平成 20 年度

プロジェクト研究報告書

バーナム暗号に用いる
疑似乱数生成アルゴリズムの比較

1090302　　青野 裕太

指導教員　　福本 昌弘

2009 年 3 月 5 日

高知工科大学 情報システム工学科

要 旨

バーナム暗号に用いる
疑似乱数生成アルゴリズムの比較

青野 裕太

ネットワークに接続された環境では，個人情報などの他者に知られると自身の不利益につながるような情報は盗聴の危険に曝されている．特にインターネットを利用する場合，自身の管理できない経路を通り情報を送るため危険性は高まる．そのため暗号方式を用いて情報を暗号化することで安全性な通信が提供されている．暗号方式の一つにバーナム暗号がある．バーナム暗号は，鍵の秘匿により解読不可能が証明されている強力な暗号であるが，真性乱数を一度きりの使い捨て鍵として用いており，暗号化を行う度に，真性乱数を生成する必要がある．真性乱数の生成はソフトウェアだけの環境では不可能なため利便性に問題がある．利便性を重視する場合，疑似乱数の利用が考えられるが，すべての疑似乱数生成アルゴリズムがランダム性の高い乱数を生成できるとは限らない．そこで，本研究ではバーナム暗号に利用可能な疑似乱数を生成できる生成アルゴリズムをバランス性の観点から比較を行い，すべての生成アルゴリズムが高いバランス性を持つことを確認している．

キーワード　バーナム暗号，疑似乱数，バランス性

目次

目次

図目次

表目次

第 1 章

序章

1.1 背景と目的

ネットワークに接続された環境では，個人情報などの第三者に知られると自身の不利益につながるような情報に限らず，さまざまな情報が盗聴の危険に曝されている．特にインターネットを介しての通信を行う場合，自身の管理の及ばない通信経路を通ることになり，送信者と受信者の間で第三者に盗聴される危険性が高くなっている．そのため，情報を守る手段として AES[*1]や RSA[*2]などの暗号方式を用いて安全な通信を提供している [1]．この暗号方式の一つにバーナム暗号がある．バーナム暗号は真性乱数を使い捨ての鍵と用いて，平文と鍵の排他的論理和を取ることで暗号化する方式であり．鍵の秘匿により解読不可能が実現されている [2]．強固な暗号であるが，実際に利用する場合には，鍵に関して問題がある．鍵となる真性乱数は，熱雑音などの物理乱数を元に平滑化処理を行い生成されるため，ソフトウェアのみの環境では生成できないため利便性に問題がある．そこで利便性を重視するとソフトウェアのみでも生成可能な疑似乱数を利用することが考えられるが，すべての疑似乱数生成アルゴリズムが真性乱数のようなランダム性を持つ乱数を生成できるとは限らない．そこで，本研究では，バーナム暗号へ適用が可能なランダム性を持つ疑似乱数を得るために，疑似乱数のバランス性について既存の疑似乱数生成アルゴリズムの比較を行い，どの疑似乱数生成アルゴリズムが適しているかを確認することを目的とする．

[*1] Advanced Encryption Standard

[*2] 開発者 Rivest,Shamir,Adleman の 3 人の頭文字

1.2　本論文の構成

本論文では，バーナム暗号に利用可能な乱数を生成できる疑似乱数生成アルゴリズムについて乱数の持つバランス性の観点から比較を行う．

第 2 章では，バーナム暗号の特徴や処理の流れを説明し，バーナム暗号の欠点について述べる．また，暗号鍵として用いる乱数について真性乱数と疑似乱数に分けて説明する．

第 3 章では，バランス性の観点から疑似乱数生成アルゴリズムとそれらのアルゴリズムにより生成された疑似乱数をバーナム暗号に用いた場合について比較と行いその結果から考察する．

第 4 章で本論文の結論を述べる．

第 2 章

暗号方式

2.1　はじめに

本章では，ストリーム暗号について説明し，この暗号方式の一つであるバーナム暗号について説明する．

2.2　暗号方式

共通鍵暗号方式は，1 つの鍵を使って暗号化と復号を行う暗号方式である．

ストリーム暗号は，同じ鍵を使って平文を 1 ビットまたは 1 バイト単位で逐次的に暗号化を行う方式である．

ブロック暗号は，同じ鍵を使って平文を一定のビットまとまりをブロックとし，ブロックごとに暗号化を行う方式である．ブロック暗号の特徴はブロックごとに換字暗号や転置暗号といった異なる暗号化を組み合わせ処理を行う．

