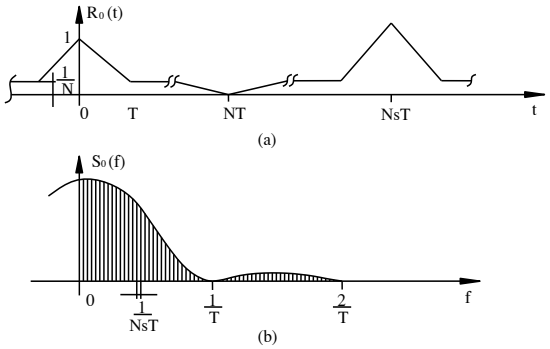
Biến đổi toán học R(k) ta được:

$$R(k) = \frac{A - D}{A + D} = \frac{A + D - 2D}{A + D} = 1 - 2\frac{D}{A + D} = 1 - 2P_O(1) \approx 0 \quad (21)$$

Vậy, dãy ra của bộ xáo trộn {O} thỏa mãn hai thuộc tính ngẫu nhiên:

$$\begin{cases} P_O(1) \approx P_O(0) \approx \frac{1}{2} \\ R(k) \approx 0 \end{cases}$$

Có thể thấy các vạch phổ sít nhau hơn và do đó đồ thị phổ của dãy khá thấp, tức là đồ thị phổ của dãy được truyền có đặc tính khá gắn với đồ thị phổ của một nhiễu trắng.



Hình 3. *Hàm tự tương quan của tín hiệu sau xáo trộn*

Xét hàm tự tương quan của dãy ra bộ trộn trong trường hợp tổng quát:

Gọi $\{i_n\}$ là dãy vào, với $\{i_n\} \in 0,1,\ 0 < n < \infty$, và I(d) là biến đổi d của dãy này. Đa thức h(d) bậc m là đa thức sinh của bộ trộn và S(d) có bậc nhỏ hơn m đặc trưng cho trạng thái ban đầu của bộ trộn. Dãy ra có dạng như sau:

$$O(d) = \frac{I(d) + S(d)}{h(d)} = O_f(d) + O_a(d) \qquad (22)$$

Trong đó, $O_f(d) = I(d)/h(d)$ và $O_a(d) = S(d)/h(d)$ lần lượt biểu diễn hưởng ứng tự do và hưởng ứng cưỡng bức của bộ trộn trong không gian d.

Khi I(d) và S(d) là độc lập, nghĩa là được chọn một cách độc lập, thì từ tính chất tuyến tính của bộ trộn, ta có: hưởng ứng cưỡng bức và hưởng ứng tự do là độc lập với nhau.

Đặt:

$$D^{-1}\left[O_f(d) + O_a(d)\right] = o_n^f + o_n^a = o_n \qquad (24)$$

là biến đổi d ngược của dãy ra, trong đó o_n là dãy ra. ACF của dãy nhị phân ra được định nghĩa là:

$$R(k) = E\left\{a(o_n).a(o_{n+k})\right\} \qquad (25)$$

Trong đó, $a(o_n)$=1 khi o_n=1 và $a(o_n)$=-1 khi o_n=0. Người ta chứng minh được rằng: ACF của một dãy ngẫu nhiên, có thể được tính như sau

$$\begin{aligned} R(k) &= 1 - 2\left[P\{a(o_n) = -1, a(o_{n+k}) = 1\} + P\{a(o_n) = 1, a(o_{n+k}) = -1\}\right] \\ &= 1 - 2P\{a(o_n) \oplus a(o_{n+k}) = 1\} \\ &= 1 - 2P_1\{a(o_n) + a(o_{n+k})\} \end{aligned}$$

$$R(k) = 1 - 2P_1\{a(o_n) + a(o_{n+k})\} \qquad (26)$$

Trong đó, $\oplus$ là phép cộng mô-đun 2, ở đây ta ký hiệu là + vì các phép tính liên quan đến dãy nhị phân dĩ nhiên là mô-đun 2, và $P_1\{x_n\}$ là xác suất 1 trong dãy x_n.

