通信情報保護學會誌
第3卷 第4號, 1993. 12

DES의 선형 해독법에 관한 해설(II)†

김 광 조*

요 약

본 해설은 1993년 1월 28일 부터 1월 30일까지 일본의 전자통신정보학회 산하 정보 시큐리티 연구회가 연례적으로 개최하는 SCIS'93(Symposium on Cryptography and Information Security)에서 우수 논문상을 수상한 미쓰비시 전기㈜의 마쯔이가 발표한 DES에 관한 선형 해독법에 이어 1993년 3월 ISEC 연구회 자료에 발표한 DES의 암호문 단독공격에 대한 결과를 번역하여 소개한다.

1. 서 론

본해설은[I]에 이어서 1993년 3월 일본의 ISEC 연구회에서 마쯔이가 발표한 DES의 새로운 해독법으로서 암호문 단독공격(COA, Ciphertext Only Attack)에 대하여 소개한다. 문헌[1]에는 해독자에게 주어진 평문이 랜덤하다는 가정하에 다음의 순서에 의해 DES의 기지평문공격을 적용하였다.

Step 1 평문과 암호문 및 키사이에 유의한 선형 근사식을 구성.

Step 2 주어진 선형근사식을 키에 관한 방정식으로서 해를 구함.

이 선형 근사식은 평문의 랜덤성만을 가정하여 구해지며 본 해독법은 본질적으로 기지평문공격에 해당되나 다음의 사항을 추가로 검토하여 본다.

P1 선형근사식의 구성법

P2 해독 성공율의 기술

P3 최량의 근사식을 도출

본고에는 해독자에게 주어진 평문이 랜덤하지 않다는 가정하에 위의 문제를 검토한다.

일반적으로 평문의 수열이 랜덤하지 않다면, 원리적으로 암호문의 수열도 랜덤하지 않으며 그 비랜덤성은 키에 의존한다고 할 수 있다. 따라서 우리의 목표는 **Step 1**에서 평문의 항을 포함하지 않고 유의한 선형 근사식을 구성하고 이것을 **Step 2**에 적용함으로서 암호문 단독공격을 실현하는 것이다.

그러나 암호문 단독공격을 시행할 때 다음에 서술한 바와 같이 기지평문공격에서 고려하지 않았던 몇가지 점을 주의하지 않으면 안된다.

W1 F함수의 입력치의 분포는 라운드에 따라 변화한다.

W2 고려하여야 할 모든 대상과 평문의 총수가 감소한다.

W3 키의 전수 탐색법도 일반적으로 계산량이 증가한다.

W1은 제 1단 F 함수에의 입력값이 어떤 특성을 갖는

* 한국전자통신연구소 실장

† 본 해설의 제 1부는 한국통신정보보호학회지 제3권 제3호 참조

분포를 나타내지만 라운드를 중첩함에 따라 중간 데이타는 점차 랜덤하게 되므로 엄밀하게 각 라운드의 근사 방법은 달리하여야 함을 의미한다.

또한, **W2**는 평문의 랜덤성을 가정하면 해독자가 다루어야 할 서로 다른 평문의 총수는 2^{64} 보다 적어진다는 것을 의미한다. **W3**는 일반적으로 암호문 단독공격의 조건아래 해독자에게 구체적인 평문이 한개도 주어져있지 않으므로 키의 전수 탐색법을 행하는 경우에도 1개의 키를 기각하는 데 반드시 1개의 암호문만으로 충분하지 않다는 것을 의미한다.

이상의 이유로 일반적으로 평문에 관한 가정에 따라 해독 방법을 찾지 않으면 안된다. 따라서 본고에는 평문에 관한 특성이 아래의 3가지 경우로 가정하고 각각에 대한 해독 가능성을 고찰한다.

Case 1 평문이 랜덤한 ASCII부호로 구성됨.
Case 2 평문이 자연 영문의 ASCII 부호로 구성됨.
Case 3 평문의 각 비트는 독립적으로 각각 확률 80%로 0(또는 1)을 취함.

본고에는 위의 상황을 8단 DES에 적용한 결과, **Case 1**은 2^{38}개, **Case 2**는 2^{29}개, **Case 3**은 2^{30}개의 암호문만으로 해독자는 1개의 평문도 모르고 비밀키를 높은 확률로 추정하는 것이 가능하다. 더욱이, 16단의 DES에 대하여도 2^{56}개 보다 적은 암호문으로 암호문 단독공격이 성립하는 상황을 구체적으로 제시한다.

2. 준 비

그림 1과 그림 2는 본고에서 취급하는 DES의 암호화 처리부 및 F함수의 구성도이다. 문헌[1]에서는 초기 전치 IP와 최종 전치 IP^{-1}을 생략하고 초기 전치후의 64비트 정보를 평문이라고 하였으나, 본고에는 평문의 비트 위치를 고려할 필요가 있을 때, IP 수행전의 본래의 평문을 참 평문이라고 불러 구별하며 최종 전치는 항상 생략한다.

본고에는 다음의 기호를 사용하며 특히 단수에 의존하지 않는 경우는 단수를 표시하는 첨자를 생략한다. 또한 각 그림에 있어서 오른쪽은 하위로하고 특히 최하위 비트는 0번째 비트로 약속한다.

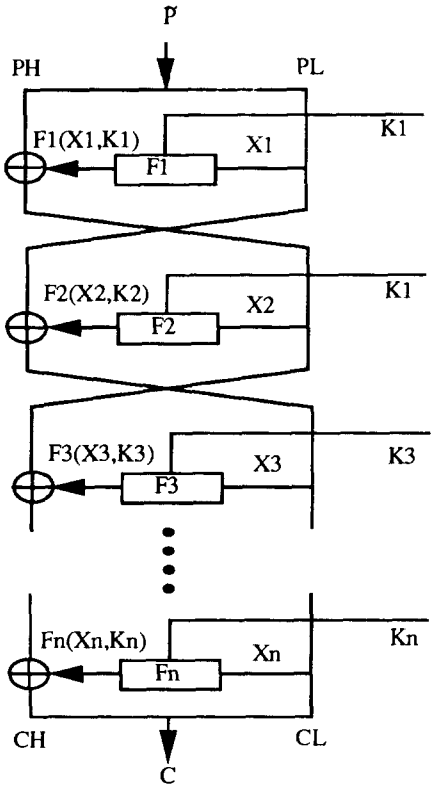


그림 1. DES 암호화 과정

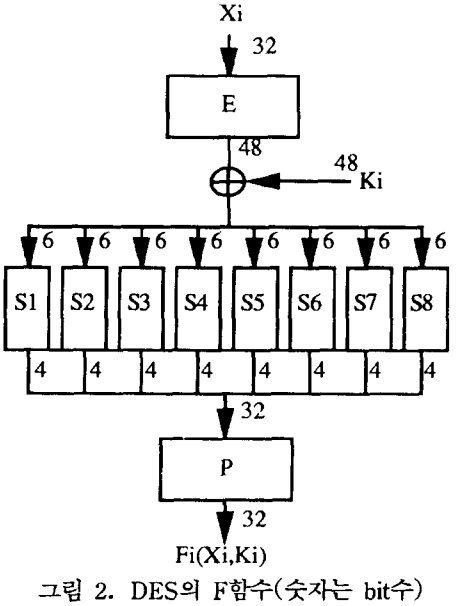


그림 2. DES의 F함수(숫자는 bit수)

