

共通鍵ブロック暗号 CIPHERUNICORN-A の安全性に関する詳細調査報告

概要

本資料は、共通鍵ブロック暗号 CIPHERUNICORN-A の安全性に関する詳細報告書である。本暗号を、乱数検定、差分攻撃、線形攻撃、高階差分攻撃、補間攻撃、鍵衝突攻撃、それぞれについて詳細に検討した結果、CIPHERUNICORN-A の安全性について解読に結び付くような深刻な問題点は、差分攻撃法に対する若干の不安を除き、導かれなかった。各攻撃法に対する安全性に関する結果は、以下の様にまとめられる。

乱数検定：

4 段以上で、今回用意した全ての検定項目で乱数と区別できない。

差分攻撃：

提案者が自己評価書において、差分確率を求めるための単純化として用いた暗号は、攻撃者に対して最も有利な仮定を与えた場合、16 段でも解読可能であり、この単純化された暗号の強度を根拠にした CIPHERUNICORN-A についても、差分攻撃に対する安全が十分示されているとは言えない。

線形攻撃：

自己評価書の記述にやや疑問点があるものの、線形攻撃に対する安全性に問題はない。

高階差分攻撃：

高階差分攻撃に対する問題点はない。

補間攻撃：

補間攻撃に対する問題点はない。

鍵衝突攻撃：

自己報告書の記述にやや難があるものの、暗号そのものの安全性については問題ない。

1 乱数検定

暗号の安全性評価法の一つとして、与えられた入力に対する暗号の出力を疑似乱数として見た場合に、その列が、真の乱数列との区別可能かどうかによって、その暗号が安全であるかどうかを判断するという報告例がいくつかみられる。そもそも、現代共通鍵ブロック暗号の代表格である差分攻撃法 / 線形攻撃法にしても、ある一定段数までに、入力差分と出力差分との関係、あるいは入出力ビットの線形式の値を調べることにより、それぞれ真の乱数とは著しく性質が異なることを利用して、最終段の鍵の値を推定する方法であるから、この「真の乱数列と区別可能かどうか」という基準は、暗号の安全性に密接に結び付いていると見てよく、この基準をクリアしない暗号は、解読される危険があると考えられる。その意味から、提案者によって提出された安全性評価書に見られるようなビットアバランシュ性等の初等統計量調査による調査は、十分に意味があると思われる。

さて、安全性評価書によれば提案者らが行なった検定項目は次の 4 項目であるとされている。

- アバランシュ
- ビットバランス
- 入出力間関連
- 出力間関連

提案者らはこれらの項目について本暗号と AES の最終候補の一段の性質を検定した所、いずれも相関や偏りはみられなかったと結論づけている。

提案者が行なったこれらの検定を検証する意味を含め、本報告書では下記に示す統計量調査を行なった結果について報告する。

AES 選考の際に NIST が行なった 189 種類の乱数検定法を用いて計算機による検定実験を行ない、何段以上であればこれら全ての乱数検定に合格するかを調査する。

行なった検定項目は次の通りである。

- Frequency,
- Block-Frequency,
- Cusum(2 種),
- Runs,
- Long-Run,
- Rank,
- FFT,
- Aperiodic-Template(148 種),
- Periodic-Template,
- Universal,

- Apen,
- Random-Excursion(8種),
- Random-Excursion-V(18種),
- Serial(2種),
- Lempel-Ziv,
- Linear-Complexity,

なお、各々の乱数検定項目の詳細については [1] を参照して頂きたい。今回の検定で行なった検定実施方法ならびにデータサンプリング方法は以下の通りである。

検定実施方法：

Hamming weight が 2 以下の 128 bit データ 132112 Byte に対し、そのデータを入力値とし、相異なる拡大鍵 300 個によって生成された暗号文 (全体で約 40 MByte) のそれぞれの乱数性を調べ、その合格率を調査する。

合格率の閾値は NIST による検定と同じく 0.9633 以上、すなわち、300 個のデータの中で 96.33 % 以上が乱数と判定された場合、そのテストは合格であるとする。

実験の結果を、図 1 から 図 6 に示す。各々のグラフについて、横軸は各検定テスト項目 1 ~ 189 番目に相当し、縦軸は、集められたデータの乱数性非棄却率である。非棄却率が 1 に近い程乱数性が高いことを示している。図 4 から 図 6 までは、縦軸を 0.96 から 1.0 へ縮尺を変えているので注意して頂きたい。

実験の結果から、今回行なったこれらの乱数検定では、1, 2 段では暗号の出力を乱数と区別できる可能性が極めて大きい、3 段目の出力ではほとんどの項目について乱数検定に合格し、さらには 4,5 段以上では (標準仕様である 16 段も含めて) その出力を乱数と区別することは難しいと判断される。

なお F 関数が十分ランダムであるような Feistel 暗号の場合、3 段以上で Random Permutation, 4 段以上で Strongly Random Permutation となるという性質が知られており [2]、本実験による 1, 2 段目の出力が乱数と区別可能であるという性質は、CIPHERUNICORN-A が Feistel 構造であるということに由来しているものと思われ、CIPHERUNICORN-A 独自の持つ性質ではないものと考えられる。



图 1: 1 段 [成功率 0.0 ~ 1.0]