Như vậy, ta có:

$$\begin{aligned} &P\{a(o_n) \oplus a(o_{n+k})\} = \\ &= P\left\{\left(a^f(o_n) + a^f(o_{n+k})\right) + \left(a^a(o_n) + a^a(o_{n+k})\right) = 1\right\} \\ &= P\left\{x_n^f(k) + x_n^a(k) = 1\right\} = P_1\left\{x_n^f(k) + x_n^a(k)\right\} \end{aligned}$$

Trong đó:

$$x_n^f(k) = a^f(o_n) + a^f(o_{n+k})$$

$$x_n^a(k) = a^a(o_n) + a^a(o_{n+k})$$

Ta giả thiết tín hiệu vào và tín hiệu do bộ xáo trộn tạo ra là độc lập thống kê

Do $a^a(o_n)$ và $a^a(o_{n+k})$ là hai dãy m lệch pha nhau nên $x_n^a(k)$ cũng là một dãy m, ta có:

$$P_1\left(x_n^a(k)\right) = P_0\left(x_n^a(k)\right) = \frac{1}{2} \text{ khi } m >> 1$$

Và:

$$q_1 = P\left\{x_n^f(k) = 1\right\} = P_1\left\{x_n^f(k)\right\}$$

$$q_0 = 1 - q_1$$

Vì $x_n^f(k)$ và $x_n^a(k)$ là độc lập, nên ta có:

$$P_1\left\{x_n^f(k) + x_n^a(k)\right\} = P_1.q_0 + P_0.q_1 = \frac{1}{2}$$

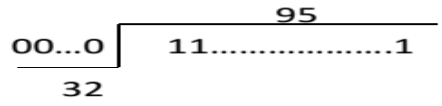
Do đó:

$$R(k) = 1 - 2P_1\left\{x_n^f(k) + x_n^a(k)\right\} = 0\ . \qquad (27)$$

Nói cách khác, tín hiệu vào sẽ được ngẫu nhiên hóa một cách hiệu quả không phụ thuộc vào tính chất thống kê của nó.

Có thể thấy rõ điều này khi thực hiện mô phỏng quá trình ngẫu nhiên hóa tín hiệu, sử dụng hai bộ trộn số mắc nối tiếp, mỗi bộ trộn số là một LFSR gồm 7 trigơ, tạo được dãy m

gồm 127 bit. Tín hiệu vào là dãy 127 bit nhị phân, có dạng bất kỳ, tuy nhiên ở đây chỉ thực hiện mô phỏng với tín hiệu vào xấu. Ở trường hợp này, tín hiệu vào là dãy 127 bit, trong đó: gồm 95 bit '1', 32 bit '0', tối đa có 95 bit '1' liên tiếp và 32 bit '0' liên tiếp, có một lần chuyển mức tín hiệu.



Hình 4. *Dãy tín hiệu vào xấu*

Dãy tín hiệu vào được đưa qua bộ trộn số thứ nhất, xáo trộn tín hiệu lần 1, tín hiệu ra của bộ trộn số thứ nhất được đưa vào bộ trộn số thứ hai, xáo trộn tín hiệu lần 2. Tín hiệu ra của bộ trộn số thứ hai là dãy tín hiệu ngẫu nhiên được truyền đi. Trong đó: m_1 là đa thức sinh của bộ trộn 1, m_2 là đa thức sinh của bộ trộn 2

N_1 : ký hiệu số bit '1' trong dãy tín hiệu ra.

N_0 : ký hiệu số bit '0' trong dãy tín hiệu ra.

N_1 / N_0 : tỷ số số bit '1' trên số bit '0'.

M_1 : ký hiệu số bit '1' tối đa liên tiếp (cụm '1' tối đa).

M_0 : ký hiệu số bit '0' tối đa liên tiếp (cụm '0' tối đa).

TR: ký hiệu số lần chuyển mức của dãy tín hiệu ra.

Ở trường hợp này: tín hiệu vào là dãy 127 bit, trong đó: gồm 95 bit '1', 32 bit '0', tối đa có 95 bit '1' liên tiếp và 32 bit '0' liên tiếp, có một lần chuyển mức tín hiệu.