- P : 평문 64 비트
- C : 암호문 64 비트
- P_H, P_L : 평문의 상위 32비트, 하위 32 비트
- C_H, C_L : 암호문의 상위 32비트, 하위 32 비트
- X_i : 제 i 단의 F 함수의 32 비트 입력
- K_i : 제 i 단의 확대키 48 비트
- $F_i(X_i, K_i)$: 제 i 단의 F 함수
- $A[i]$: A의 제 i 번째 비트
- $A[i, j, \cdots, k] : A[i] \oplus A[j] \oplus \cdots \oplus A[k]$

3. 선형 해독법의 개요

문헌[1]에 의거 선형 해독법의 개요에 대하여 간단히 서술한다.

선형 해독법(Linear Cryptanalysis)의 목표는 랜덤하게 주어진 평문 P와 대응하는 암호문 C 및 키 K에 대하여 유의 확률 p로 다음식과 같은 형태의 선형 근사식을 구성하는 데 있다.

$$P[i_1, i_2, \cdots, i_a] \oplus C[j_1, j_2, \cdots, j_b] = K[k_1, k_2, \cdots, k_c] \quad (1)$$

여기서 $i_1, \cdots, i_a, j_1, \cdots, j_b, k_1, \cdots, k_c$는 고정된 비트 위치이고, 유의 확률이란 $p \neq 1/2$를 의미한다. 실제 이 식의 구성에 성공하면 암호 해독자는 주어진 평문과 대응하는 암호문의 쌍으로 부터 좌변을 계산하고 그 값의 분포도에 따라 우변의 키 1 비트 $K[k_1, k_2, \cdots, k_c]$를 의미있게 추정이 가능하다. $|p-1/2|$ 값이 큰 선형 근사식을 이용하여 해독을 시도하면, 필요한 평문수가 감소하므로 이것은 해독자에게 바람직한 상황이 된다. 식(1)의 형태 중 $|p-1/2|$ 값이 가장 큰 것을 최량 표현이라 부르고 이때의 p를 최량 확률이라 부르고 DES의 경우 구체적인 결과를 [1]에 제시하였다.

실제 해독에 있어서 복수의 키 비트를 효율적으로 구하기 위해 $n-1$단까지의 근사를 이용하는 방법을 취한다. 즉, 최종단을 복호함으로서 식의 중간에 F함수로 포함한 $n-1$단의 선형 근사식을 다음과 같이 구성한다.

$$P[i_1, i_2, \cdots, i_a] \oplus C[j_1, j_2, \cdots, j_b] \oplus F_n(C_L, K_n)[l_1, l_2, \cdots, l_d] = K[k_1, k_2, \cdots, k_c] \quad (2)$$

여기서 이 식에 잘못된 K_n을 대입하면 그 유의성을 상실한다. 즉, 식(2)의 성립 확률이 1/2에 근접한다. 따라서 해독자는 각 K_n에 대하여 실제 좌변을 계산하고, 그 계산 결과의 분포를 관찰하고 유의 확률로 K_n 및 $K[k_1, k_2, \cdots, k_c]$를 추정할 수 있다. DES에 대해서 위의 선형 근사식을 구체적으로 구성하기 위하여 우선 S-box의 선형성을 나타내는 지표로서 다음의 정의로 부터 출발한다.

DES는 8개의 S-box $S_1, \cdots, S_8$은 각각 6비트의 입력을 갖고 있으므로 입력 패턴의 총수는 $2^6=64$이다. 각 S-box에 대하여 몇개의 입력 비트의 Exclusive Or값과 몇개의 출력 비트의 Exclusive Or값이 64개 중 몇개가 일치하는 가를 우선 조사한다.

정의 1 *S-box* S_a (a=1, 2, ⋯, 8)에 있어서, S_a의 입력 64개 중에 $1 \leq \alpha \leq 63$으로 마스크된 입력 비트 위치와 $1 \leq \beta \leq 15$로 마스크된 출력 비트의 *Exclusive Or*값이 일치되는 경우의 수를 $NS_a(\alpha, \beta)$로 표현한다. 즉,

$$NS_a(\alpha, \beta) = \#\{x \mid 0 \leq x < 64, \bigoplus_{s=0}^{5} (x[s] \cdot \alpha[s]) = \bigoplus_{t=0}^{3} (S_a(x)[t] \cdot \beta[t])\}$$

여기서 $\cdot$는 비트 단위의 곱셈을 의미한다.

이 값이 32가 되지 않는 경우 S-box의 입출력 비트간에 상관이 있다고 생각한다. S-box의 구성법에 의해 $NS_a(x,y)$는 항상 짝수값을 가지며, x가 1,32 또는 33의 경우 $NS_a(x,y)$는 32가 된다.

임의의 S-box의 선형 근사식은 F함수의 선형 근사식으로 확장 가능하다. 또한, F함수의 선형 근사식을 적절히 중첩시킴으로서 알고리즘 전체의 선형 근사식을 구할 수 있다. 이때 그 근사 확률은 계산 가능하다.

보조정리 1. Piling-up Lemma

독립적인 확률 변수 $X_i (1 \leq i \leq n)$이 확률 p_i로 0, 확률 $1-p_i$로 1을 취할 때, $X_1 \oplus X_2 \oplus \cdots \oplus X_n = 0$이 되는 확률은 다음 식과 같이 얻어진다.

$$2^{n-1} \prod_{i=1}^{n} (p_i - 1/2) + 1/2 \quad (3)$$

식(2)의 해독 성공률은 다음과 같이 계산 가능하다. 간단히 $F(X, K)[l_1, l_2, \cdots, l_d]$에 영향을 주는 키 비트 집합은 특정한 1개의 S-box에 입력되는 6비트만이라고 가정하자. 본고에는 실제 이 경우만을 취급한다.

이때, K의 64개의 후보 $K_i(i=1,2,\ldots,64)$에 대하여 $F(X, K_i)[l_1, l_2, \cdots, l_d] = F(X, K)[l_1, l_2, \cdots, l_d]$가 성립할 확률을 q_i로 하면 다음이 성립한다.

