L-007

S-box による xorshift 乱数の非線形化
Nonlinearization of xorshift random numbers with S-box

劉 忠達† 佐々木 慶文†
Zhongda Liu Yoshifumi Sasaki

1. はじめに

Xorshift [1]乱数生成法は排他的論理和とビットシフト演算のみで疑似乱数（以降，乱数）を生成するので，非常に高速である．これまでに，xorshift をベースした高速な乱数生成法[2][3]が次々と提案されているが，これら乱数生成法は，内部状態であるコンパニオン行列が推測しやすく，そのままで暗号に応用することができない．そこで，本研究では，DES 暗号に使われている一方向性関数 S-box を用いて xorshift 乱数の非線形化を行った．更に均等分布をするようにメルセンヌ・ツイスタの調律と呼ぶ変換を実施した．評価実験では，非線形化した乱数に対して NIST 乱数検定を行い，元の xorshift 乱数と比較した．その結果，乱数の性質が向上していることがわかった．

2. Xorshift 乱数

整数は w ビットのバイナリベクトルで表される（$w=32$ *或は* 64）．w ビットのベクトル x と行列L^aとの掛け算$L^a x$は C 言語のビット左シフト演算 x<<a で実現できる$(a<w)$．L は$w\times w$バイナリ行列であり，対角項は 1，それ以外は 0 になる．L^aは対角項から右上に a 個ずれた場所に 1 になっている行列である．そして，XOR 処理を追加した x^(x<<a)は$(I+L^a)x$を実現することができる（I は単位行列）．同様に右シフト演算を用いて，x^(x>>b)で$(I+R^b)x$を表現できる．適切なa,b,c（表 1）を選び，$T=(I+L^a)(I+R^b)(I+L^c)$で構成したコンパニオン行列Tは可逆行列である．Tのなす一般線形群の位数は2^w-1である．$Tx, T^2x, \cdots, T^k x=x$ $(k=2^w-1)$ によって，周期が2^w-1の乱数を生成することができる．

表 1　a,b,c（の候補（$w=32$）

1, 3, 10	1, 5 ,16	1, 5, 19	1, 9, 29	1, 11, 6	1, 11, 16	1, 19, 3	1,21,20	1,27,27
2, 5, 15	2, 5, 21	2, 7, 7	2, 7, 9	2, 7, 25	2, 9, 15	2,15,17	2,15,25	2,21, 9
3, 1,14	3, 3,26	3, 3,28	3, 3,29	3, 5,20	3, 5,22	3, 5,25	3, 7,29	3,13, 7
3,23,25	3,25,24	3,27,11	4, 3,17	4, 3,27	4, 5,15	5, 3,21	5, 7,22	5, 9,7
5, 9,28	5, 9,31	5,13, 6	5,15,17	5,17,13	5,21,12	5,27, 8	5,27,21	5,27,25
5,27,28	6, 1,11	6, 3,17	6,17, 9	6,21, 7	6,21,13	7, 1, 9	7, 1,18	7, 1,25
7,13,25	7,17,21	7,25,12	7,25,20	8, 7,23	8,9,23	9, 5,14	9, 5,25	9,11,19
9,21,16	10,9,21	10,9,25	11,7,12	11,7,16	11,17,13	11,21,13	12, 9,23	13, 3,17
13, 3,27	13, 5,19	13,17,15	14, 1,15	14,13,15	15, 1,29	17,15,20	17,15,23	17,15,26

3. Xorshift 乱数の非線形化

内部状態であるコンパニオン行列 T を決めるパラメータは表 1 の通りに限定されるため，鍵全数探索攻撃などでパラメータが解明される可能性が非常に高く，そのままでは暗号に応用することができない．そこで，本研究では，図 1のように，DES暗号に使われている一方向性関数 S-box[4]を用いて xorshift 乱数の非線形化を行う．

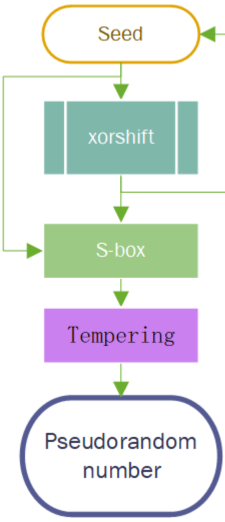

図 1　非線形化の流れ

32 ビットのシードを xorshift に入力すると，内部のコンパニオン行列 T によって線形変換されたベクトルが出力される．この出力を新たなシードとして xorshift に入力し，次のベクトルを得る．このようにして，ベクトルのシーケンスが生成される．

S-box は DES 暗号[4]の換字（substitution）処理を行う．1 つの S-box では，内部の換字表によって 6 ビットのブロックを 4 ビットに変換する．本研究では，xorshift の入力と出力からなる 48 ビットのベクトルを 8 つの S-box で換字処理し，32 ビットの乱数列$y_0, y_1, y_2, \cdots$ を生成する．しかし，S-box の換字処理によって乱数列に偏りが生じたため，メルセンヌ・ツイスタ[5]の調律行列 S をかけて$Sy_0, Sy_1, Sy_2, \cdots$を出力し，均等分布をするように改良した．

