# 音声及びダウンサンプリング時の折り返し雑音の乱数性の検証

川村　大河　　黒澤　安奈　　長瀬　智行

弘前大学理工学部　〒036-8224　青森県弘前市大字文京町　3
E-mail: nagase@eit.hirosaki-u.ac.jp

**あらまし**　今日の暗号技術において、乱数生成技術はさまざまな局面で必要不可欠な技術であり、性能の良い暗号論的乱数生成器は現在でも渇望されている。特に、真の乱数を発生させるためには高価な機材等が必要であり、安価に実現できる真の乱数発生器はさまざまな場面で必要とされている。そこで、身近な自然現象である音を利用し、真の乱数もしくは、それに順ずる乱数列を発生させることの可能な乱数生成器を製作することはできないかと考えた。本研究では、そのための前段階として、音声及び、音声をダウンサンプリングした際に発生する折り返し雑音の乱数性について検証を行った。

# Folding Noise Based Random Generator for Improving Cryptography Systems

Taiga Kawamura, Anna Kurosawa and Tomoyuki Nagase

Faculty of Science and Engineering, Hirosaki University
3 Bunkyo-cho, Hirosaki-shi, Aomori, 036-8224 Japan
E-mail: nagase@eit.hirosaki-u.ac.jp

**Abstract** Random number generation techniques have numerous applications in science and engineering where unpredictable values are generated. Some of security applications such as cryptography require a very high degree of randomness. In this report, we introduce a new method for generating real randomness based on physical phenomenon such as folding noise to generate a highly unpredictable sequence.

## 1. はじめに

昨今の暗号技術において、乱数生成技術は必要不可欠であり、それはさまざまな局面にて必要とされている。例えば、暗号化の際の鍵などはランダムな数字が好ましいとされ、その最たる例が使い捨てパッドであることは暗号の世界では常識となっている。そして、著名な擬似乱数生成器として挙げられる線形合同法や、ラグ付きフィボナッチ法、メルセンヌツイスター法[1]などは、暗号論的乱数生成器としては、その性質上、そのような用途に使用することは不可能である。

それゆえ、良質な暗号論的乱数生成器は現在でも渇望されており、真の乱数、及びそれに近い性質を持つ乱数生成器の研究は現在でも絶えず行われている。中でも真の乱数は、自然現象を利用するという性質上、発生させるために高価な機材を必要とするものが多く、安価で容易に真の乱数を生成する乱数生成器は現在でも渇望されているのではないだろうか。

上述の背景により、本研究は、身近な自然現象である「音」というものを利用して真の乱数、及びそれに準ずる性質を持った良質な乱数を生成する乱数生成器の実現を目指し、その実現のためにまずは音声そのもの及びそれに準ずるビット列の乱数的性質について評価するものである。

ただし、乱数はその性質上、絶対の評価を下すことができず、評価基準も多々あり、曖昧である。そのため本研究では、その乱数性評価に NIST(米国標準技術研究所)より発表された Special Publication 800-22[2](以下 NIST 乱数検定)を使用して統計的評価を行う。

# 2. 本研究で扱う音声ファイルの形式及び性質

## 2.1. ファイル形式

標準的なコンピュータにおいて、音声は一般的に「WAVE」という形式で表現されることが広く普及しており、その実態は Pulse Coding Modulation (PCM：パルス符号変調)によって標本化されたものである。これは、「音声」と認識された波形を、標本化、量子化し、得られた信号の大きさを二進の数値データとして表現したものである。この WAVE というフォーマットを論ずるときに必要となるものが、サンプリング周波数と、ビット数である。サンプリング周波数というのは、音声の波形を一秒間に何回標本化するかを示した数字である。例えば市販されている一般的な音楽 CD は、サンプリング周波数が 44.1kHz に設定されており、つまりこれは一秒間に 44100 回標本化を行っているということである。ビット数というのは、サンプリングの際に、ひとつのデータを何ビットで量子化するかを示した数値であり、一般的な音楽 CD は 16 ビットである。これは波形の振幅を 65536 段階で表現するということである。その性質を考えればわかるように、サンプリング周波数を高くすれば高くするほどその音声ファイルは、高周波の音まで表現することができるようになり、ビット数を高くすれば高くするほど、よりダイナミックな音を表現することができるようになる。