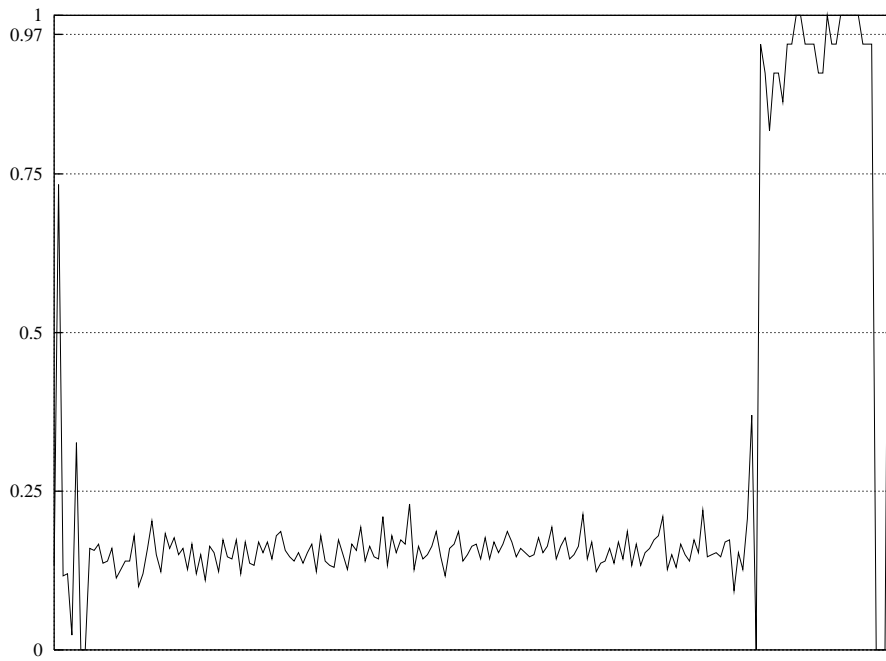


图 2: 2 段 [成功率 0.0 ~ 1.0]

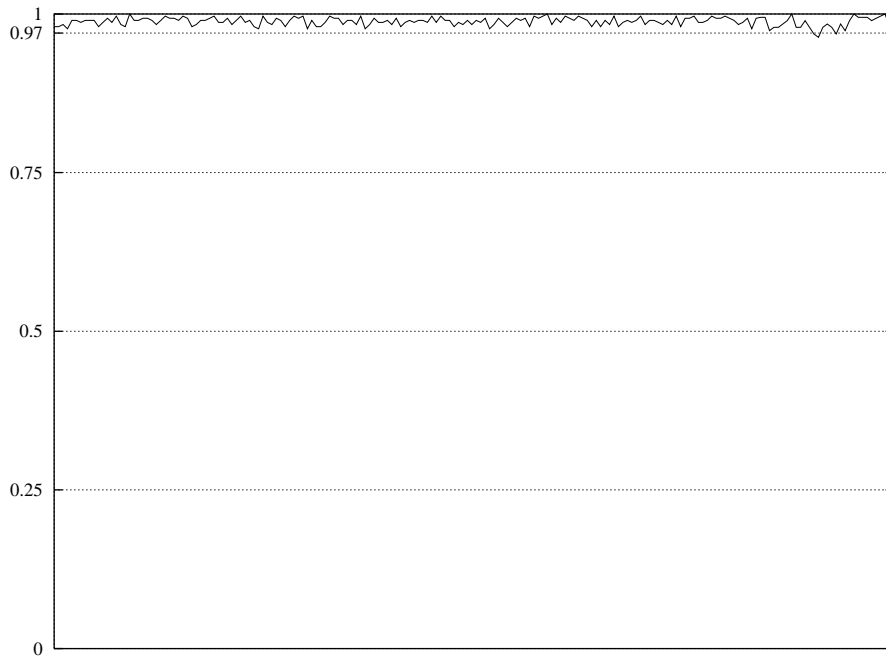


图 3: 3 段 [成功率 0.0 ~ 1.0]

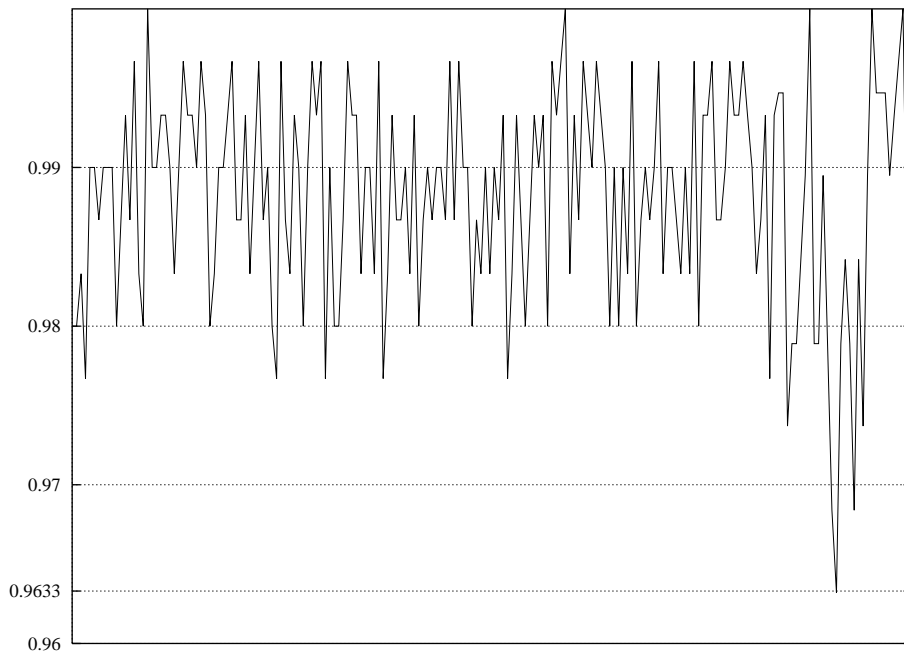


图 4: 3 段 [成功率 0.96 ~ 1.0]