보조정리 2. N개의 랜덤한 기지 평문을 이용하여 식(2)를 풀때, 그 해독 성공 확률(키 K의 6비트와 식(2)의 우변을 바르게 추정할 확률)은 적어도 다음의 식으로 주어진 값의 이상이다.

$$\int_{x=-2\sqrt{N}|p-1/2|}^{\infty} \left(\prod_{K_i=K} E_i(x)\right) \frac{2}{\sqrt{\pi}} e^{-x^2/2}dx \quad (4)$$

여기서, Π는 K과 서로 다른 모든 K의 후보 K_i에 대한 것으로, $E_i(x)$는 다음의 식으로 정의되는 함수이다.

$$E_i(x) = \int_{-x-4\sqrt{N}(p-1/2)q_i}^{x+4\sqrt{N}(p-1/2)(1-q_i)} \frac{2}{\sqrt{\pi}} e^{-x^2/2}dy \quad (5)$$

4. 암호문 단독공격의 예-1

암호문 단독공격의 첫번째 예로서 평문에 관한 다음의 가정을 하고 8단 DES에 적용한다.

가정 1. 평문의 각 바이트는 랜덤한 ASCII코드(각 바이트의 제7비트는 항상 0)이다.

그림 3과 같이 8단 DES의 선형 근사표현에서 출발한다. 그림에서 제2단 F함수의 선형 근사는 $NS_8(2,8) = 30$에서, 제3단 및 제7단 F함수의 선형근사를 $NS_7(4,8) = 36$에서, 제4단 및 제6단 F함수의 선형 근사는 $NS_8(2,12) = 28$에서 각각 구한 것이다. 그 결과 평문의 1비트를 포함한 다음의 선형근사식을 구하였다.

$$P_L[27] \oplus C_H[27] \oplus C_L[0] \oplus F_8(C_L, K_8)[27] = K_2[1] \oplus K_3[8] \oplus K_4[1] \oplus K_6[1] \oplus K_7[8] \quad (6)$$

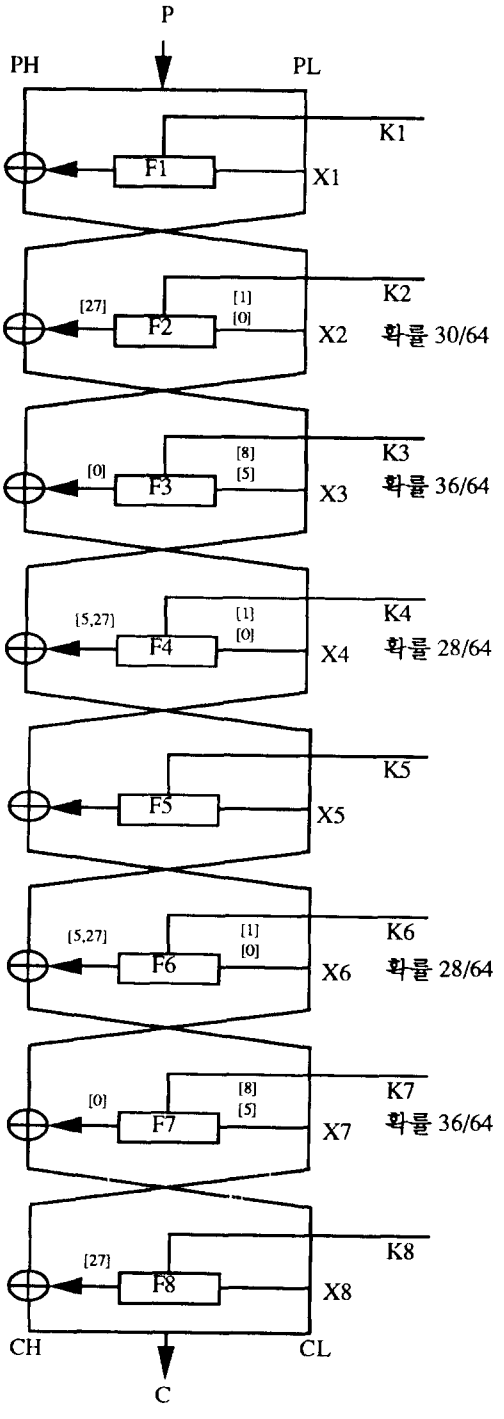


그림 3. 8단 DES의 암호문 단독 공격(1-1)

식(6)의 성립확률은 Piling-up Lemma에 의해

$$1/2 + 2^4(-2/64)(4/64)^2(-4/64)^2 = 1/2 - 2^{-17}$$

이다. 이경우 평문의 어떤 분포를 가정하였으므로 엄밀히는 각 단에 있어서 F함수의 입력 값을 랜덤하지는 않다. 그러나, 그림 3에서의 선형근사는 제 2단 이후 F 함수의 근사만을 이용하여 구성하였으므로 그의 입력값은 실제 랜덤하지 않다고 보아도 좋다. 그러면 여기서 DES의 초기전치 IP를 고려하면 $P_L[27]$은 초기 전치를 하기 전의 참 평문의 제39비트 $\hat{P}[39]$에 상당하고 이것은 가정에 의해 항상 0이 된다. 따라서 식(6)은 평문을 포함하지 않는 것이라고 생각해도 좋다. 식(6)에 있어 $F_8(C_L, K_8)[27]$에 영향을 주는 키 비트는 K_8의 제 0비트부터 5비트까지 총 6비트이므로 결국 해독자는 주어진 암호문만으로 6비트와 식(6)의 우변의 1비트를 구할 수 있게 된다.

다음은 해독의 성공확률을 조사하자. 이확률을 식(4)를 이용하여 계산하면 표 1과 같다. 이 결과는

표 1. 식 (6)의 해독 성공 확률

N	식(4)의 계산치
$4\|p-1/2\|^{-2}$	0.74
$8\|p-1/2\|^{-2}$	0.93
$16\|p-1/2\|^{-2}$	0.99

기지평문공격의 예와 같이 실제는 계산치보다 좋은 성공확률을 얻을 수 있다. 결국 $8(2^{-17})^{-2} = 2^{37}$개 정도의 암호문이 주어지면 평문의 1비트도 몰라도 이 7비트의 해독에 대부분의 경우 성공한다. 그러나 그림 4에서 보듯이 8단 DES의 선형 근사로부터도 식(6)과 실질적으로 같은 확률로 성립하는 표현식을 구할 수 있다. 그림에 있어서 제2단 및 제6단 F 함수의 선형근사는 $NS_6(4.8) = 34$부터, 제3단 및 제5단 F함수의 선형 근사는 $NS_1(2.2) = 30$에서 구한 것이다.

$$P_L[28] \oplus C_H[28] \oplus F_8(C_L, K_8)[28] = K_2[14] \oplus K_3[43] \oplus K_5[43] \oplus K_6[14] \quad (7)$$