前述したように、サンプリング周波数を高くすると、より高周波の音を表現することができるのだが、それは裏を返せばサンプリング周波数によって表現できる周波数は制限されているということである。例えば 100Hz の音声波形をサンプリング周波数が 10Hz、つまり一秒間に 10 回標本化して表現したとする。一秒間に 100 回振動する波を一秒間に 10 回の標本化で表現できるはずがないのは議論するまでもないだろう。サンプリング周波数で表現することができる最高の周波数をナイキスト周波数と呼び、サンプリング周波数を f、ナイキスト周波数を $f_n$ とすると $f_n=f/2$ によって表される。例えば一般的な音楽 CD であれば、サンプリング周波数が 44.1kHz なので、ナイキスト周波数は 22.05kHz である。これは人間の可聴域よりも高い周波数であり、音楽 CD のサンプリング周波数はそれを考慮した値である。

## 2.2 折り返し雑音

折り返し雑音はサンプリング周波数を変換する際に発生する雑音である。

サンプリング周波数を $f_1$ から $f_2(f_1<f_2)$へと変換、つまりアップサンプリングする時の基本処理は、

各サンプルの後に $f_2/f_1$-1 個の 0 を追加することであるが、単純に考えると、0 とそれ以外の値が交互に現れる為、ただアップサンプリングした場合、波形はギザギザのままである。これは折り返し雑音を拾っているということである。

逆に、サンプリング周波数を $f_1$ から $f_2$($f_2$<$f_1$)へと変換する際、つまりダウンサンプリングする時も、変換前よりも変換後の方がナイキスト周波数は低くなるため、表現しきれなかった高周波の信号が折り返し雑音となって現れる。図 1 を見ると、2 つの正弦曲線が標本化によって、全く同じ標本列を生成しうることがわかるだろう。

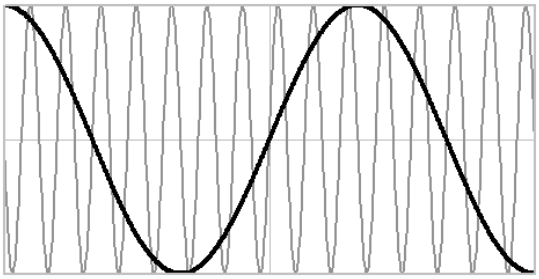

図 1　ダウンサンプリング時の折り返し雑音発生の原理

本研究では、音声のファイルそのものの他に、このダウンサンプリングの際の折り返し雑音を抽出したものについても、検討を行っていく。

### 2.3. WAVE ファイルの構造

WAVE ファイルは一般的に 44 バイトのヘッダを持ち、それぞれ、サンプリング周波数や、ビット数等、基本的な情報が記述されている。これらはもちろん、実際の音声のデータでは無いので、本研究ではこの部分をカットし、データ部だけを評価するものとする。

データ部の構造は、チャンネル数、ビット数によって異なり、例えば 16 ビットのステレオ(2ch)信号ではデータは 1 ブロック 4 バイトで表現される。1・2 バイト目、つまり最初の 16 ビットが右チャンネルのデータであり、次の 16 ビットが左チャンネルのデータとなる。

今回はモノラルについては考えず、ステレオの信号に対して検討を行っていく。

## 3. 評価方法

### 3.1. 準備

今回検定することとなる標本系列は 2 つである。一つ目は WAVE ファイルそのもの。二つ目は WAVE ファイルをダウンサンプリングしたものから抽出した折り返し雑音である。前者は、サンプリング周波数 44.1kHz、ビット数 16 ビット、2ch。後者はサンプリング周波数 4000Hz、ビット数 16 ビット、2ch である。

折り返し雑音は、用意した WAVE ファイルをハイパスフィルターに通し、変換後のナイキスト周波数以下の周波数をカットした後、ダウンサンプリングを行う。これで折り返し雑音のみの音声ファイルができる。

用意された二つの音声ファイルから、ヘッダ部である、冒頭の 44 バイトを削除し、それを標本系列とする。

### 3.2. 乱数検定

本研究にて使用する NIST 乱数検定について説明する。NIST 乱数検定では、表 1 に示す 15 種類の検定法が示されており、それら全ては、0 と 1 からなる乱数列を対象としている。

各検定は標準正規分布またはカイ 2 乗分布に基づいて行われ、それにより p-value と呼ばれる値が算出される。p-value とは、真の乱数生成器が検定を行っている系列よりもランダムでない系列を生成する確率と解釈でき、個々の検定に対して、p-value<0.01 の時に良い乱数列ではない