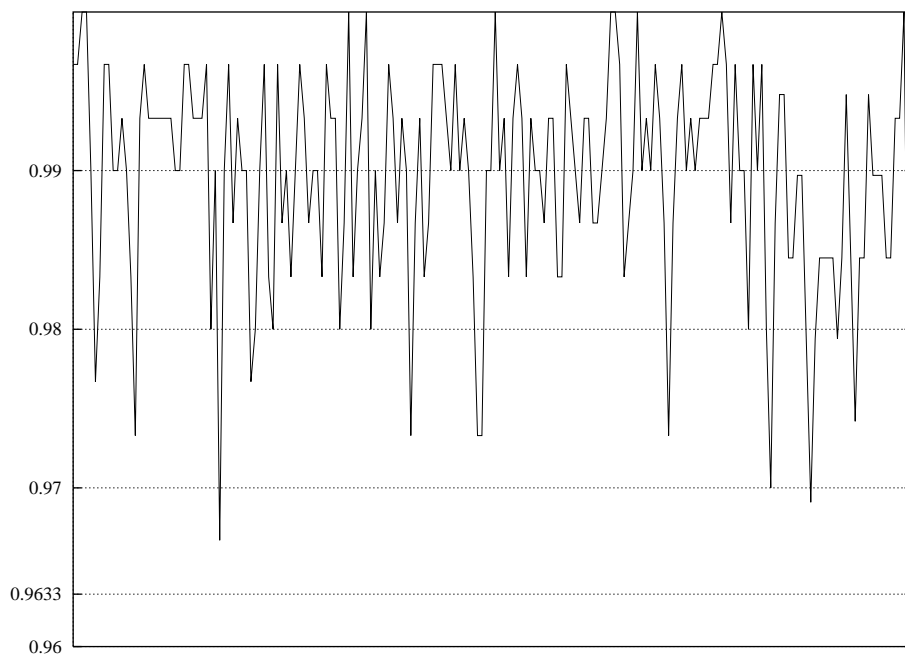


图 5: 4 段 [成功率 0.96 ~ 1.0]

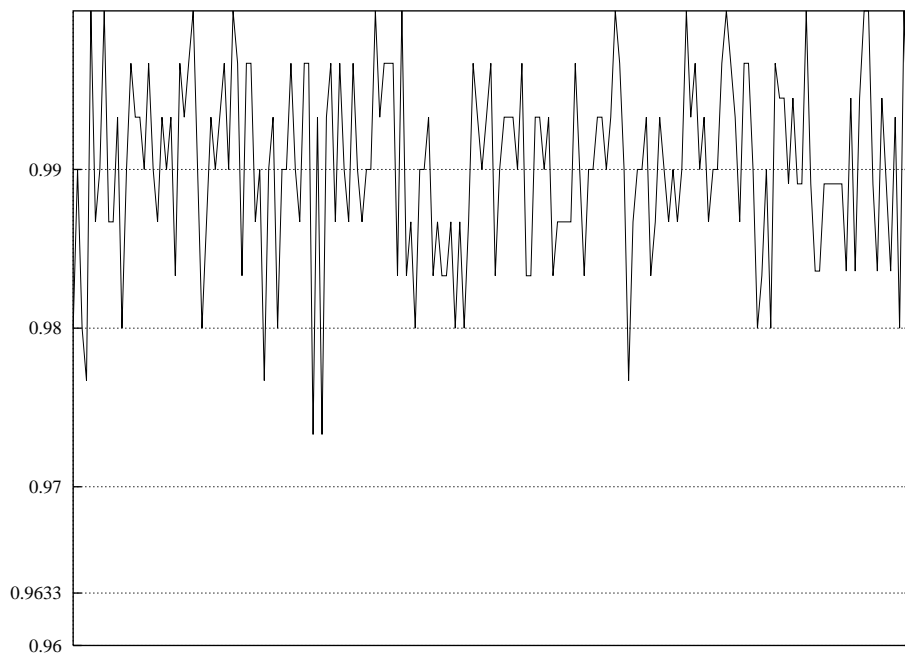


图 6: 5 段 [成功率 0.96 ~ 1.0]

2 差分攻撃

本暗号は F 関数内部の演算において 32 bit 加算さらに 32bit 定数乗算等の算術演算を用いていることから、従来から知られている手法では厳密に差分確率を求めることは、困難であると思われる。設計者による安全性解析では、F 関数に対して次のような単純化を行なった関数 mF を定義し、 mF 関数を用いた Feistel 型ブロック暗号の差分攻撃に対する安全性を評価している。

1. 算術加算は排他的論理和に置き換える
2. 定数乗算は、32 bit データの上位 1 バイトへ入力ビットを集める処理、すなわち下位 3 バイトの排他的論理和を上位 1 バイトへ排他的論理和する処理とする。
3. A3 関数は、バイト全体に差分が入る場合とそうでない場合の場合分けをし、バイト単位で差分値が通過するものとみなす。

単純化された暗号の強度評価結果を根拠として、もとの暗号の安全性を示そうとする場合、暗号の安全性が増すような変形操作や、攻撃者に不利となる仮定が必要なものであってはならない。このような立場から、上記の単純化に関する正当性について、意見を述べておく。

- 1 については、排他的論理和に置き換えてもなお安全であることが示されれば、変換前の暗号の安全性についても保証されたと考えていいと思われる。ただし、本暗号のように算術加算が拡大鍵に対してなされている場合には、鍵の値によって差分確率に差が生じ、weak key となる場合があることに注意が必要である [3]。
- 2 については、乗算に用いる定数の決め方からその性質を近似しようと考えてのことであると思われる。つまり入力 32 ビット全てが出力上位 1 バイトに影響する、という定数の設計方針から、その性質のみをこの方法で抽出するための単純化であると考えられる。乗算に関する差分確率の計算は今だ研究途上にあり、正確に求めることは困難であるため、乗算については何らかの近似が必要となる。本近似方法によって、安全性が増すような構造になっているとは思われないことから、本近似による安全性の検証は、安全性を検証するための近似手段として、唯一の方法というわけではないと思われるが、(例えば、完全に取り去ってしまうという単純化もある) 認められるものと思われる。
- 3 については、評価書の記述だけでは正確には意味を読み取ることが出来なかった。そこで図 3.3 や例などから以下のような近似方法であると解釈する。A3 関数はビット単位の全単射線形関数で、バイト単位の最小分岐数は、簡単な考察から 4 であることが分かるため、バイト単位で考えた場合には、その分岐数が 4 以上であるような任意の入出力差分が通るものとする。本単純化操作も、攻撃者に有利な仮定であることは明らかであり、本方法による設計者の立場からの安全性評価結果にも妥当性があると思われる。