윗식의 성립확률은 Piling-up Lemma에 의해

$$1/2 + 2^3(2/64)(-2/64)^2 = 1/2 + 2^{-17}$$

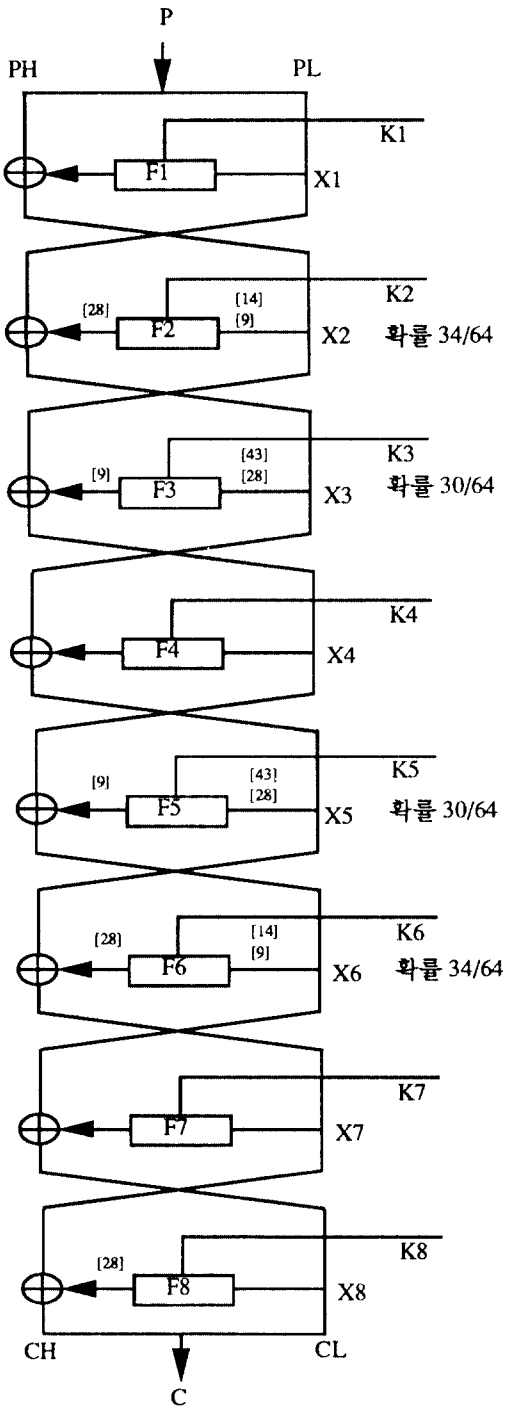


그림 4. 8단 DES의 암호문 단독 공격(1-2)

임을 알 수 있다. 윗식에도 평문의 1비트가 있으나 이 $P_L[28]$은 참 평문의 제 31비트 $\hat{P}[31]$에 상당하고 항상 0이 된다. 본 식에서 $F_8(C_L, K_8)[28]$에 영향을 주는 키 비트는 K_8의 제 12비트부터 제 17비트까지의 6비트이고 해독자에게 주어진 암호문만으로 이 6비트 및 식(7)의 우변의 1를 구할 수 있게 된다. 이 식을 이용한 경우 해독 성공확률은 식(4)에 의해 표 2와 같이 구해진다. 이 결과, $16(2^{-17})^{-2} = 2^{38}$개

표 2. 식 (7)의 해독 성공 확률

N	식(4)의 계산치
$8\|p-1/2\|^{-2}$	0.79
$16\|p-1/2\|^{-2}$	0.92
$32\|p-1/2\|^{-2}$	0.98

정도의 암호문으로 이 7비트의 해독에 대부분 성공한다.

위의 2가지 해독을 독립적으로 병렬로 실행할 수 있으므로 결국 2^{38}개의 암호문이 주어지면 해독자는 평문을 1개도 모르고 키의 14비트를 높은 확률로 산출할 수 있음을 의미한다.

5. 암호문 단독공격의 예-2

실제적인 예로 평문에 관하여 다음의 가정을 하여 전절과 같이 8단 DES에 대한 암호문 단독공격을 해보자.

가정 2. 평문의 자연 영문을 ASCII 부호로 표현한 것이다.

이 조건에서 해독은 그림 5에서 나타난 선형근사식에서 출발한다. 여기서 제 2단 F함수의 선형 근사는 $NS_5(16,14) = 42$에서, 제 3단 $NS_1(4,4) = 30$에서, 제 4단 및 6단은 $NS_5(16,15) = 12$에서, 제 7단은 $NS_1(8,4) = 26$에서 각각 구한 것이다. 이 결과 8단 DES에 대한 선형 근사식은 다음식과 같이 구해진다.

$$P_L[7,18,24] \oplus C_{II}[7,18,24,29,30] \oplus C_L[15] \oplus F_8(C_L, K_8)[7,18,24,29,30] = K_2[22] \oplus K_3[44] \oplus K_4[22] \oplus K_6[22] \oplus K_7[45] \quad (8)$$

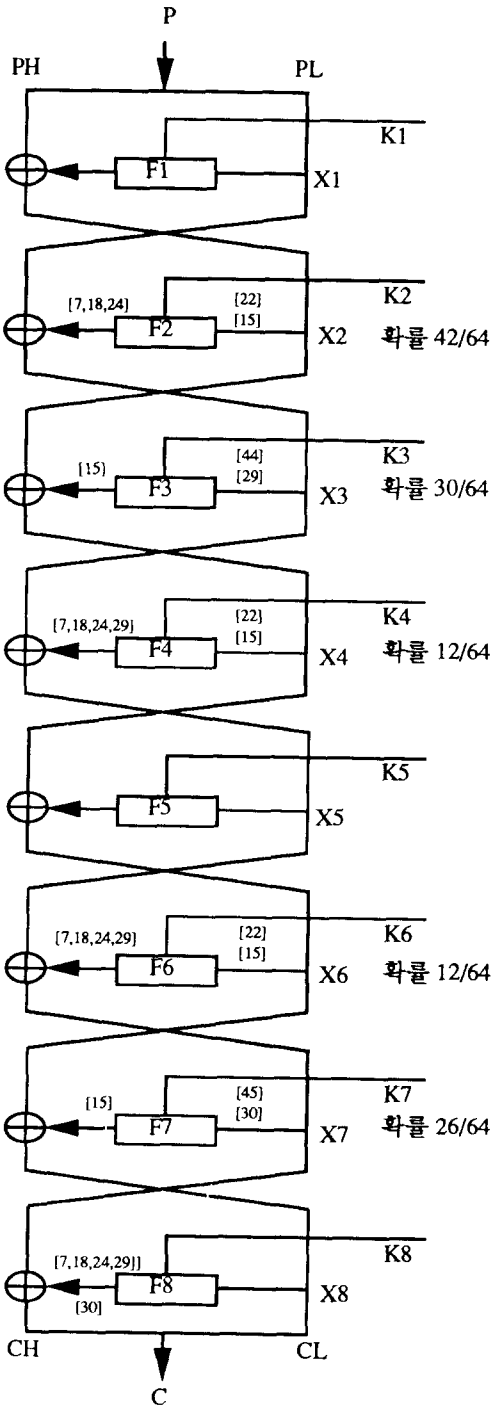


그림 5. 8단 DES의 암호문 단독 공격(2)

