1E-03

DESにおけるS-boxの初期配置に関する考察

山田　隆行†，小松　健太†
高知工業高等専門学校†

１．はじめに

ブロック暗号においては，非線形変換が安全性の要であり，このため，DES では，出力から逆関数（逆変換）を使って元のデータを簡単には計算できないようにs-Boxを用いて，単純な四則演算やシフトだけでは実現不可能な処理を行っている．AES 暗号は，仕様書において構成法を公開していることからその安全性について検証することができるが，DES については，構成法や初期配置について公開されていない[1]．

本論文では，DES 暗号の初期値の配置について，統計的観点から考察を行う．

２．DESについて

AES暗号において，SubBytes処理は一般的にS-box と呼ばれ，AES の仕様書で各ブロックに対して有限体GF(2^8)上における乗法逆元を計算し，その結果をアフィン変換するという構成法を公開している[1]．一方，DES では，S-box の作成法が公開されておらず[1]，また，初期配置の選び方についての言及もない．

これは，IBM が提案した Lucifer という方式を基に，米国国家安全保障局（NSA）が修正を加えたためといわれているが，NSA の修正点の１つにS-box の変更があり，NSA が S-box に暗号解読用の裏口を仕込んだのではないかという疑惑が長年持たれてきた[1]．

本論文では，S-box が注意深く設計されたものなのか出力の偏りが最小となるように求めた配置による S-box（偏り最小の S-box）を作成して，オリジナルの S-box（初期配置の S-box）との比較行う．そして，２つの S-box を使用した DES により作成された出力を乱数検定することで，オリジナルのS-boxランダム性能について考察する．

３．ランダム配置のS-box

初期配置の S-box は，オリジナルの 8 個の S-box をそのまま使用する．また，オリジナルのS-box の S_1 から S_8 まで各 8 テーブルに 0 から 15 が 4 個ずつ計 16×4＝64 個設定されていることから，初期値として 0 から 15 までの 4 個ずつ並べた数列に時刻情報を種にした Fisher-Yates shuffle アルゴリズムを用いてランダムにシャッフルしたランダム配置のS-Box を作成する．

４．排他的論理和表

j 番目の S-box を S_j，S_jに対する２つの入力を B_j, B_j^*，(B_j, B_j^*)を６ビットのペアとしたとき，

$B_j \oplus B_j^*$：(S_jの)入力の排他的論理和

$S_j(B_j) \oplus S_j(B_j^*)$：($S_j$の)出力排他的論理和

任意の $B_j' \in (Z_2)^6$ に対して，

$\delta(B_j')$：$B_j \oplus B_j^*$が B_j'となるペアの集合

と定義する．

このとき，$\delta(B_j')$は 2^6=64 個のペアを含み，

$\delta(B_j') = \{(B_j, B_j \oplus B_j') : B_j \in (Z_2)^6\}$

となることから，$\delta(B_j')$において，S_jの出力の排他的論理和を計算でき，さらに結果の表(排他的論理和表)を作成することができる[2]．

各 S-box の S_jについて，すべての 6bit 入力 2^6=64 に対する排他的論理和表を作成する．攻撃への耐性等を考慮すると各入力に対する排他的論理和表の各入力に対する各度数は均等($2^6/2^4$=4)となることが望ましいが，偏りがあることが知られている[2]．この偏りの度合いは，χ^2値として観測度数を n_i，期待度数を m_i とすると以下の式で計算することができる．

$$\chi^2 = \sum_{k=1}^{N} \frac{(n_i - m_i)^2}{m_i}$$

５．偏り最小のS-box

3 に示す方法でランダム配置の S-box を繰り返し作成しながら，00000 から 111111 までのすべての 6bit 入力 2^6=64 個における χ^2値を 4 に示す方法で計算し，χ^2値が前回より小さい場合には乱数の種を記録し，随時更新して行くことで，最小の χ^2 値を求める．このとき，繰り返し回数についてはpの一致推定量を$\hat{p}$とすると母比率pの95%信頼区間の推定式

$$\hat{p} - 1.96 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + 1.96 \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

から 384 回以上が必要ということが導き出される．

したがって，1000(>384)回繰り返せばオリジナルの S-box の χ^2値 χ_0 より小さな値を見つけ出せることが統計的に期待できる．そこで，ランダム配置の S-box の作成を 1000 回繰り返し，χ^2値の更新を繰り返して最後に記録されている χ^2値を求める最小 χ^2 値とし，このときの乱数の種から作成される配置の S-box を偏り最小の S-box とする．この一連の手順と操作を表１に示す．

Proposal of self-certification method for electronic documents using hashes
† Takayuki Yamada, Kenta Komatsu
National Institute of Technology, Kochi College