2.2.1　バーナム暗号

バーナム暗号は，図 2.1 のような構成で暗号通信を行う．この方式は，2 値乱数発生器で生成された乱数を暗号鍵とし，平文と鍵の排他的論理和演算を行い暗号化する．復号するためには，暗号化に用いた鍵と同じものが必要となるため，この鍵を秘密裏に共有しておく必要がある．復号は，暗号文と暗号鍵の排他的論理和をとることで復号する．

バーナム暗号はストリーム暗号に属し，図 2.2 平文と鍵の各ビットごとに排他的論理和を

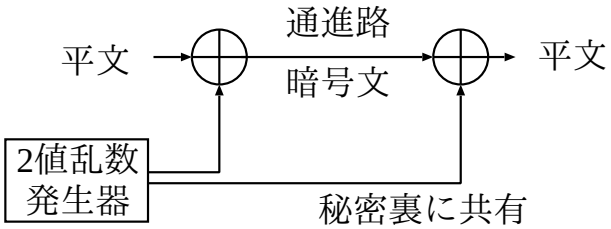

図 2.1　バーナム暗号を用いた通信の構成

取り暗号化を行うシンプルな方式である．しかし，鍵の秘匿により解読不可能となる非常に強力な暗号方式の一つである．強力である要因の一つに，鍵として真性乱数を用い，その鍵を 1 度限りの使い捨てとしている点が挙げられる．しかし，バーナム暗号の鍵長は平文の長さ以上必要なため，暗号化するごとに長い乱数列を生成する必要がある．真性乱数はどのような環境でも容易に生成できるものではないため実用性が低く一般には使用されていない．そこで鍵生成の問題を真性乱数の代わりに疑似乱数を用いることで実用性も持たせられると考えられる．

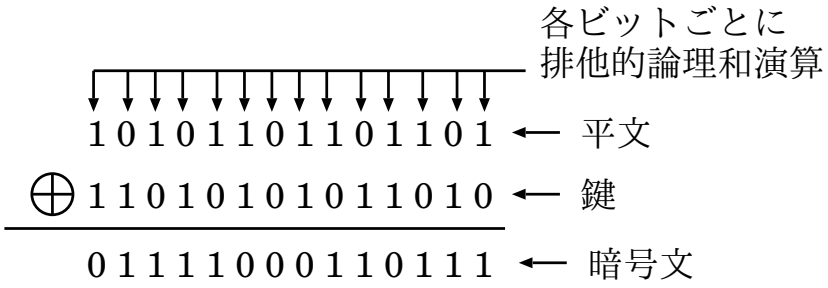

図 2.2　バーナム暗号の暗号化

第 3 章

暗号に用いられる乱数と安全性

3.1　はじめに

本章では，真性乱数の特徴と問題点を述べる．また，疑似乱数についても特徴と問題点を述べ，4 つの疑似乱数生成アルゴリズムを示す．最後に乱数を暗号へ用いる場合に必要な安全性について述べる．

3.2　真性乱数

真性乱数は，非再現性や予測不可能性の性質を持った乱数であり，0 と 1 の発生確率がおのおの 1/2，各ビットは他の部分と独立しているという特徴がある．真性乱数は，疑似乱数の初期値や暗号化の鍵として用いられており，同じ乱数列を再現できない点や次に発生するビットを予測できない点では，安全性に優れているが，生成方法が困難である問題がある．真性乱数の生成方法は図 3.1 のように熱雑音などの予測や再現が不可能な自然現象から物理乱数を得た上で平滑化処理を行い生成される．物理乱数は統計的な性質に関して 0,1 の出現バランスが取れていないため，平滑化処理を行い出現バランスを整える必要がある．暗号に用いる場合，非再現性や予測不可能性の性質は高い安全性につながるが，ソフトウェアのみで容易に生成できないため，環境に依存してしまい利便性が低いという問題がある．