さて、S-box の設計法から各 T_i 関数の差分確率の最大値は、全て 2^{-6} となる。自己評価書ではこのような単純化を施した暗号について、バイト単位の差分探索を行なった結果、図 7 のようなケースが差分特性確率 DP_{mF} を最大にし、その値は

$$DP_{mF} = (2^{-6})^2 = 2^{-12}$$

であると報告している。一般に、F 関数の差分近似の出力差分値が 0 でない場合、その差分確率を q としたときの R 段の差分確率は、最大で $q^{R \times 2/3}$ となることから、16 段における最大差分確率 DCP は、

$$DCP = (DP_{mF})^{(16 \times 2/3)} = 2^{-132}$$

表 1: 提案者による 16 段評価

段数	1	2	3	4	5	6	7	8
差分確率	2^{-12}	2^{-12}	1	2^{-12}	2^{-12}	1	2^{-12}	2^{-12}
全体	2^{-12}	2^{-24}	2^{-24}	2^{-36}	2^{-48}	2^{-48}	2^{-60}	2^{-72}
段数	9	10	11	12	13	14	15	16
差分確率	1	2^{-12}	2^{-12}	1	2^{-12}	2^{-12}	1	2^{-12}
全体	2^{-72}	2^{-84}	2^{-96}	2^{-96}	2^{-108}	2^{-120}	2^{-120}	2^{-132}

表 2: より特性確率を大きくできる 16 段評価

段数	1	2	3	4	5	6	7	8
差分確率	1	2^{-12}	2^{-12}	1	2^{-12}	2^{-12}	1	2^{-12}
全体	2^0	2^{-12}	2^{-24}	2^{-24}	2^{-36}	2^{-48}	2^{-48}	2^{-60}
段数	9	10	11	12	13	14	15	16
差分確率	2^{-12}	1	2^{-12}	2^{-12}	1	2^{-12}	2^{-12}	1
全体	2^{-72}	2^{-72}	2^{-84}	2^{-96}	2^{-96}	2^{-108}	2^{-120}	2^{-120}

であると報告している。

この報告については、いくつか疑問点がある。

提案者は 16 段における DCP の値は $2^{-132} = (2^{-12})^{11}$ 、つまり 11 段分の近似が必要であるとしている。自己評価書ではこれ以上詳しくは述べられていないが、上記で求められた差分近似を、表 1 のように配分したものと思われる。

一方で表 2 のように配分すれば 16 段の差分確率の値が

$$DCP = (DP_{mF})^{10} = 2^{-120}$$

となる。つまりこの mF の差分近似式を用いた 16 段 Feistel 暗号の差分確率は、 2^{-120} 以下であることしか導かれない。この値は 2^{-128} より大きいため、自己評価書における記述を、差分確率に対する安全性を示すための状況証拠とすることは出来ないものと思われる。

ただし、上記の差分近似式が存在する為には、図 7 で示された

$$F : \Delta a \rightarrow \Delta b$$

に対して、同じ成立確率をもつ次の Fm 関数の差分近似式

$$F : \Delta b \rightarrow \Delta a$$

が存在するという強い仮定が必要であり、この仮定はあまり現実的ではないため、暗号の強度を厳密に評価するには適さないかもしれない。一方で、出力差分が 0 の差分近似があれば、この差分近似式単独で複数段の Feistel 暗号の差分近似式を導くことが可能であり、より正確な評価に近くなるものと考えられる。そこで、提案者が用いている F 関数の単純化 mF について、出力差分が 0 となるようなバイト単位の差分近似式を探索した結果、図 8 の様な近似式が見つかった。

た。なお T_k については、すべての k の中で最小となるものを用いている。この図に示される近似式では、active となる (入力に差分がある) T 関数の個数は 3 個であるから、3 個全ての T 関数の入力が互いに独立であるとすれば、 $(2^{-6})^3$ 以下となるが、これらの active な T のうち、一つ目と三つ目については、両 T 関数の間に拡大鍵が、陽には挿入されておらず、入力に強い相関がある可能性が否定できないために、現段階では、この近似式の成立確率は、 $(2^{-6})^2 = 2^{-12}$ 以下であることしか保証されない。さらに、この近似式で差分が 0 でない入力ビット数はわずか 16 ビットであるから、1 段目ならびに 15,16 段目の拡大鍵のうちそれぞれ $16 \times 2 = 32$ ビットを推定する、いわゆる (1+2)- 段消去型攻撃を用いることができる。

本攻撃で推定する鍵ビット数は、32 ビットであり、用いられる 13 段差分近似式 DCP_{13} の成立確率は、最高で

$$DCP_{13} = (2^{-12})^6 = 2^{-72}$$

となるから、攻撃者に対し最も有利な仮定を設定した場合には最低で 2^{72} 個の選択平文で解読できることになる。この結果は、 F 関数を mF 関数によって単純化する方法では、差分攻撃に対する安全性を示せないことを示している。

なお、本暗号について単純化をせず F 関数自身を用いた場合には、定数乗算や鍵加算等に現れる極めて高い非線形性、あるいは A3 関数の拡散性等によって、実際には、図 8 で示されるような差分近似式は存在しないか、あるいは成立確率がもっとずっと小さいものしか存在しないと予想される。本結果は暗号の単純化の仕方に問題があるのであり、現状では差分攻撃に対する安全性を保証するために十分な情報が示されているわけではない、ということを示しているに過ぎず、直ちに CIPHERUNICORN-A の安全性について否定的な結果を導くものではないことをコメントしておく。今後は、より実際の F 関数に近い関数を用いて、正確な差分確率を計算する必要があるものと思われる。

