암호 알고리즘에 관한 연구

이호현* · 박종민* · 조범준**

*조선대학교 컴퓨터공학과 · **조선대학교 컴퓨터공학부

Research about Cipher Algorithm

Ho-Hyun Lee* · Jong-Min Park* · Beom-Joon Cho**

*Dept. of Computer Engnieering, Chosun University

**Dept. of Computer Engnieering, Chosun University

E-mail : pjm@macrodata.co.kr

요 약

현대사회는 전자, 통신, 컴퓨터, 반도체 기술들의 비약적인 발전에 따른 고도의 정보통신과 정보처리에 기반을 둔 정보화 사회로 모든 업무가 컴퓨터에 의존하고 있는 실정이다. 또한 컴퓨터 네트워크와 데이터베이스의 이용 및 기술의 괄목할 만한 발전에 따라 수많은 사용자가 컴퓨터를 이용하여 동시에 자료를 수집, 검색, 처리, 전송, 저장할 수 있게 되었다.

그러나 컴퓨터를 이용한 정보의 처리 및 사용이 편리해진 반면에 자료의 노출이 상대적으로 심해져, 비밀을 요하는 자료의 보안 문제가 대두 되었다. 이에 따라 자료의 보안을 위하여 컴퓨터가 나오기 전의 방법과는 다른 컴퓨터를 이용한 새로운 암호 방법들이 제시되어 계속적으로 발전되고 있다. 이에 본 논문에서는 향후 연구에 사용할 암호 알고리즘을 선택하기 위하여 관용 암호 알고리즘 중에서 DES와 3DES, 공개키 암호 알고리즘 중 RSA, 타원곡선 암호 알고리즘에 관하여 연구하였다.

ABSTRACT

Modern society is information-oriented society that allow fetters in electron, telecommunication, computer, highly Information-Communication and information processing by great development of semi-conductor technologies. All businesses are depending on computer. Also, Great many user according to development who is worth watching eagerly of computer network and utilization of database and technology could collect, search, handle, transmit and store data at the same time using computer.

But, while processing and use of information that use computer become convenient, exposure of data became serious relatively. For these reason, Security problem of data that need the secret rose. Accordingly, new encryption methods to use computer for security of data are presented and are developed continuously. Studied about DES, 3DES, RSA, ECC algorithm to select cipher algorithm to use in research hereafter in this treatise.

키워드

공개키, RSA, DES, 타원곡선

Ⅰ. 서 론

현 사회는 수많은 정보를 필요로 하는 고도의 정보화 시대이다. 때문에 정보를 수집 생산하고 처리하기 위하여 컴퓨터의 이용이 날로 증대되고 있다. 또한 컴퓨터 네트워크와 데이터베이스의 이용 및 기술의 괄목할 만한 발전에 따라 수많은 사용자가 컴퓨터를 이용하여 동시에 자료를 수집, 검색, 처리, 전송, 저장할 수 있게 되었다.

그러나 컴퓨터를 이용한 정보의 처리 및 사용이 편리해진 반면에 자료의 노출이 상대적으로 심해져 비밀을 요하는 자료의 보안 문제가 대두되었다.

이에 따라 자료의 보안을 위하여 컴퓨터가 나오기 전의 방법과는 다른 컴퓨터를 이용한 새로운 암호 방법들이 제시되어 계속적으로 발전되고 있다.

본 논문에서는 향후 연구에 사용할 암호 알고리즘을 선택하기 위하여 암호화 키와 복호화 키가 같은 대칭키 암호 방식인 관용 암호 알고리즘과, 암호화 키와 복호화 키가 다른 공개키 암호 알고리즘으로 암호 알고리즘을 분류하고, 관용 암호 알고리즘 중에서 DES와 3DES, 공개키 암호 알고리즘 중 RSA, 타원곡선 암호 알고리즘에 관하여 연구하였다.

II. 관용 암호 알고리즘

관용 암호 알고리즘은 암호화와 복호화에 동일한 키를 사용함으로 공통키 암호 알고리즘 또는 암호화와 복호화 과정이 대칭적이어서 대칭 암호 알고리즘이라고도 한다.

관용 암호 알고리즘은 수천년 전부터 사용되어 오고 있는 암호 방식으로 평문의 문자를 다른 문자로 환자(치환) 또는 문자의 위치를 바꾸는 전치과정으로 구성된다.