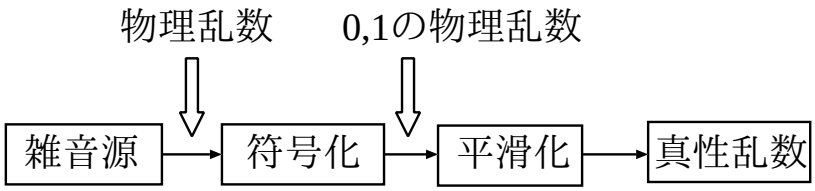

図 3.1　真性乱数の生成方法

3.3　疑似乱数

疑似乱数は，決定性アルゴリズムにより生成された乱数である．真性乱数と比較してソフトウェアでも容易に生成できる利点があるため，利便性に関しては真性乱数よりも高く利用環境にあまり影響しない．しかし，暗号化の鍵などに用いる場合は，いくつかの問題がある．疑似乱数生成器により生成された乱数には周期があることであり，周期が短いと出力パターンがすぐに見付かり次に生成される乱数が予測されてしまうことがある．次に，疑似乱数はシードと呼ばれる初期値を元に乱数を生成しており，同じシードを用いると同じ乱数が生成される．そのため，同じシードは使えないため，生成する度に新しいシードと交換する必要がある．疑似乱数生成アルゴリズムにはいくつか種類があり，生成速度や周期などの特徴に違いがある．

暗号には，0 または 1 からなる 2 値乱数を用いる．今回利用する生成アルゴリズムは整数値で乱数を生成するため，0 または 1 を生成する工夫が必要である．

3.3.1　線形合同法

線形合同法は，漸化式

$$X_{n+1} = (A \times X_n + B) \bmod M \tag{3.1}$$

により乱数を生成する．式中の A，B，M は定数であり，M の値が最大周期となる．また，A，B，M の間には以下の関係がある．

- B と M が互いに素である
- $A-1$ は M の全ての素因数で割りきれる
- M が 4 の倍数である場合，$A-1$ も 4 の倍数である

線形合同法により生成された疑似乱数の下位ビットはランダム性が低いという特徴があるため，最上位ビットを出力とする．

3.3.2　線形フィードバックシフトレジスタ

線形フィードバックレジスタ (LFSR) は，2 値系列の乱数列を高速に生成できる疑似乱数生成アルゴリズムである．図 3.2 のような構成により乱数を生成する．a_L は LFSR の有線

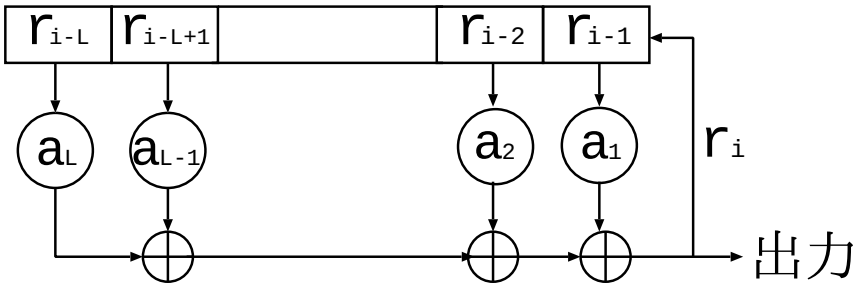

図 3.2　線形フィードバックシフトレジスタ

状態を示しており，$a_L = 1$ のときは線が結ばれていることを表し，$a_L = 0$ のときは切断されていることを示す．L 個のレジスタからなる LFSR は 2^L 個の状態が存在し，全レジスタの値が 0 である状態を除いた $2^L - 1$ 個の状態が全て現れるとき，出力系列は最も長い周期 $2^L - 1$ を持つ．

3.3.3　メルセンヌ・ツイスタ

メルセンヌ・ツイスタの特徴は長周期性と生成速度である．メルセンヌ・ツイスタは，漸化式

$$X_{k+n} = X_{k+m} \oplus (x_k^u | x_{k+1}^l)A, (k = 0, 1, 2, ...) \tag{3.2}$$

により乱数を生成する．$(x_k^u|x_{k+1}^l)$ は x_k の上位 $(w-r)$ ビットと x_{k+1} の下位 r ビットを並べて得られる $GF(2)^w$ の元であり，こうして得られる数列の最大周期は $2^{nw-r}-1$ となる．n,m は定数であり，$n > m > 0$ の関係がある．また，A は

$$A = \begin{bmatrix} & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ a_{w-1} & a_{w-2} & \cdots & \cdots & a_0 \end{bmatrix} \tag{3.3}$$

のような行列である．

以上の漸化式や行列により整数の乱数が生成される．メルセンヌ・ツイスタには下位ビットがランダムでないなどの問題は見付かっていないが，今回は最上位ビットを出力とする．