3 線形攻撃

本暗号は、算術加算、定数乗算、データ依存テーブル参照等によって、差分攻撃の時と同様、線形攻撃に対する安全性を差分確率の計算によって正確に見積もることが難しい。提案者は、差分攻撃と同様、F 関数を単純化した mF 関数を用いて、本暗号の安全性を検証している。mF 関数への単純化については、差分攻撃と同様、安全性を評価する上での正当性があるものと思われる。自己評価書によれば、mF 関数の線形近似特性をバイト単位のマスク値で探索した所、図 10 のパターンで線形特性確率が最も大きくなり、その確率 LP_{mF} は以下の値となったと報告している。

$$\begin{aligned} LP_{mF} &= ((S_0S_1S_2S_3)での入力マスク = 0)^4 \times ((S_0S_1S_2S_3)での入力マスク \neq 0)^1 \\ &\quad \times ((S_1S_2S_3)での入力マスク \neq 0)^3 \\ &= 2^{-2.71 \times 4} \times 2^{-2.39 \times 1} \times 2^{-3.08 \times 3} \\ &= 2^{-22.47} \end{aligned}$$

一般に入力マスクが 0 で線形確率が q であるような線形パターンを用いた R 段の Feistel 型暗号の線形特性確率は $q^{R \times 1/2}$ で与えられることから、提案者らは本暗号 16 段の線形特性確率 LCP の下界として、

$$(2^{-22.47})^{16 \times 1/2} = 2^{-179.76}$$

が示されるため、mF 関数を用いた暗号であっても線形攻撃に対し、安全であると結論づけている。

この報告についてはいくつか疑問点がある。

まず、新たに探索した結果、mF 関数について、提案者が示すものより線形確率が大きな線形パターンが見付けられたので報告する (図 11)。本パターンの線形特性確率は、

$$LP_{mF} = ((S_0S_1S_2S_3)での入力マスク = 0)^8 = 2^{-2.71 \times 8} = 2^{-21.68}$$

である。なお、このパターンにおいて active となる 8 個の T 関数の入力それぞれに対し、独立の拡大鍵値を加算しているわけではないため、入力に強い相関関係があるものと思われる、実際の線形確率は、より大きくなることが予想されるが、各 T 関数の入力値の従属性に関する解析が複雑であるため、詳細に求めることが困難である。

本暗号に対して、この線形パターンを用いて線形攻撃を行なう場合、鍵の全数探索より少ない計算量で解読するには、15 段の線形近似と、1- 段消去型攻撃法を用いることが出来る (図 12)。15 段の線形確率は、 $(2^{-21.68})^7 = 2^{-151.76}$ であるから、本攻撃で必要となる既知平文数は、 $2^{151.76}$ 以上となり、今回新たにみつかった線形パターンを用いても、線形攻撃は成功しない。

今回、F 関数を mF 関数で単純化した暗号について線形攻撃に対する安全性を再評価した結果、安全性に関する深刻な問題は発見できなかったものの、提案者による値よりも若干確率が大きくなる線形パターンが見付かった。なお、仕様通りの F 関数についての線形攻撃に対する安全性については、現時点では問題点は何も見付かっていないが、今後はより F 関数に近い形で厳密な評価を行なった方がよいと思われる。

4 高階差分攻撃

提案者は自己評価書において、F 関数の低次項 (1 次 ~ 4 次) の項数を鍵値を変えて計算し、その項数の最大値、最小値、平均値を求め、ランダムな関数の場合、理論的に導かれる期待値から差がほとんどないことを導き、F 関数の代数次数は最低でも 4 次であり、5 段以上では解読に必要な平文を用意することは不可能であると導いている。

本結果は、F 関数を近似すること無く、低次の項数を計算した上で導いたものであり、本暗号を高階差分攻撃を用いて解読することは不可能であるとの結論には妥当性があると思われる。

5 補間攻撃

提案者は自己評価書において、F 関数で用いる 4 種類の S-box の入出力を $GF(2^8)$ の多項式で表現した時の項数を計算し、最低でも 252 項以上であったことから、暗号全体を多項式で表現するために必要な平文を用意することは不可能であり、補間攻撃に対して十分に安全であると期待していると述べている。

本暗号の S-box は、 $GF(2^8)$ の冪乗演算を利用していることから、S-box 単体では、有理式表現を用いた攻撃に対する耐性を検討しなければならないところではあるが、F 関数全体としては、定数乗算や、拡大鍵の算術加算等、別の代数系が用いられていることによって、多項式表現が極めて難しくなっているものと思われる。よって補間攻撃に対しては、段数が少ない場合を除き攻撃が成功することはないとの考えは、妥当性があると思われる。

6 鍵衝突攻撃

提案者は自己評価書において、鍵衝突攻撃は暗号器の内部構造によらず、秘密鍵ビット長のみ依存する、と述べているが、これは、鍵スケジュール部が理論的にランダムな関数であることを仮定した場合に行なう考察であって、本件のように具体的な鍵スケジュール部が存在する場合は、解析が必要であり、記述として不十分であると感じられる。

さて、本暗号についていえば、MT 関数が全単射関数であることと、Feistel 構造を利用した拡大鍵生成法であることから、秘密鍵から、連続する [鍵ビット長 / 32] 個の拡大鍵への 1 対 1 写像が導かれる。このような写像が存在することから、本暗号では拡大鍵生成部に関して鍵衝突が起こらないことが理論的に示される。

参考文献

- [1] “Random Number Generation and Testing,” National Institute of Standard Technology, URL<http://csrc.nist.gov/rng/rng2.html>
- [2] M. Luby and C. Rackoff, “How to Construct Pseudorandom Permutations from Pseudorandom Functions,” *SIAM Journal on Computing*, 17(2), pp.373–386, 1988.
- [3] S. Moriai, K. Aoki, and K. Ohta “Key-Dependency of Linear Probability of RC5,” *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Science*, E80-A(1), 1997.