と判断され、それ以外の系列が良い乱数列であるとされる。

標準正規分布に基づいて検定が行われる場合、p-value は以下の関数 erfc (complementary error function)を用いて計算される。

$$\mathbf{erfc}\ (z) = \int_z^{\infty} \frac{2}{\sqrt{\pi}} e^{-x^2} dx \qquad (1)$$

カイ 2 乗分布に基づいて検定が行われる場合、以下の関数 igamc (incomplete gamma function)を用いて p-value が計算される。

$$\mathbf{igamc}\ (a, z) = \frac{1}{\Gamma(a)} \int_z^{\infty} e^{-t} t^{a-1} dt \qquad (2)$$

$$\Gamma(a) = \int_0^{\infty} e^{-t} t^{a-1} dt \qquad (3)$$

表 1　NIST 乱数検定に含まれる検定法一覧

| | 検定名 |
|---|---|
| 1 | 一次元度数検定 |
| 2 | ブロック単位の頻度検定 |
| 3 | 累積和検定 |
| 4 | 連の検定 |
| 5 | ブロック単位の最長列検定 |
| 6 | 2 値行列ランク検定 |
| 7 | 離散フーリエ変換検定 |
| 8 | 重なりのないテンプレート検定 |
| 9 | 重なりのあるテンプレート検定 |
| 10 | Maurer のユニバーサル統計検定 |
| 11 | 近似エントロピー検定 |
| 12 | ランダム偏差検定 |
| 13 | 種々のランダム偏差検定 |
| 14 | 系列検定 |
| 15 | 線形複雑度検定 |

各検定では、複数の標本系列(NIST では 1000 程度を推奨)に対し検定を行い、
(1)p-value が 0.01 以上になる比率（proportion）
(2)p-value の一様性(Uniformity)
によって乱数列の評価を行う。

(1)については、p-value が 0.01 以上になる計列数は正規分布 $N(\mu, \sigma 2)$に従うと考え、$\mu \pm 3\sigma$ 以内に収まれば良いとする。

(2)については、区間[0,1)を 10 分割し、各区間に属する p-value の個数が均等であるかどうかをカイ 2 乗分布によって検定する。具体的には、乱数列の個数が m 個の場合、1≦i≦10 について、$F_i$ を区間[(i-1)/10,i/10)に属する p-value の個数とするとき、次式を計算し、

$$\chi^2 = \sum_{i=1}^{10} \frac{(F_i - m/10)^2}{m/10} \qquad (4)$$

igamc(9/2, $\chi^2/2$)を求め、それが 0.0001 以上となった場合、乱数生成器であると判定される。

## 3.3. 実験結果

本研究では、擬似乱数生成系の検定方法に関する調査　調査報告書[3]　に習い、10 万ビット 500 本、10 万ビット 1000 本、100 万ビット 500 本に対して、それぞれ乱数検定を行った。検定結果は表 2、3 に示すとおりである。
なお、表の中の数値は、P においては、p-value が 0.01 以上になる系列数を、U においては igamc(9/2, $\chi^2/2$)の値を、「x」は複数ある検定結果において、ひとつも合格しなかったことを、「-」は系列長が推奨値に満たないため、検定を行わなかったことを表している。
網掛けになっているものはミニマムセット[4]に含まれる検定方法である。

500 本の標本系列数で検定した場合の(1)proportion の合格基準は 488 以上、1000 本の標本系列数の場合は 980 以上となっている。

表 2、3 を見ると、音声ファイル及び、それから生成された折り返し雑音を抽出したもの、いずれに対しても、NIST 乱数検定においてほとんどの検定について、「良い乱数性を持った乱数列で

はない」と判断されていることがわかる。

しかしまた、一部の検定項目を見ると、すべてのビット列に対して「良い乱数性を持った乱数列ではない」と判断されているわけではないこともわかる。

表 2　音声データの検定結果

P：合格比率(合格数)　U：一様性
A. 10 万ビット×500 本
B. 10 万ビット×1000 本
C. 100 万ビット×500 本
x:FAILURE

| | A | | B | | C | |
|---|---|---|---|---|---|---|
| | P | U | P | U | P | U |
| 1 | 445 | 0 | 856 | 0 | 500 | 0 |
| 2 | 374 | 0 | 684 | 0 | 0 | 0 |
| 3 | x | x | x | x | | x |
| 4 | 1 | 0 | 1 | 0 | 0 | 0 |
| 5 | 323 | 0 | 534 | 0 | 499 | 0 |
| 6 | 98 | 0 | 165 | 0 | 0 | 0 |
| 7 | 392 | 0 | 726 | 0 | 99 | 0 |
| 8 | − | − | − | − | x | x |
| 9 | − | − | − | − | 0 | 0 |
| 10 | − | − | − | − | 0 | 0 |
| 11 | 242 | 0 | 399 | 0 | 0 | 0 |
| 12 | − | − | − | − | | x |
| 13 | − | − | − | − | | x |
| 14 | | | | | | x |
| 15 | − | − | − | − | 500 | 0 |