1. DES 암호 알고리즘

DES는 평문 64비트를 암호문 64비트로 변환시키는 암호 방식으로 64비트의 키를 사용하고 있다. 이 키는 8비트마다 패리티(Parity) 비트 하나씩을 포함하고 있어 DES의 암호화 과정에는 56비트만이 적용된다.

DES 암호 알고리즘의 기본 동작은 전치, 환자와 mod 2 연산으로 구성되어 있다. 다시 전치는 평형 전치, 확대 전치 그리고 축약 전치 등의 세 종류가 있으며 환자는 S-box라는 환자 장치에서 이루어진다. 전치와 환자 그리고 mod 2 연산으로 구성된 DES의 암호화 과정은 그림 1과 같다.[1]

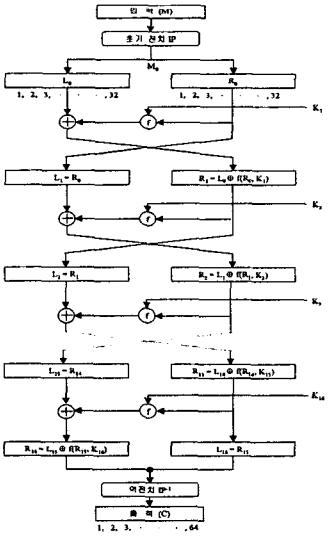

그림 1. DES 암호화 과정

DES의 동작 과정을 살펴보면 다음과 같이 크게 세 가지 과정으로 나눌 수 있다.[2]

①. 평문 M의 64비트는 초기 전치(initial permutation) IP를 거쳐 $IP(M) = M_0$는 32비트씩 나누어져 L_0, R_0로 나누어진다. 초기 전치 IP는 평형 전치로 58번째 비트를 1번째 비트로 50번째 비트를 2번째 비트로 64비트의 위치를 변경시킨다.

② 초기 전치 출력 R_0, L_0는 아래와 같은 함수 계산을 16회 반복한다.

$$L_i = R_{i-1}$$
$$R_i = L_{i-1} \oplus f(R_{i-1}, K_i)$$

$\oplus$는 mod 2 연산을 하는 EX-OR를 의미하며 L_i, R_i는 그림 1의 32비트씩의 중간 데이터를 말한다. 서브키 K_i는 48비트의 DES 암호화키로써 $K_1, K_2, \ldots, K_{16}$의 값은 서로 다르며 16회 암호화 과정에 사용된다. 함수 f는 S-Box를 포함한 환자 과정을 의미한다.

③ R_{16}과 L_{16}은 초기전치의 역전치인 IP^{-1}를 거쳐 64비트의 암호문이 된다. 즉 $C = IP^{-1}(R_{16}, L_{16})$으로 L_{16}과 R_{16}이 서로 반대로 되어 있는 것에 유의해야 한다. IP^{-1}는 IP의 역전치로, IP에서 58번째 비트가 1번째 비트로 전치되었기 때문에 반대로 IP^{-1}에서는 1번째 비트가 58번째 비트로 전치되어야 한다.

마찬가지로 IP에서 50번째 비트가 2번째 비트로 전치되었기 때문에 IP^{-1}에서는 2번째 비트가 50번째 비트로 전치되어야 한다.

2. 3DES 암호 알고리즘

DES 암호 방식의 특기할 만한 특징 중에 하나가 보수 특성이다. 평문의 보수와 키의 보수를 DES 암호 방식에 입력시키면 다음 특징을 나타내게 된다.[3]

$$C = E_K(M), \overline{C} = E_{\overline{K}}(\overline{M})$$

이 결과는 S-Box가 비선형적이라는 것을 생각하면 의외의 결과이다. 그러나 DES 암호 방식의 보수 특성은 DES 암호 방식의 사용에 치명적인 약점으로 작용하지는 않는다. 즉, 평문과 암호문 한 쌍을 알고 있을 때 소모적 공격을 줄일 수 있는 것은 아니다. 하지만 선택 평문 공격에서 보수 특성을 이용하면 소모적 공격 시간을 줄일 수 있다. 즉 평문 M과 $C_1 = E_K(M)$, $\overline{C_2} = E_K(\overline{M})$이 주어지면, $C_2 = E_{\overline{K}}(M)$이 되어 키 K의 변화에 따른 결과 값으로 C_1, C_2 어느 쪽을 찾을 수 있어 키 공격 시간을 줄일 수 있다.