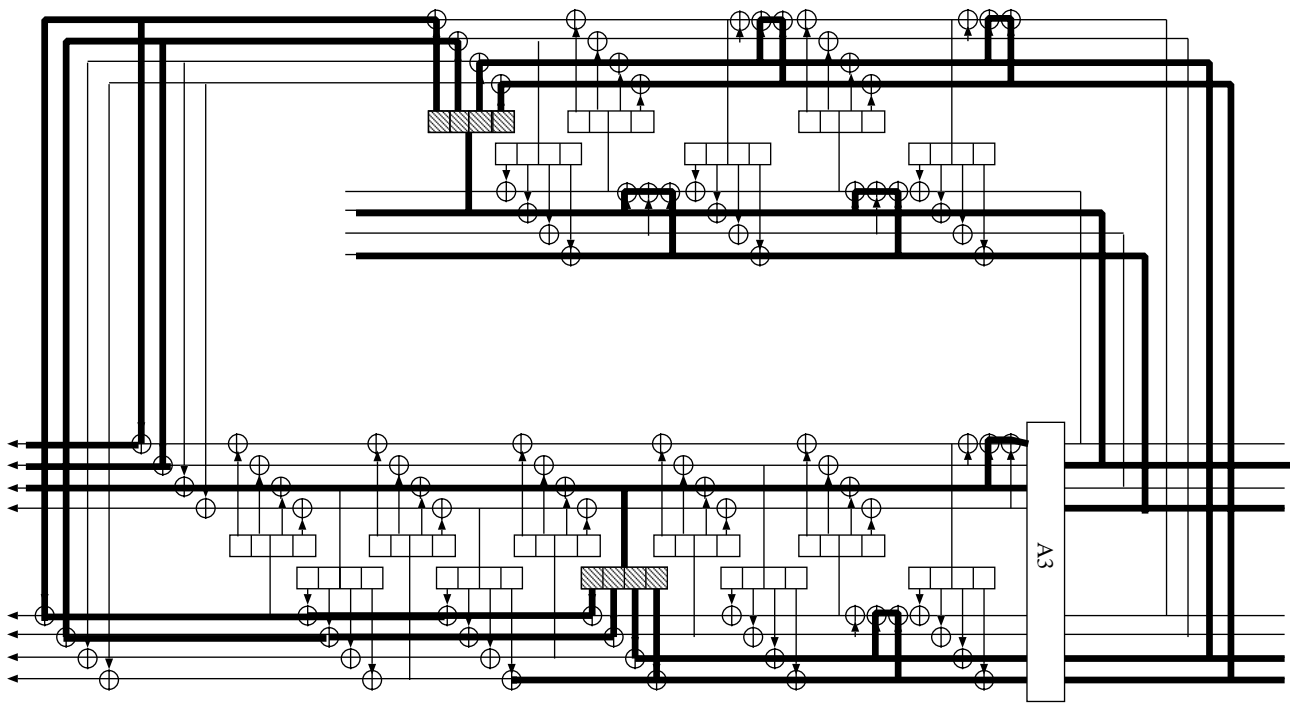


図 7: 提案者らによる最良差分近似

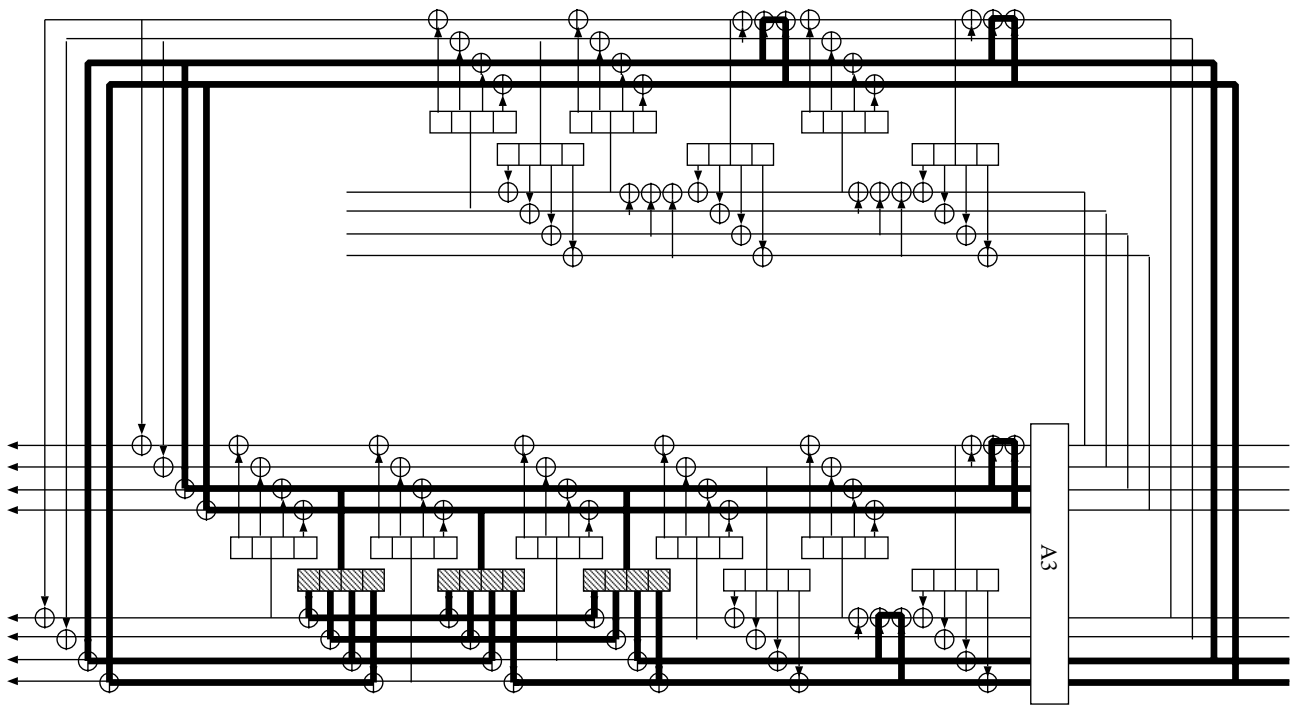


図 8: 出力差分が 0 となる差分近似