윗식의 성립확률은 Piling-up Lemma에 의해

$$1/2 + 2^4(10/64)(-2/64)(-20/64)^2(-6/64) = 1/2 + 1.46 \times 2^{-11}$$

임을 알 수 있다. 식(8)은 전절과 같이 제 2단 이하의 F함수만을 근사하여 얻은 것이므로 각 F함수의 근사에 있어서 그 입력치는 랜덤하다고 가정한다. 식(8)에 나타난 $F_8(C_L, K_8)[7, 18, 24, 29, 30]$ 에 영향을 주는 키 비트는 S-box S_2에 입력되는 K_8의 제 36에서 41비트와 S-box S_5에 입력되는 K_8의 제 18에서 32비트로 총 12비트이다. 이 중 S_5만 $NS_5(16, 15)=12$을 이용하여 근사시키면 다음의 선형 근사식을 얻는다.

$$P_L[7, 18, 24] \oplus C_H[7, 18, 24, 29, 30] \oplus F_8(C_L, K_8)[30] = K_2[22] \oplus K_3[44] \oplus K_4[22] \oplus K_6[22] \oplus K_7[45] \oplus K_8[22] \quad (9)$$

윗식의 성립확률은 Piling-up Lemma에 의해

$$1/2 + 2(-20/64) \times 1.46 \times 2^{-12} = 1/2 - 1.83 \times 2^{-12}$$

임을 알 수 있다. 여기서 초기 전치 IP를 고려하면 식(9)에 포함된 평문의 3비트 $P_L[7]$, $P_L[18]$, $P_L[24]$는 각각 참 평문의 1, 45, 63번째 비트에 해당한다. 따라서 해독자에 필요한 정보는 평문을 바이트 단위로 생각할 때 제 1, 5, 7비트의 분포상황이다. 여기서 제 7비트는 가정에 의해 항상 0이 되기 때문에 결국 제 1 및 5비트의 Exclusive Or 값을 알면 된다. 이것을 찾기 위해 다량의 영문을 조사한 결과, 0이 될 확률은 최대 0.35임을 알 수 있었다.

이 값을 가정하고 식(9)에서 평문항을 소거한 식(10)을 얻었고,

$$C_H[7, 18, 24, 29, 30] \oplus F_8(C_L, K_8)[30] = K_2[22] \oplus K_3[44] \oplus K_4[22] \oplus K_6[22] \oplus K_7[45] \oplus K_8[22] \quad (10)$$

이 식의 성립확률은

$$1/2 - 2(0.35-0.5) \times 1.83 \times 2^{-12} = 1/2 + 1.10 \times 2^{-13}$$

이 된다. 식(10)에서 $F_8(C_L, K_8)[30]$에 영향을 주는 키 비트는 K_8의 제 36비트에서 41비트까지 6비트이므로 결국 해독자는 주어진 암호문만으로 이 6비트 및 식(10)의 우변 1비트를 구할 수 있다. 이 해독의 성공확률은 표3과 같다. 전절에서와 같이 $8(1.10 \times 2^{-13})^{-2} = 1.65 \times 2^{28}$개의 암호문이 주어진다면 7비트의 해독은 대부분의 경우 높은 확률로 성공한다고 할 수 있다. 식(8)에서 $F_8(C_L, K_8)[30]$은 해독자에게 이제는 아는 값이고 미지의 값은 K_8의 제 18에서 23까지의 6비트만이 된다(식(10)의 우변을 구하면 식(8)의 우변은 이 6비트에서 자연적으로 구하게 됨). 이때 식(8)에서 평문의 항을 제외한 식이 성립할 확률은

$$1/2 + 2(0.35-0.5) \times 1.46 \times 2^{-11} = 1/2 + 1.75 \times 2^{-13}$$

이므로 여기서 6비트를 구하는데 필요한 평문의 수는 약 $8(1.75 \times 2^{-13})^{-2} = 1.31 \times 2^{27}$개라고 생각된다(이 경우 미지수는 6비트 뿐이므로 그 해독 성공확률은 식(4)보다 높으며 상세한 내용은 생략한다). 결국 2^{29}개 정도의 암호문이 주어지면 해독자는 평문의 한 비트도 모르고 암호화 키의 15비트를 알 수 있게 된다.

표 3. 식 (10)의 해독 성공 확률

N	식(4)의 계산치
$4\|p-1/2\|^{-2}$	0.70
$8\|p-1/2\|^{-2}$	0.92
$16\|p-1/2\|^{-2}$	0.99

6. 암호문 단독공격의 예-3

평문중 특정 비트를 지정하지 않고 모든 비트가 동일한 확률을 갖는다고 가정하고 8단 DES를 대상으로 암호문 단독공격의 가능성을 고찰한다.

가정 3. 평문의 64비트 중 어떤 비트 위치에 있어서 1이 나타날 확률은 20%이다.

이 경우 8단 DES암호의 근사 표현의 예로 그림 6에 나타내었다. 이것은 8단 DES의 최량 표현이고 기지평문공격에 사용되었던 것이다. 여기서, 제 1단에 있어서 다음식의 좌변을 우변으로 치환하여 근사를

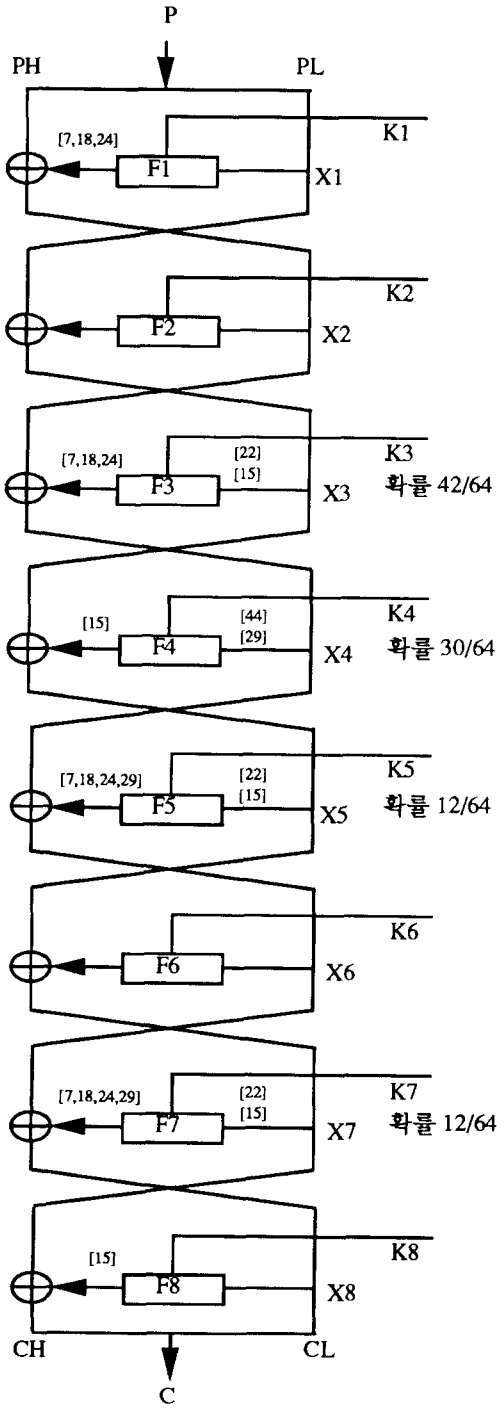


그림 6. 8단 DES의 암호문 단독 공격(3-1)

하였다.

$$F_1(P_L, K_1)[7,18,24] = P_L[12,16] \oplus K_1[19,23] \quad (11)$$

윗식은 $NS_5(34,14) = 16$에서 구한 것으로 그 성립 확률은 P_L이 랜덤하다는 가정하에 0.25이다. 그러나 가정 3이 성립하는 평문은 랜덤하지 않으므로 S-box에 입력되는 데이타의 분포는 그 값 이전에 Exclusive Or를 한 K_1의 값에 의존한다.