3.3.4　XorShift

XorShift は，排他的論理和演算とビットシフト演算のみで疑似乱数を生成するアルゴリズムである．周期は 2^k-1, $(k = 32, 64, 96, 128, 160, 192)$ となる．

周期 $2^{128}-1$ の場合の XorShift の演算は，4 つの初期値 x, y, z, w を用いて，x を左に 11bit シフトさせたものと x 自体の排他的論和をとり変数 t へ代入する．次に x に y を代入，y に z を代入，z に w を代入し値を保持する．最後に w 右に 19 ビットシフトさせたものと w 自体の排他的論理和をとり，t を右に 8 ビットシフトさせたものと t 自体の排他的論理和をとったのも排他的論理和をとる．XorShift も特に問題は指摘されていないが，最上位ビットを出力とする．

3.4　乱数の安全性

暗号に利用する乱数として求められる性質には，過去に生成された乱数列に規則性がなく，生成される乱数列を予測できない予測不可能性がある．真性乱数であれば，この性質を満たしているため暗号鍵として用いたとしても規則性が見付けられず，解読される危険性は

低くなる．しかし，疑似乱数の場合は決定性アルゴリズムにより乱数を生成しているため，この性質を満たすことは難しく，容易に大量の乱数を生成できないため実用的ではない．そこで，0,1 が等頻度で出現するバランス性，乱数列のあるビットが他の一部と独立している無相関性などの評価尺度を用いることで，予測不可能性を持つ乱数よりは比較的容易に生成できるようにして実用性を高めている．

安全性の評価尺度の一つとしてバランス性がある．バランス性は 0,1 の出現頻度が 1:1 となることが理想であり，真性乱数であればこの性質を持っているが，疑似乱数の場合生成アルゴリズムによっては偏りがある可能性がある．バランス性が低い乱数列は 0 もしくは 1 の出現頻度に偏りが生じ，乱数列とはいえなくなってしまう．

バランス性の他にも長周期性，無相関性，非線形性，線形複雑度の大きさなど性質がある．暗号学的に安全な乱数であればこれらしかし，これらの条件を全て満たす疑似乱数が必ずしも暗号学的に安全な乱数であるとは限らない．

第 4 章

疑似乱数生成アルゴリズムの比較

4.1 はじめに

本章では，暗号に利用する乱数の安全性とその要素の一つであるバランス性について説明し，各アルゴリズムにより生成さられた疑似乱数のバランス性の比較を行う．また，これらの疑似乱数を実際にバーナム暗号に適用した場合に暗号文のバランス性の比較を行う．最後に比較結果の考察を行う．

4.2 バランス性の比較方法

各アルゴリズムを用いて生成した 1024bit の疑似乱数 2000 組を対象にバランス性について比較を行う．まず，疑似乱数のみを対象に比較を行う．図 4.1 のように 1024bit の疑似乱数について 1 の出現回数を求め，理論値との最大誤差と平均誤差を求める．ここでの理論値は 1024bit の半分となる 512bit である．

また，図 4.2 のように 1024bit の 1bit 目から 1024bit 目までの各 bit について，1 の出現回数を求める．こちらの場合は各 bit に出現する 1 の割合が 1/2 に近いほどバランス性が高いと考えれる．ここでの理論値は，2000 組を対象としているため，1000 個となる．

次に各アルゴリズムにより生成された疑似乱数をバーナム暗号に適用した場合について，疑似乱数のみの場合と同様に 0,1 の出現回数を求め，理論値との最大誤差と平均誤差を求める．バーナム暗号に適用する場合，平文を考慮する必要があるため，平文を固定した場合と固定しないで毎回変更する場合とに分けて比較を行う．