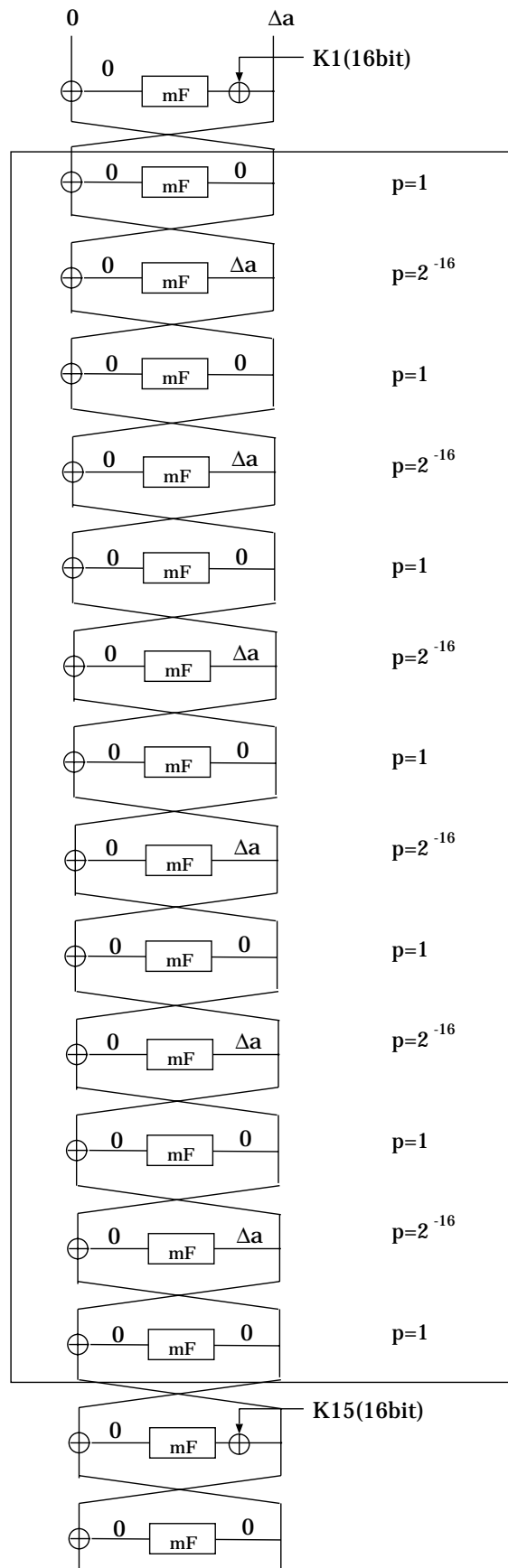


图 9: (1+2)- 段消去型差分攻撃

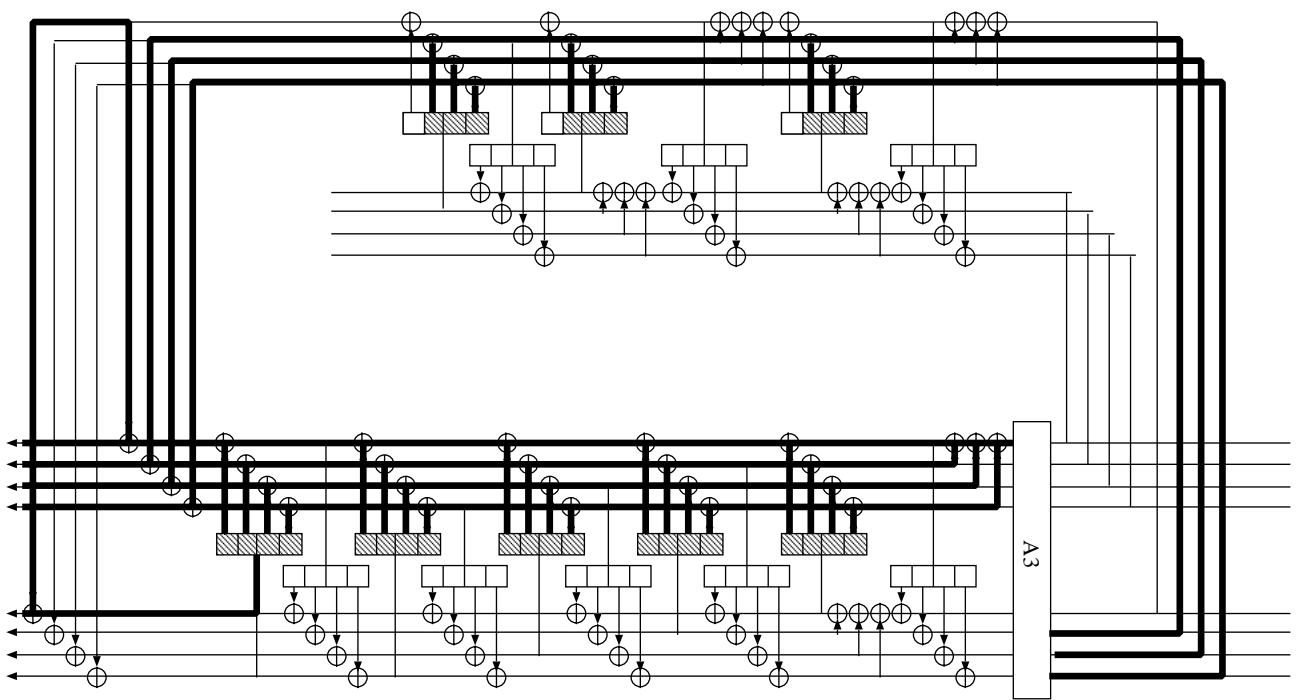


図 10: 提案者らによる最良線形近似