Bảng 1. *Tín hiệu vào và ra sau xáo trộn*

m_1	m_2	N_1	N_0	N_1/N_0	M_1	M_0	TR
10000011	10000011	95	32	2.97	95	32	1
	10010001	75	52	1.44	9	4	65
	10101011	69	58	1.19	8	4	65
	11000001	67	60	1.12	8	8	54
	11010101	55	72	0.76	10	8	59
	11110001	51	76	0.67	5	10	60
	10001001	77	50	1.54	10	8	55
	10011101	53	74	0.72	7	10	57
10010001	10101011	69	58	1.19	8	5	65
	11000001	53	74	0.72	7	11	56
	11010101	59	68	0.87	6	9	62
	11110001	63	64	1.02	5	5	72
	10001001	55	72	0.76	4	6	67
	10011101	53	74	0.72	6	7	57
10101011	11000001	51	76	0.67	4	10	60
	11010101	55	72	0.76	4	9	64
	11110001	61	66	0.92	10	8	56
	10001001	75	52	1.44	8	4	63
	10011101	51	76	0.67	4	8	65
11000001	11010101	61	66	0.92	5	9	63
	11110001	61	66	0.92	8	5	57
	10001001	63	64	0.98	9	8	62
	10011101	67	60	1.12	4	4	72
11010101	11110001	63	64	0.98	6	8	55
	10001001	57	70	0.81	5	5	64
	10011101	65	62	1.05	7	4	62
11110001	10001001	41	86	0.48	4	14	49
	10011101	61	66	0.92	6	5	64
10001001	10011101	59	68	0.87	5	5	63

Đây là trường hợp mà tín hiệu vào xấu nhất so với các trường hợp khác. Sau khi được xáo trộn, tín hiệu thu được có các thuộc tính:

- Số lần chuyển mức tín hiệu của dãy vào nhỏ nhất. Dãy ra hầu hết đều đạt gần 50% số chuyển mức có thể.

- Tỷ lệ chênh lệch N_1 / N_0 của dãy vào cao (2.97 lần), nhờ xáo trộn, dãy ra có tỷ lệ N_1 / N_0 thấp hơn nhiều.

- Tín hiệu vào gồm hai bước chạy, bước chạy '0' chiều dài 32 và bước chạy '1' chiều dài 95. Chiều dài bước chạy của tín hiệu sau xáo trộn nhỏ hơn nhiều.

Một số trường hợp kết hợp các đa thức đặc trưng khác nhau để tạo bộ xáo trộn, thu được dãy tín hiệu ra với các thuộc tính đặc biệt tốt:

Khi chọn đa thức đặc trưng của các LFSR là hai đa thức đối ngẫu: $h_1(d) = 1 + d^6 + d^7$ và $h_2(d) = 1 + d + d^7$ tín hiệu thu được sau xáo trộn có: 67 bit '1', 60 bit '0', tỷ lệ N_1 / N_0 là 1.12, có hai bước chạy độ dài 8 (một bước chạy gồm 8 bit '1' và một bước chạy gồm 8 bit '0'), còn lại là các bước chạy có độ dài từ 5 trở xuống, có 54 lần chuyển mức tín hiệu.

Khi chọn đa thức đặc trưng của các LFSR là $h_1(d) = 1 + d^3 + d^7$

Và $h_2(d) = 1 + d + d^2 + d^3 + d^7$ tín hiệu thu được sau xáo trộn có: 63 bit '1', 64 bit '0', tỷ lệ N_1 / N_0 là 0,98, có hai bước chạy độ dài 5 (một bước chạy gồm 5 bit '1' và một bước chạy gồm 5 bit '0'), còn lại là các bước chạy có độ dài nhỏ hơn, có 72 lần chuyển mức tín hiệu.

Khi chọn đa thức đặc trưng của các LFSR là $h_1(d) = 1 + d + d^7$

và $h_2(d) = 1 + d^3 + d^4 + d^5 + d^7$ tín hiệu thu được sau xáo trộn có: 67 bit '1', 60 bit '0', tỷ lệ N_1 / N_0 là 1.12, có năm bước chạy độ dài 4 (bốn bước chạy gồm 4 bit '1' và một bước chạy gồm 4 bit '0'), còn lại là các bước chạy có độ dài nhỏ hơn, có 72 lần chuyển mức tín hiệu.

Khi chọn đa thức đặc trưng của các LFSR là

$h_1(d) = 1 + d + d^2 + d^3 + d^7$

và $h_2(d) = 1 + d^3 + d^4 + d^5 + d^7$ tín hiệu thu được sau xáo trộn có: 61 bit '1', 66 bit '0', tỷ lệ N_1 / N_0 là 0.92, có một bước chạy độ dài 6 (gồm 6 bit '1'), hai bước chạy độ dài 5 (gồm 5 bit '0'), còn lại là các bước chạy độ dài nhỏ hơn, có 64 lần chuyển mức tín hiệu.