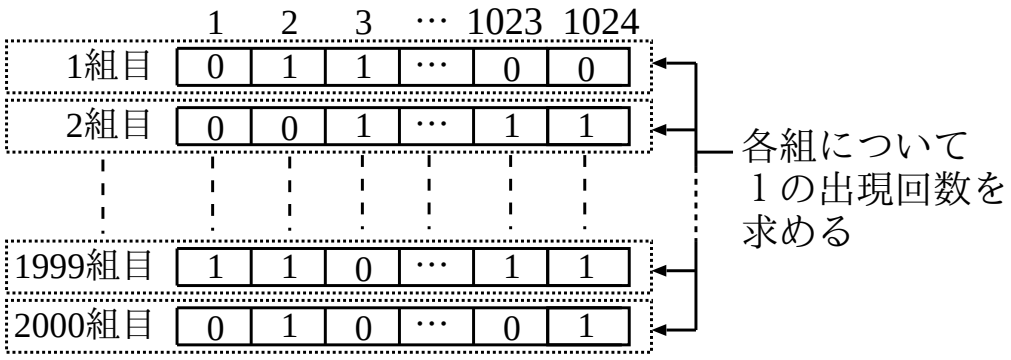

図 4.1　各組の出現回数

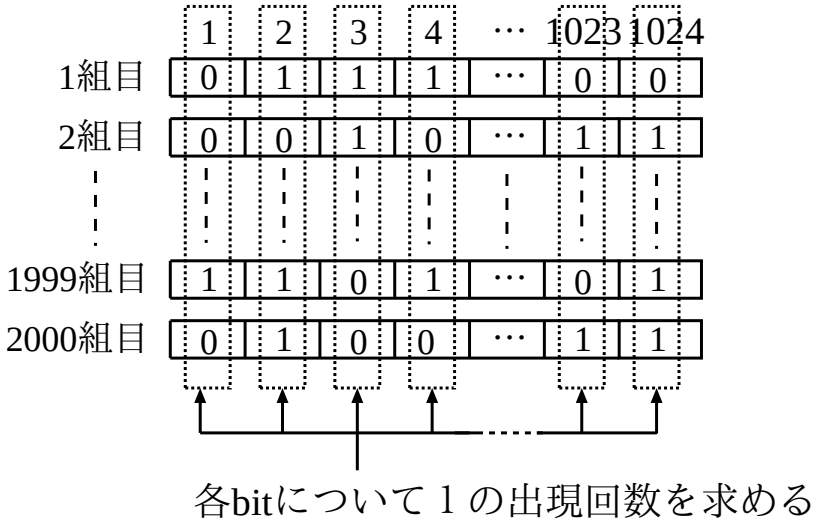

図 4.2　各ビットの出現回数

4.3　バランス性の比較結果

各アルゴリズムにより生成された 2000 組の疑似乱数 1024bit の理論値との最大誤差と平均誤差を表 4.1 に示す．最大誤差は大きいもので 60 個の差があるため，実際には 0 と 1 の出現比率は 452 : 562 となる．しかし平均的な誤差の観点ではどのアルゴリズムも誤差が小さいためバランス性は高いと考えられる．

次に，各アルゴリズムにより生成された 2000 組の 1024bit の各 bit についての 1 が出現

4.3　バランス性の比較結果

表 4.1　最大誤差と平均誤差

アルゴリズム	最大誤差	平均誤差
線形合同法	43	0.7435
線形フィードバックシフトレジスタ	53	0.7110
メルセンヌ・ツイスタ	60	0.1775
XorShift	56	0.8170

する回数を表 4.2 示す.

表 4.2　各 bit の 1 の出現回数

アルゴリズム	最大出現数	最小出現数	平均出現数
線形合同法	1047	954	1001
線形フィードバックシフトレジスタ	1069	907	1001
メルセンヌ・ツイスタ	1097	927	999
XorShift	1098	922	998

平均出現数は各アルゴリズムともに理論値となる 1000 個と誤差がほとんどなく，どのアルゴリズムもバランスがとれていると考えれる.

平文を固定しバーナム暗号に適応した場合の最大誤差と平均誤差を表 4.3 に示す．最大誤

表 4.3　平文固定の場合の最大誤差と平均誤差

アルゴリズム	最大誤差	平均誤差
線形合同法	70	0.0235
線形フィードバックシフトレジスタ	53	0.4930
メルセンヌ・ツイスタ	58	0.2415
XorShift	54	0.0250

差は大きいもので 70 個の差があるため，実際には 0 と 1 の出現比率は 442 : 572 となる．しかし，平均的な誤差の観点ではどのアルゴリズムも誤差が小さいためバランス性は高いと考えられる．

平文を毎回変更しバーナム暗号に適応した場合の最大誤差と平均誤差を表 4.4 に示す．最

表 4.4　平文変更の場合の最大誤差と平均誤差

アルゴリズム	最大誤差	平均誤差
線形合同法	56	0.4265
線形フィードバックシフトレジスタ	55	0.3090
メルセンヌ・ツイスタ	55	0.0065
XorShift	61	0.0955