이러한 DES 문제점을 해결하기 위해 2개의 키 쌍을 갖는 3DES 암호 알고리즘이 개발 되었다.

3DES의 동작 과정은 2개의 키 쌍을 적용하여 먼저 첫 번째 키로 평문을 암호화하고, 2번째 키로 이 평문을 복호화한 다음 다시 첫 번째 키로 암호화 하여 강한 암호문을 얻게 된다. 여기에 사용되는 암호화, 복호화 과정은 DES의 과정을 그대로 사용하게 된다.

III. 공개키 암호 알고리즘

1. RSA 암호 알고리즘

1976년 Diffie와 Hellman이 공개키 암호 방식의 개념을 발표한 후, 1978년 MIT의 Rivest, Shamir와 Adleman이 처음 RSA라고 하는 유망한 공개키 암호 방식을 제안하였다.[4]

이들은 합성수의 소인수 분해의 어려움을 이용하여 RSA 암호 방식을 실현하였다. 가입자는 백 자리 크기 이상의 두 개의 소수 p, q를 선택하여 $n = p \cdot q$를 계산한다. 이 때 p와 q를 알고 있는 사람은 n을 계산하기 쉽지만 n만 알고 있는 사람은 n으로부터 p와 q를 찾는 소인수 분해는 어렵다.[5]

p와 q를 선택하여 n을 계산한 다음 Euler 함수 값 $\phi(n) = (p-1)(q-1)$와 서로소인 K_e를 선택한다. 다시 유클리드 알고리듬을 이용하여 다음 식을 만족하는 K_d를 계산한다.

$$K_e \cdot K_d \equiv 1 \bmod \phi(n)$$

K_e와 n은 공개 목록에 등록하여 공개하고 K_d는 비밀리에 보관한다. 즉, K_e는 공개 암호화 키가 되고 K_d는 비밀 복호화 키가 된다. RSA 암호 방식의 구성 절차와 암호화, 복호화 과정은 그림 2와 같다.

가입자 B에게 평문 M을 비밀리에 전달하려는 가입자 A는 공개목록에서 가입자 B의 공개 암호화 키 K_{e_B}를 찾아, 암호문 $C \equiv M^{K_{e_B}} \bmod n_B$를 계산하여 가입자 B에게 전송한다. 가입자 B는 가입자 A로부터 수신한 암호문 C를 자신이 비밀리에 보관하고 있는 복호화 키 K_{d_B}로 평문 $M \equiv C^{K_{d_B}} \bmod n_B$을 복원한다.

물론 가입자 B가 가입자 A에게 평문을 비밀리에 전송하려면 가입자 A의 공개 암호화 키 K_{e_A}를 공개 목록에서 찾아 암호문 $C \equiv M^{K_{e_A}} \bmod n_A$를 계산하여 가입자 A에게 전송한다. 가입자 A는 가입자 B로부터 수신한 암호문 C를 자신이 비밀리에 보관하고 있던 비밀 복호화 키 K_{d_A}로 평문 $M \equiv C^{K_{d_A}} \bmod n_A$을 복원한다.

RSA 암호 방식의 복호화 과정은 공개 암호화 키 K_e와 비밀 복호화 키 K_d가 $\bmod \phi(n)$상에서 서로 역수 관계가 있으므로

$K_e \cdot K_d \equiv 1 \bmod \phi(n)$ 임의의 t에 관하여

$K_e \cdot K_d = t\phi(n)+1$ 이 성립한다.

암호문 C는 $C \equiv M^{K_e} \bmod n$이므로 복호화 과정을 적용하면 다음 식이 성립한다.

(단, $\gcd(M,n) = 1$이라고 가정한다.)

$$C^{K_d} = (M^{K_e})^{K_d} \bmod n = M^{K_e K_d} \bmod n$$
$$= M^{t\phi(n)+1} \bmod n = (M^{\phi(n)})^t \cdot M \bmod n$$
$$\equiv 1^t \cdot M \bmod n$$
$$(\because M^{\phi(n)} \equiv 1 \bmod n, \text{ if } (M,n) = 1)$$
$$= M \bmod n$$

위 식에서 알 수 있는 바와 같이 복호화 키 K_d로 암호문 C로부터 평문 M을 복원할 수 있다.