여기서 본 절에는 이점을 엄밀히 조사하면 식(11)의 좌변이 0이 되는 확률 $p(K_1)$을 K_1의 값으로 분류한다(표 4). 이 확률은 S-box S_5에 입력되는 K_1의 제 18비트부터 23비트만 의존한다는 것을 주의하자. 표 4는 좌측에 이 6비트값, 우측에 $p(K_1)$을 나타내었다.

표 4. $p(K_1)$의 분포

K_1	$p(K_1)$	K_1	$p(K_1)$	K_1	$p(K_1)$	K_1	$p(K_1)$
0	0.326	16	0.236	32	0.405	48	0.356
1	0.333	17	0.284	33	0.614	49	0.582
2	0.633	18	0.330	34	0.372	50	0.190
3	0.675	19	0.536	35	0.402	51	0.287
4	0.394	20	0.242	36	0.687	52	0.586
5	0.399	21	0.298	37	0.697	53	0.602
6	0.690	22	0.309	38	0.636	54	0.288
7	0.766	23	0.590	39	0.690	55	0.367
8	0.709	24	0.594	40	0.751	56	0.670
9	0.449	25	0.324	41	0.652	57	0.594
10	0.798	26	0.566	42	0.606	58	0.338
11	0.678	27	0.410	43	0.438	59	0.264
12	0.695	28	0.371	44	0.728	60	0.609
13	0.598	29	0.307	45	0.522	61	0.386
14	0.739	30	0.367	46	0.510	62	0.268
15	0.798	31	0.556	47	0.610	63	0.295

그림 6에 있어서 다음의 근사를 한 단은 제 3단이므로 여기서 그 입력값은 비교적 랜덤한 값으로 되어 있다고 생각된다. 제 1단에 있어서 $F_1(P_L, K_1) = 0$로 근사하고 3단 이후에 있어서 통상의 선형 근사를 하므로 평문과 암호문을 연결하는 다음의 근사식을 구할 수 있다.

$$P_H[7,18,24] \oplus C_H[15] \oplus C_L[7,18,24,29] \oplus F_8(C_L, K_8)[15] = K_3[22] \oplus K_4[44] \oplus K_5[22] \oplus K_7[22] \quad (12)$$

이 식의 성립확률은 Piling-up Lemma에 의해

$$1/2 - 2^4(p(K_1)-1/2)(10/64)(-2/64)(-20/64)^2 = 1/2 - (p(K_1)-1/2) \times 1.95 \times 2^{-8}$$

이 된다. 가정에 의해 임의의 비트 위치 i에 대하여 $P[i]=0$이 확률 80%로 성립하므로 식(12)의 3개소에 나타난 $P[i]$를 0으로 하여도 근사식의 유의성을 잃지 않는다.

이 근사식은 식(13)과 같으며

$$C_H[15] \oplus C_H[7,18,24,29] \oplus F_8(C_L, K_8)[15] = K_3[22] \oplus K_4[44] \oplus K_5[22] \oplus K_7[22] \quad (13)$$

이 식의 성립확률은

$$1/2 - 2^3(0.8-0.5)^3(p(K_1)-1/2) \times 1.95 \times 2^{-8} = 1/2 - (p(K_1)-1/2) \times 1.69 \times 2^{-10}$$

이 된다. 식(13)은 평문에 관한 비트가 포함되어 있지 않으므로, 해독자는 암호문만으로 키의 7비트를 구할 수 있다. 이때 해독자에 필요한 암호문 수는

$$8\left((p(K_1)-1/2) \times 1.69 \times 2^{-10}\right)^{-2} = (p(K_1)-1/2)^{-2} \times 1.40 \times 2^{21}$$

개가 되는 것을 알 수 있다(이 경우 식(4)의 계산결과는 [1]참조). 예를 들면, 2^{29}개의 암호문으로 이 7비트를 해독할 경우 K_1이

$$\sqrt{(1.40 \times 2^{21})/2^{29}} = 0.074 < |p(K_1)-1/2|$$

을 만족하면 거의 틀림없이 성공한다. 이와 같이 키는 표4에 의하면 57개가 존재한다. 또한 2^{30}개의 암호문으로 이 7비트를 해독한 경우 K_1은

$$\sqrt{(1.40 \times 2^{21})/2^{30}} = 0.052 < |p(K_1)-1/2|$$

를 만족하면 좋다. 이러한 키는 표4에 의하면 60개가 존재한다.

위의 결과로 식(13)을 이용하여 7비트의 키를 찾는데는 2^{29}개의 암호문을 이용한 경우 57/64=89%로 성공하고, 2^{30}개의 암호문을 이용한 경우 60/64=94%로 성공한다고 할 수 있다. 식 (13)은 기지평문공격에서 이용한 근사식에서 출발한 것으로 위의 관계식은 일반적으로 몇개를 구할 수 있으므로 이것은 병렬로 실행하여 해독하면 다음과 같이 성공확률이 계산 가능하다.

예를들면, 그림 7에 나타난 근사식을 이용하여 같은 방법의 해독을 시행해 보자. 마찬가지로 $F_1(P_L, K_1)[7,18,24,29]=0$이 성립할 확률 $q(K_1)$을 조사해 보자(표5). 이 결과 간단한 계산에 의해

표5. $q(K_1)$의 분포

K_1	q(K_1)	K_1	q(K_1)	K_1	q(K_1)	K_1	q(K_1)
0	0.223	16	0.477	32	0.281	48	0.731
1	0.232	17	0.708	33	0.281	49	0.780
2	0.542	18	0.796	34	0.297	50	0.773
3	0.346	19	0.796	35	0.239	51	0.739
4	0.232	20	0.650	36	0.281	52	0.765
5	0.315	21	0.731	37	0.476	53	0.837
6	0.404	22	0.796	38	0.254	54	0.739
7	0.507	23	0.845	39	0.332	55	0.796
8	0.261	24	0.424	40	0.457	56	0.738
9	0.204	25	0.631	41	0.261	57	0.712
10	0.580	26	0.768	42	0.327	58	0.685
11	0.292	27	0.719	43	0.178	59	0.524
12	0.204	28	0.401	44	0.261	60	0.654
13	0.169	29	0.458	45	0.261	61	0.703
14	0.523	30	0.719	46	0.235	62	0.524
15	0.327	31	0.719	47	0.220	63	0.662