大誤差は大きいもので 61 個の差があるため，実際には 0 と 1 の出現比率は 451 : 573 となる．しかし，平均的な誤差の観点ではどのアルゴリズムも誤差が小さいためバランス性は高いと考えられる．

次に，平文を固定して暗号化を行った 2000 組の 1024bit の各 bit についての 1 が出現する回数を表 4.5 に示す．平均出現数は各アルゴリズムともに理論値となる 1000 個と誤差が

表 4.5　平文固定：各 bit の 1 の出現回数

アルゴリズム	最大出現数	最小出現数	平均出現数
線形合同法	1046	953	999
線形フィードバックシフトレジスタ	1093	931	1000
メルセンヌ・ツイスタ	1065	921	999
XorShift	1098	922	1000

ほとんどなく，どのアルゴリズムもバランスがとれていると考えれる．

次に，平文を変更して暗号化を行った 2000 組の 1024bit の各 bit についての 1 が出現す

る回数を表 4.6 に示す.

表 4.6　平文変更：各 bit の 1 の出現回数

アルゴリズム	最大出現数	最小出現数	平均出現数
線形合同法	1069	914	1001
線形フィードバックシフトレジスタ	1085	927	999
メルセンヌ・ツイスタ	1078	923	1000
XorShift	1063	929	999

平均出現数は各アルゴリズムともに理論値となる 1000 個と誤差がほとんどなく，どのアルゴリズムもバランスがとれていると考えれる.

4.4　考察

どのアルゴリズムもバランス性の観点からは高いバランス性を持っていることが確認できる．また，バーナム暗号に適用した場合も同様に高いバランス性を持っており，暗号文に 0 または 1 の偏りは見付からない．このことからバランス性に関しては，どのアルゴリズムを使っても暗号文に偏りがでないと考えられる.

4.5　まとめ

本章では，暗号に利用する乱数の安全性とその要素の一つであるバランス性について説明し，各アルゴリズムにより生成さられた疑似乱数のバランス性の比較を行った．また，これらの疑似乱数を実際にバーナム暗号に適用した場合に暗号文のバランス性の比較を行った．比較の結果，どのアルゴリズムを用いてもバランス性の高い疑似乱数が生成され，それらの乱数をバーナム暗号に適用した場合も暗号文には偏りが出ないことが分かった.

第 5 章

結論

5.1　本研究のまとめ

真性乱数を鍵として暗号化を行うバーナム暗号がある．しかし，真性乱数の生成は容易に行えないため利便性に問題がある．そこで利便性を考慮すると疑似乱数の利用が考えられるが，すべての疑似乱数生成アルゴリズムが真性乱数のようなランダム性を持つ乱数を生成できるとは限らないという問題がある．そこで既存の疑似乱数生成アルゴリズムの中から線形合同法，線形フィードバックシフトレジスタ，メルセンヌ・ツイスタ，XorShift の 4 つをバランス性の観点から比較を行った．各アルゴリズムにより生成された疑似乱数には大きな偏りは見られなかった．また，これらの疑似乱数をバーナム暗号に適用して，平文を固定した場合と毎回変更した場合について，バランス性を調べたが，偏りは見付からなかった．バランス性の観点からはどのアルゴリズムもバーナム暗号に利用できると考えられることが分かった．

5.2　今後の課題

バランス性の観点からはどのアルゴリズムも生成する疑似乱数に偏りが見られないため，0,1 の出現頻度が直接危険につながるとは考えづらい．しかし，安全性を重視するため，非線形性，線形複雑度の大きさ，無相関性などの他の性質も考慮する必要がある．

謝辞

高知工科大学情報システム工学科福本昌弘准教授に心より感謝し厚くお礼申し上げます．副査をして頂いた坂本明雄教授にも感謝致します．同研究室の大学院博士課程の佐伯幸郎様，福冨英次様，劉立剛様，大学院修士課程の金井宏一郎様にも感謝致します．ともに研究に取り組んだ 4 年生にも感謝します．

参考文献

[1] 岡本栄司，“暗号理論入門,” 共立出版株式会社，p.62，1993

[2] 若林宏，“図解入門よくわかる最新暗号技術の基本と仕組み,” 株式会社秀和システム，p.47，2005

[3] 大矢雅則，“情報数理入門,” 株式会社サイエンス社，1999

[4] 伏見正則，“乱数,” 財団法人東京大学出版会，1989

[5] 岡本龍明，山本博資，“現代暗号,” 産業図書株式会社，1997