一次元度数検定の結果を見てみると、音声データの 100 万ビット×1000 本の結果を除いて、8 割から 9 割のビット列が p-value の値について合格と判定されながら、総合的に見ると(1)proportion の合格比率に僅かに届いていないようである。(2)Uniformity については、すべての結果において不合格である。

電子政府情報セキュリティ技術開発事業によって導出されたミニマムセット[3]については、折り返し雑音の 100 万ビット×500 本における線形複雑度が唯一合格しているのみとなっている。

表 3　折り返し雑音の検定結果

P：合格比率(合格数)　U：一様性
A. 10 万ビット×500 本　　B. 10 万ビット×1000 本
C. 100 万ビット×500 本
x:FAILURE

| | A | | B | | C | |
|---|---|---|---|---|---|---|
| | P | U | P | U | P | U |
| 1 | 462 | 0 | 925 | 0 | 443 | 0 |
| 2 | 423 | 0 | 843 | 0 | 153 | 0 |
| 3 | x | x | x | x | x | x |
| 4 | 451 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 59 | 0 |
| 6 | 0 | 0 | 680 | 0 | 110 | 0 |
| 7 | 350 | 0 | 699 | 0 | 13 | 0 |
| 8 | − | − | − | − | x | x |
| 9 | − | − | − | − | 0 | 0 |
| 10 | − | − | − | − | 14 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | − | − | − | − | x | x |
| 13 | − | − | − | − | | |
| 14 | x | x | x | x | x | x |
| 15 | − | − | − | − | 489 | 0.269 |

音声ファイルと、そこから生成した折り返し雑音について比較してみると、一部を除いたすべての結果について、元の音声よりも、折り返し雑音の方が良い結果になっているのがわかる。ただし、ブロック単位の最長連検定や、連の検定については元の音声ファイルの方が良い結果になるようである。

これらの結果から判断すると、音声データも、折り返し雑音のデータも、NIST 乱数検定においては「良い乱数性を持った乱数生成器ではない」と判断されている。しかし、データの部分部分によっては、「良い乱数列である」と判断されているものもあり、音声や、それから発生した折り返

し雑音が乱数としての使用に耐えないか、については更なる検証を進めなければ判断できないだろう。

ただし、NIST 乱数検定によって示された検定項目や、合否基準は絶対のものではなく、すべての検定に合格したときに限り、その標本系列がよい乱数列であるという評価を下すことは正しいとは言えない。例えば、真のランダム性を持った乱数列を NIST 乱数検定に通した場合、全ての検定項目において合格する確立は、二項分布を用いると約 54%、正規分布を用いて求めた場合は 78%となる[3]。

## 4. まとめ

様々な音を混ぜて作り上げた音声のデータは、ある程度の乱数性を持ちながら、それ自体は人間の「意志」やある一定の「規則性」が介在する可能性のある、極めて評価の難しい標本系列である。今回の結果を踏まえて、どのような音声データだと乱数性が高いのか、について調べてみる必要がある。

また、サンプリング周波数の数値ごと、ビット数の数値ごと、折り返し雑音を抽出せずに生成元の音声に混ぜた場合、音声データとその折り返し雑音との差分データなど多数の調べるべき標本系列がある。これらについても適宜評価を行い、音声を元に乱数生成する場合、どのような演算を行うのが最も効率的で高品質なのか、音声をどのように利用するべきかなどについてじっくり検討を行っていく必要がある。

## 文 献


[1] M. Matsumoto and T. Nishimura, “Mersenne Twister: A 623-dimensionally equi distributed uniform pseudorandom number generator,” ACM Trans. on Modeling and Computer Simulation Vol. 8, No. 1, Jan., pp.3-30, 1998.

[2] Andrew Rukhin, et al., “A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications,” NIST Special Publication 800-22 Ver.1a, April, 2010.

[3] 廣瀬勝一,京都大学,擬似乱数生成系の検定方法に関する調査　調査報告書,2004.

[4] 電子政府情報セキュリティ技術開発事業「擬似乱数検証ツールの調査開発」調査報告書、情報処理振興事業協会セキュリティセンター、2003 年 2 月.