확률 $1/2 - (q(K_1)-1/2) \times 1.01 \times 2^{-11}$로 성공하는 식(14)를 얻는다.

$$C_H[15] \oplus C_L[7,18,24] \oplus F_8(C_L, K_8)[15] = K_3[22] \oplus K_4[44] \oplus K_5[22] \oplus K_7[22] \quad (14)$$

따라서 2^{29} 또는 2^{30} 개의 암호문으로 키 7비트를 해독할 경우, 간단한 계산에 의해 각각

$$|q(K_1)-1/2| > 0.247, \quad |q(K_1)-1/2| > 0.175$$

를 만족하면 높은 확률로 성공한다고 할 수 있다. 그러나 표4와 표5에 의하면, $|q(K_1)-1/2| > 0.074$

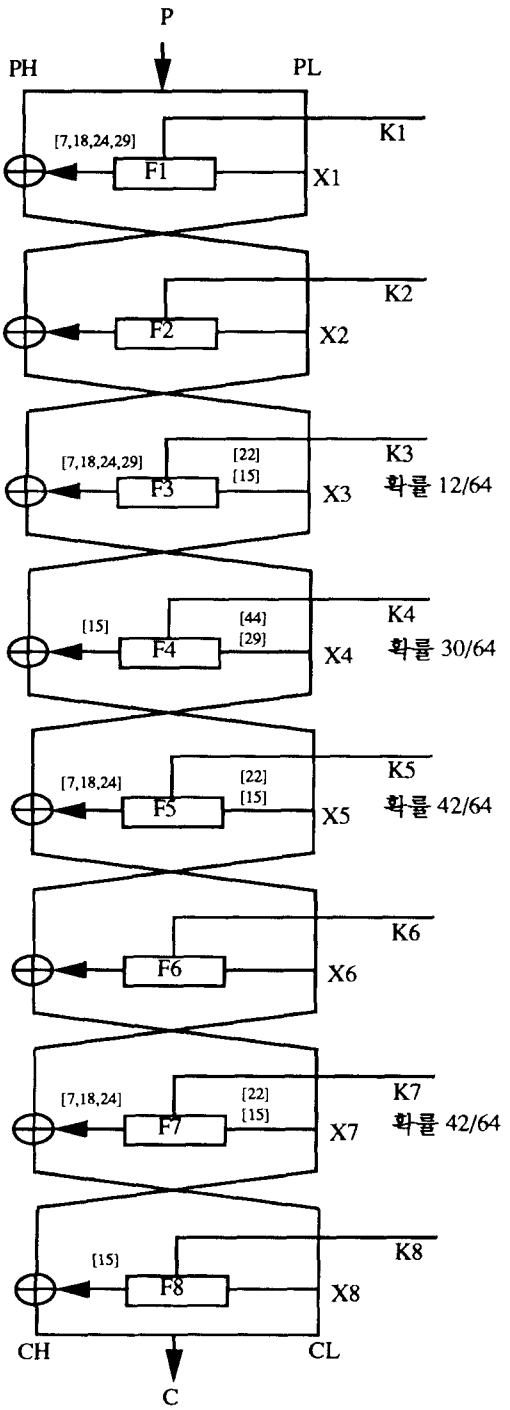


그림 7. 8단 DES의 암호문 단독 공격 (3-2)

또는 $|q(K_1)-1/2|>0.247$를 만족하는 키는 총 62개이고 따라서 2개의 해독을 병렬로 실행하면 62/64 = 97%로 어느 한쪽이 성공한다고 말할 수 있다. 또한, 2^{30}개의 암호문으로 부터 키의 7비트를 구한 경우, $|q(K_1)-1/2|>0.052$ 또는 $|q(K_1)-1/2|>0.175$가 성립할 필요가 있으나 표4와 표5에 의하면 64개 모두의 키가 어느 한쪽도 만족한다. 따라서 이 2개의 해독을 병렬로 수행하면 K_1에 의존하지 않고 어느 쪽의 한쪽이 틀림없이 성공한다.

7. 암호문 단독공격의 예-4

DES의 암호문 단독공격의 마지막 예로서 16단 DES에 적용 가능한 선형 근사식을 찾아본다. 우선 그림 8에 나타난 16단 DES의 선형 근사를 고찰한다. 이것은 16단에서 최량 표현으로 랜덤하게 주어진 평문과 대응하는 암호문에 대하여 확률 $1/2+1.19\times 2^{-22}$로 성립한다.

$$\begin{aligned} &P_H[7,18,24]\oplus P_L[12,16]\oplus C_H[15]\\ &\quad\oplus C_L[7,18,24,29]\oplus F_{16}(C_L,K_{16})[15]\\ &\quad=K_1[19,23]\oplus K_3[22]\oplus K_4[44]\oplus K_5[22]\\ &\quad\oplus K_7[22]\oplus K_8[44]\oplus K_9[22]\oplus K_{11}[22]\\ &\quad\oplus K_{12}[44]\oplus K_{13}[22]\oplus K_{15}[22] \end{aligned} \tag{15}$$

이 식을 이용하여 기지 평문공격으로 16단 DES의 암호화 키의 최초 7비트, 즉 K_{16}의 제 42부터 47비트까지의 6비트 및 식(15)의 우변 1비트를 해독하는데 필요한 평문의 수는

$$8(1.9\times 2^{-22})^{-2}=1.41\times 2^{46}$$

가 된다[1]. 이 식에는 평문의 5비트만이 나타나 있으므로 키의 7비트를 해독하기 위하여는 이 확률 정보가 주어져 있으면 좋다. 예로서 이 5비트가 독립적으로 확률 80%로 0가 되고 다른 비트는 독립적으로 확률 50%로 0이 된다고 가정하자.

이 경우 엄밀히는 전절에서 고찰한 바와 같이 제 1단 F함수의 근사는 별도로 고려할 필요가 있다. 그러나, 이 경우 제 1단 F함수의 입력중 랜덤하지 않는 것은 2비트 뿐이므로 실제 그 차이는 문제가 되지 않는다. 여기서 식(15)에서 부터 출발한다.