表１．最小χ^2値を求める手順と操作

手順	操作
1	ランダム配置： 3に示す方法により,ランダム配置のS-boxを作成
2	排他的論理和の表の作成： 2^6=64の入力に対する2^4=16の排他的論理和表を作成
3	統計検定： 排他的論理和表の各度数に対してχ^2検定を行う
4	χ^2値の更新： 前回より小さい場合,χ^2値を更新し乱数の種を記録
5	繰り返し： 上を1000回繰り返し最後に記録されるχ^2値を求める

６．乱数検定

出力系列のランダム性を乱数検定により評価する．統計的検定法の１つに米国商務省標準技術局（NIST）が公開している NIST Special Publication 800-22（SP800）がある．

SP800の検定では，乱数生成器の出力2元乱数系列を16種類の検定法計189個の試験で検定を行う．この結果，各検定ごとに検定で出力される統計量の正規分布もしくは，χ^2分布において，それよりも偏った統計量が発生する確率 *p-value* が得ることができ，

1．*p-value*の一様性

2．*p-value* > 0.01になる割合

をもとに乱数列の評価を行う[3]．

1では，[0,1]を10の区間に分割し，分割した区間ごとの頻度の一様性をχ^2検定により得られた*p-value*により，0.0001以上ならば乱数列は良い乱数であると判断する．

2では，標本の数をmとした時，0.01以上となる*p-value*の数の割合が

$$0.99 \pm 3\sqrt{\frac{0.99 \times 0.01}{m}}$$

の範囲に入っている場合には，乱数列は良い乱数であると判断する[3]．

この結果，著しく偏った乱数を発生するような乱数生成器を不適と判断することができる．

７．検証実験

5に示す方法により，初期配置の各S-box(S_1～S_8)における出力のχ^2値χ_0より小さなχ^2値を各々導出した．このときの乱数の種から生成したS-BoxをオリジナルのDESに組込み，乱数検定を行う．SP800による検定では平文として大量の入力データが必要となるため，インターネット上で公開されている青空文庫の書籍から表2に示す10種類の書籍のテキストデータを使用した．

乱数検定には，表3に示す鍵を10種類用意し，2つのS-boxについて平文と鍵を変えながら6に示す試験を各10×10=100回の行い，このうち189個の試験に合格した数の平均で最良のものを判断する．2つのS-boxの試験合格数の平均を表4に示す．この結果，初期配置のS-box，偏り最小のS-boxの順となった．S-boxの出力の偏りは，最小χ^2値を求めることで改善できたが，DESの出力の偏りは，ほとんど差がなく僅差となった．これは，S-box以外（初期転置IP，拡大関数E，転置 PC-1,PC-2）等の線形変換の初期配置も全体のランダム性に影響を与えているためではないかと考えられる．

また，SP800については，合否判定はかなり厳しいといわれており，少しでも規則性の兆しが見られれば帰無仮説を棄却するという特徴があることや，乱数系列のランダム性が棄却される結果が得られという報告[4]およびいくつかの検定法に不具合が指摘されている[5]ことなども要因として挙げられる．

表２　書籍データ

a.こころ	b.吾輩は猫である		c.ドグラマグラ
d.三四郎	e.それから		f.坊っちゃん
g.蓼食う虫	h.浮雲	i.破戒	j.金色夜叉

表３　DESの鍵

鍵	16進数	鍵	16進数
1	0x33C7ED2B4D8115A0	6	0x6177F89587F2CE1C
2	0xC9317C7D6CE595FE	7	0x959F57B4D5920510
3	0xAB89694A79A94593	8	0xB160CF765B3B32A6
4	0xFA5538D09693F511	9	0xFEA9B7C188A02A26
5	0xAFF67A7AAECCEAD3	10	0xBEA585E610FABFFB

表４　試験合格数の平均

S-boxの作成方式	試験合格数の平均
初期配置のS-box	137.39
偏り最小のS-box	137.06

８．まとめ

本論文では，構成法が示されていないDESのS-boxについて，偏り最小のS-boxを作成して初期配置のS-boxと比較することで考察した．

検証結果は，S-boxを注意深く設計したという事実を明確に裏付けるには至らなかった．

今後，線形変換についても検討を考えている．

参考文献

[1]IPUSIRON,暗号技術のすべて,翔泳社,2017
[2]Douglas R.Stinson,Cryptography:theory and practice,Chapman&Hall/CRC,2006
[3]IPA/ISEC,擬似乱数検証ツールの調査開発調査報告書,2003
[4]濱野健二,NISTの乱数検定に含まれる最長連検定の修正,信学技報,107(44),pp17-21,2007
[5]IPA/NICT,CRYPTREC Report,2004