4. NIST 検定

暗号に応用する乱数列の性質を解析するには，統計的な検定手法を用いて乱数列の性質を調べることができる NIST 検定を用いるのが一般的である．

本研究が使用した NIST Special Publication 800-22 Revision 1a [5]では，以下の検定が採用されている．

(1) Frequency Test
　一様性の検定

(2) Frequency Test within a Block
　ブロック単位の一様性の検定

(3) Runs Test
　連の検定

(4) Test for Longest Run of Ones in a Block

† 石巻専修大学 理工学部
Faculty of Science and Engineering,
Ishinomaki Senshu University

ブロック単位の最長連検定
(5) Binary Matrix Rank Test
2 値行列階数検定
(6) Discrete Fourier Transform (Spectral) Test
離散フーリエ変換検定
(7) Non-overlapping Template Matching Test
重なりのないテンプレート適合検定
(8) Overlapping Template Matching Test
重なりのあるテンプレート適合検定
(9) Maurer's "Universal Statistical" Test
同じパターンの出現間隔を用いた検定
(10) Linear Complexity Test
線形複雑度検定
(11) Serial Test
パターンの長さと出現の頻度に関する検定
(12) Approximate Entropy Test
近似エントロピー検定
(13) Cumulative Sums (Cusum) Test
累積和検定
(14) Random Excursions Test
ランダム偏差検定
(15) Random Excursions Variant Test
種々のランダム偏差検定

検定ごとに $p-value$ という数値が得られる．$p-value$ は，検定した標本が真の乱数列である確率を表す．NIST 検定では，有意水準 $\alpha=0.01$ であり，$p-value<0.01$ の時に良い乱数列ではないと判断する．そして，複数の標本系列 (1000 程度を推奨している) に対し検定を行い，以下の２つ指標で評価する．

1. $p-value$ の一様性
2. $p-value \geq 0.01$ の割合

1. では，$p-value$ が区間 [0, 1) で一様に分布しているかどうかを調べる．[0, 1) を 10 の区間に分割し，分割した区間ごとの頻度が一様になっているかどうかをカイ 2 乗検定する．カイ 2 乗検定により得られた $p-value$ が 0.0001 以上ならば，良い乱数列乱数であると判断する．また、2. では、標本の数を m としたとき、0.01 以上となる $p-value$ の数の割合が $(1-\alpha)\pm 3\sqrt{\frac{(1-\alpha)\times\alpha}{m}}$ の範囲（$m=1000$ の時，約 0.9805607 以上）に入っている場合は，乱数列は良い乱数であると判断する．

本研究では，Intel Xeon Silver 4280 CPU (2.10GHz)の環境で 1000 標本の NIST 検定を行った．このために，同じシードを用いて，提案する非線形化アルゴリズムと xorshift により乱数列を生成した．乱数列のサイズは130MBとした．生成速度は xorshift が 5.4Gbps，提案手法が 0.2Gbps であった． 8 つの S-box の換字処理が速度の低下を引き起こしたと考えられる．

検定結果として，xorshift の乱数列は(5) 2 値行列階数検定において，全ての標本について失敗した．節 2 で述べたように，xorshift は 2 値行列の掛け算を行って乱数を生成するので，ある種の線形従属性がある．そのために，関連する 2 値行列階数検定が全て失敗したと考えられる．一方，提案する非線形化アルゴリズムの乱数列の一部は(2) ブロック単位の一様性の検定と(11) パターンの長さと出現の頻度に関する検定が失敗した．S-box 換字表の調整によって改善できると考えられる．

5. まとめ

本研究では，暗号に応用するために xorshift 乱数の非線形化を行った．NIST 検定によって，乱数列の性質は良好であることが分かった．課題として，乱数生成速度の改善が挙げられる．

参考文献

[1] George Marsaglia, "Xorshift rngs", Journal of Statistical Software, Vol.8, No.14 (2003).
[2] François Panneton, Pierre L'ecuyer, "On the xorshift random number generators", ACM Transactions on Modeling and Computer Simulation (TOMACS) , Vol.15, No.4 (2005).
[3] Sebastiano Vigna, "Further scramblings of Marsaglia's xorshift generators", Journal of Computational and Applied Mathematics, Vol.315 (2017).
[4] National Institute of Standards and Technology, "Data Encryption Standard", Federal Information Processing Standards Publication, No. 46 (1977)-
[5] Makoto Matsumoto, Takuji Nishimura, "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator." ACM Transactions on Modeling and Computer Simulation (TOMACS) , Vol.8, No.1 (1998).
[6] Andrew Rukhin, Juan Soto, James Nechvatal, et al. "A statistical test suite for random and pseudorandom number generators for cryptographic applications", NIST Special Publication 800-22 Revision 1a (2010).