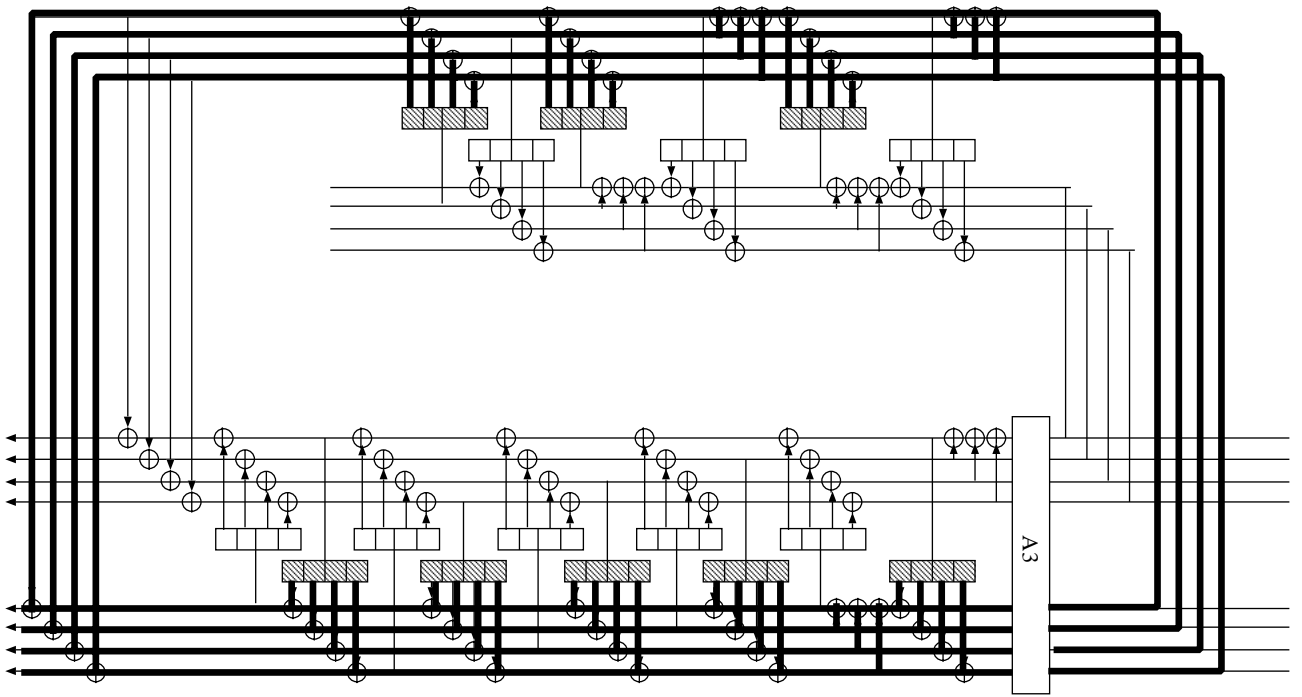


図 11: 新たに見付かった線形近似

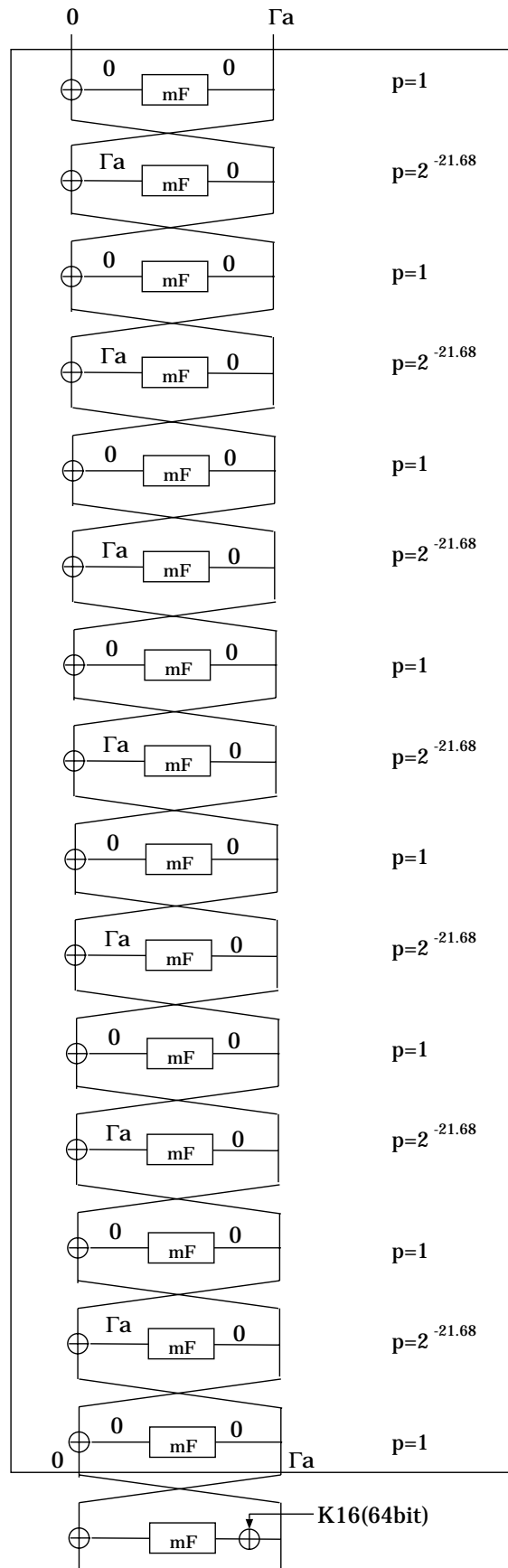


圖 12: 1 段消去型 線形攻擊