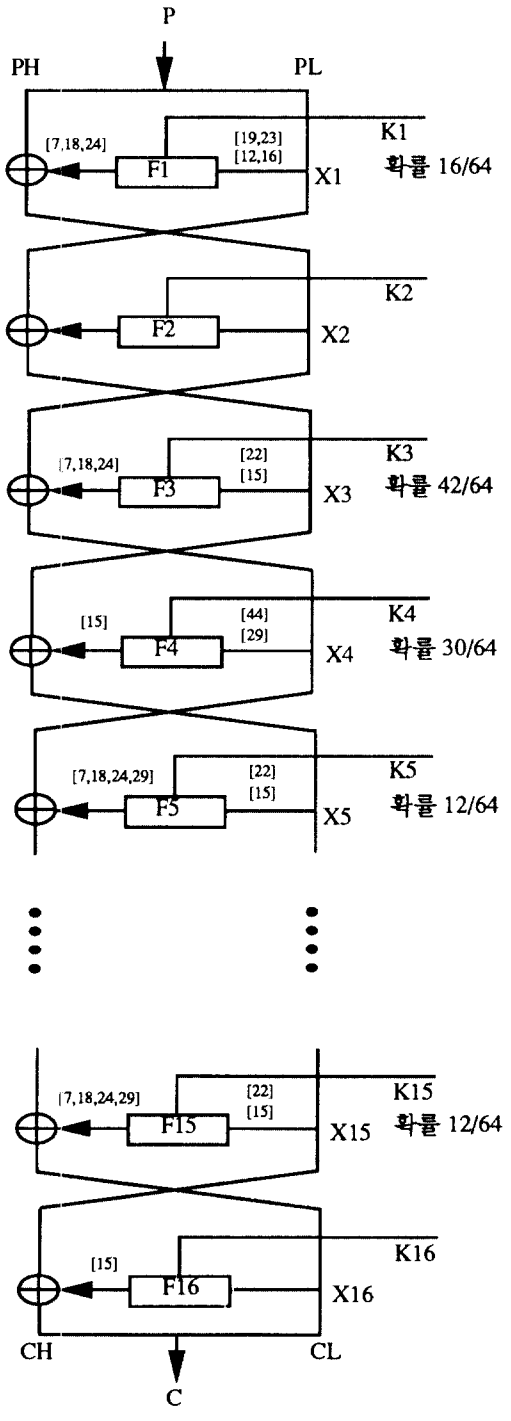


그림 8. 16단 DES의 암호문 단독 공격

이때 식(15)에 나타난 5개의 평문 비트 위치 i에 대하여는 $P[i]=0$가 확률 80%로 성립하므로 $P[i]$를 0으로 치환하여 근사한 식(16)은 그 유의성을 잃지 않는다.

$$C_H[15] \oplus C_L[7,18,24,29] \oplus F_{16}(C_L, K_{16})[15] = K_1[19,23] \oplus K_3[22] \oplus K_4[44] \oplus K_5[22] \oplus K_7[22] \oplus K_8[44] \oplus K_9[22] \oplus K_{11}[22] \oplus K_{12}[44] \oplus K_{13}[22] \oplus K_{15}[22] \quad (16)$$

식(16)은 확률

$$1/2 + 2^5(0.8-0.5)^5 \times 1.19 \times 2^{-22} = 1/2 + 1.48 \times 2^{-26}$$

으로 성립한다. 따라서 식(16)을 이용하면 해독자는 구체적인 평문을 전혀 알 필요없이 키의 7비트를 구할 수 있다. 또한 이때 해독에 필요한 암호문의 수는

$$8(1.48 \times 2^{-26})^{-2} = 1.82 \times 2^{53}$$

개가 됨을 알 수 있다.

8. 결 론

본 해설은 DES의 암호문 단독공격을 소개하였으며 본 방식은 블럭 암호에 대하여 범용적인 암호문 단독공격으로 최초의 결과이다.

본 방식은 DES에 대하여만 해석을 하였으나, 이것은 타 암호에도 적용 가능하고 이것에 대하여는 별도로 고찰한다. 다음은 본 방식에서 검토하지 않았던 문제점 및 향후 대책 등에 대하여 정리한다.

Problem 1 일반적인 최량 표현의 효율적 도출방법.

문헌[1]에는 20단까지의 DES에 대하여 평문이 랜덤하다는 가정하에 최량 근사식을 구하였으나 일반적으로 평문에 관한 확률 분포를 고려한 N단 DES의 최량 표현 및 최량 확률은 고찰하지 않았다. 본고에서 고찰한 암호문 단독공격에 있어서 8단 DES의 선형 근사는 반드시 최량의 근사식을 이용하였다고는 할 수 없다. 이와 같이 일반적인 형태에 있어 최량 근사식의 산출법은 아직 문제로 남아있다.

Problem 2 해독 성공 확률의 정확한 계산

문헌[1] 및 본고에는 실제 해독에 있어서 성공 확률은 **보조정리** 2에 의하여 계산하였다. 그러나, **보조정리** 2와 F 함수의 출력값이 키에 대하여 독립이다고 가정하고 구하였으므로 엄밀한 계산은 하지 못하였다고 할 수 있다.

실험에 의하면, 이 계산치보다 훨씬 좋은 값을 얻었으므로 **보조정리** 2의 계산 결과가 90%를 넘는 경우에는 "$8|p-1/2|^{-2}$개 (또는 $16|p-1/2|^{-2}$개)의 암호문이 주어지면 대부분의 경우 해독이 성공한다." 라고 할 수 있다. 이 확률 계산을 보다 정확히 산출하는 것이 문제라고 생각한다.

Problem 3 최량의 S-box의 구성법.

본 해독법의 성공 확률을 감소시킬 수 있는 S-box의 구성 예를 제시하는 자체는 비교적 간단하다고 생각되나 최소화될 수 있는 S-box의 구체적인 구성법은 향후의 문제이다.

여기에 일반적으로 암호학적으로 강한 S-box가 요구되는 조건, 예를들면 Differential Cryptanalysis와의 관계를 종합적으로 고려한 S-box의 구성법을 구하는 것이 중요한 문제라고 생각된다.

참 고 문 헌

1. M. Matsui, "Linear Cryptanalysis of DES Cipher(I)", Proc. of SCIS'93, SCIS93-3C, 1993.

2. E. Biham and A. Shamir, "Differential Cryptanalysis of DES-like Cryptosystems", J. of Cryptology, Vol.4, pp.3-72, 1991.

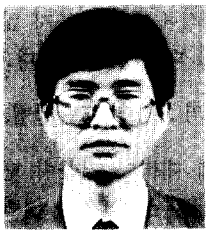
3. E. Biham and A. Shamir, "Differential Cryptanaylsis of FEAL and N-Hash", Advances in Cryptology-Eurocrypt'91, Lecture Notes in Computer Science, Vol.547, pp.1-16, 1991.

4. E. Biham and A. Shamir, "Differential Cryptanaylsis of the full 16-round DES", Crypto'92 Extended Abstracts, pp.12.1-12.5, 1992.

5. A. Tardy-Corfdir and H. Gilbert, "A Known Plaintext Attack of FEAL-4 and FEAL-6", Advances in Cryptology-Crypto'91, Lecture Notes in Computer Science, Vol.576, pp.172-182, 1991.

6. M. Matsui and A. Yamagishi, "A New Method for Known Plaintext Attack of FEAL Cipher", Eurocrypt'92 Extended Abstracts, pp. 77-87, 1992.

□ 著者紹介



金 光 兆 (正會員)

1980年 延世大學校 電子工學科(學士)
1983年 延世大學校 大學院 電子工學科(碩士)
1990年 요꼬하마 國立大學 大學院 電子情報工學科(博士)
現 在 : 韓國電子通信研究所 室長

關心分野 : 暗號學 및 應用分野, M/W通信