Như vậy, ta thấy trong các trường hợp đặc biệt của tín hiệu vào (tín hiệu vào là dãy gồm nhiều bit '1' liên tiếp và nhiều bit '0' liên tiếp), khi chọn các đa thức đặc trưng của LFSR phù hợp, tín hiệu ra sau xáo trộn sẽ có những thuộc tính gần thỏa mãn các thuộc tính của dãy ngẫu nhiên, như: số bit '1' và bit '0' chênh lệch nhau không nhiều, một số trường hợp chênh lệch giữa số bit '1' và số bit '0' không quá một bit, độ dài bước chạy giảm đáng kể, số lần chuyển mức tín hiệu đủ lớn để đảm bảo dễ dàng khôi phục tín hiệu định thời tại phía thu.

KẾT LUẬN

Trên đây, đã sử dụng phương pháp biểu diễn xáo trộn và các dãy giả ngẫu nhiên chiều dài cực đại tạo bởi các mạch ghi dịch hồi tiếp tuyến tính bằng đa thức trên trường GF(2) để phân tích các thuộc tính tín hiệu ra. Ta có thể thấy bộ xáo trộn đã làm cải thiện đáng kể chất lượng truyền dẫn. Do đó chúng được sử dụng rộng rãi trong các ứng dụng như: nhận dạng hệ thống, đồng bộ, đo lường từ xa, đoán nhận kênh, cân bằng nhiễu và nhất là mật mã (cryptography).

LỜI CẢM ƠN

Các tác giả bài báo xin trân thành cảm ơn sự hỗ trợ kinh phí nghiên cứu khoa học của Trường Đại học Mở Hà Nội thông qua đề tài cấp Trường mã số V2018-12.

TÀI LIỆU THAM KHẢO

1. R. G. Gitlin, J. F. Hayer (1975), "Timing recovery and scramblers in data transmission," *Bell.Syst.Tech. Journal*, vol 54. N^o 3, pp. 589-593, Mar. 1975.

2. Quynh L. C. (1985), PhD dissertation IIT DelHi-INDIA.

3. X. B. Liu et al (2012), "Reconstructing a Linear Scrambler With Improved Detection Capability and in the Presence of Noise" *IEEE transactions on information forensics and security*, vol. 7, no. 1, pp 208-18, February 2012.

4. Jin Zhang (2013), EPoC Scrambler, IEEE 802.3bn EPoC TF Meeting Nov. 2013.

5. H.J. Zepernick (2005), A. *Finger Pseudo Random Signal Processing Theory and Application*, John Wiley & Sons Ltd

6. P.Z. Fan and M. Darnell (1996), Sequence Design for Communications Applications, New York: Wiley.

7. Hieu Le Minh et al (2015), "Design and Analysis of Ternary m-sequences with Interleaved Structure by d-Transform", *Journal of Information Engineering and Applications*, Vol.5, No.8, pp. 97-101.

8. Quynh L Ch et al (2016), "A Hardware Oriented Method to Generate and Evaluate Nonlinear Interleaved Sequences with Desired properties", *Journal of Information Engineering and Applications*, Vol.6, No.7, pp.1-12.

9. Z.Dai, G.Gong, H.Y.Song, D.Ye (2011), "Trace Representation and Linear Complexity of Binary eth Power Residue Sequences of Period P", IEEE Trans. on information theory, Vol.57, No.3, pp 1530-1547, March 2011.

10. C.-Y. Lai and C.-K. Lo (2002), "Nonlinear orthogonal spreading sequence design for third generation DSCDMA systems", IEE Proceeding commun. vol 149 n2, pp 405-410.

11. W Golomb and G. Gong (2005), "Signal Design for Good Correlation - for Wireless Communication, Cryptography and Radar," Cambridge University Press.

12. T. Kang et al (2013), "A Survey of Security Mechanisms with Direct Sequence SpreadSpectrum Signals", Journal of Computing Science and Engineering,Vol. 7, No. 3, September, pp. 187-197.

13. R.Kazemi et al (2016), "Data Hiding Robust to Mobile Communication Vocoders", EEE transactions on multimedia, pp 1.