가입자 A	공개정보 $K_{e_A}, n_A, K_{e_B}, n_B$	가입자 B
p_A, q_A		p_B, q_B
$n_A = p_A \cdot q_A$		$n_B = p_B \cdot q_B$
$\phi(n_A) = (p_A-1)(q_A-1)$		$\phi(n_B) = (p_B-1)(q_B-1)$
$\gcd(K_{e_A}, \phi(n_A)) = 1$		$\gcd(K_{e_B}, \phi(n_B)) = 1$
$K_{e_A} \cdot K_{d_A} \equiv 1 \bmod \phi(n_A)$		$K_{e_B} \cdot K_{d_B} \equiv 1 \bmod \phi(n_B)$
$C \equiv M^{K_{e_B}} \bmod n_B$	C	$M \equiv C^{K_{d_B}} \bmod n_B$

그림 2. RSA 공개키 암호 방식

2. 타원곡선 암호 알고리즘

타원곡선을 이용한 공개키 암호알고리즘은 유한체 위에서 정의된 타원곡선 군에서의 이산 대수 문제에 기초한다. 타원곡선 암호 시스템은 1985년 N. Koblitz와 V. Miller에 의해서 처음 제시되었다. 타원곡선 암호 알고리즘은 비트 당 안전도가 타 공개키 시스템보다 효율적이라고 알려져 있다.

타원곡선상의 이산 대수 문제(Elliptic Curve Discrete Logarithm Problem)는 유한체위에 타원곡선 E가 정의되어 있을 때, 정수 a배의 관계에 있는 타원곡선 위의 두 점 P, Q를 모두 안다고 하더라도 정수 a를 계산하는 것이 어렵다는 점을 이용한 것이다.[6]

$$a * P = Q \ (P, Q \in E(F_q))$$

이 때 연산 $*$는 타원곡선 상의 배수 연산이고, P와 Q는 타원곡선 위의 점을 의미하고, a는 타원곡선상의 점이 아닌 일반 정수를 의미한다.

타원곡선 암호 시스템을 구현하기 위해서는 다양한 선택적 요소들이 있다. 타원곡선 암호시스템을 적용한 응용에 따라 이들 요소들을 적합하게 조합하는 것이 필요하다. 이들 파라미터에 따라서 안전성에 중점을 둘 것인지 성능에 중점을 둘 것인지를 결정할 수 있다. 파라미터들에는 타원곡선, 좌표계, 유한체, 기저 등이 있다.

첫째로 유한체에 대해서 살펴보면 실제 구현에 사용되는 유한체에는 표수(Characteristic)가 3보다 큰 소수인 F_p가 있으며, 표수가 2 또는 짝수인 F_{2^m}형태가 있다. 수학적인 연구에 의하면 2^m과

p가 비슷한 크기라면 그 위에 정의된 ECDLP는 대략 같은 시간 복잡도를 가지고 있다고 한다. 구현에 있어서는 하드웨어의 경우 유한체 F_{2^m}가 F_p에 비해서 효율적이다. 하지만 소프트웨어의 경우에는 F_p사용이 효율적일 수도 있다.

또한 보안적 측면에서 볼 때 유한체는 ECDLP의 복잡도에 악영향을 준다. 특별한 보안을 요구하는 응용에 있어서는 유한체 F_p의 사용이 효과적일 수도 있다.

실수에서 정의된 곡선을 사용하는 것이 아니라 유한체 위에 타원곡선을 사용하는 것은 유한개의 원소로 이루어지고 실수에서 정의된 것과는 달리 round off error가 발생하지 않기 때문에 암호학적 목적에 적합하기 때문이다.

특히 유한체 F_{2^m}의 경우에는 세 가지의 기저 중에 한 가지를 선택해야 한다. Polynomial basis 나 Normal basis나 Optimal normal basis 세 가지가 있는데 Normal basis 보다는 Optimal normal basis가 더 효율적이기 때문에 Polynomial basis 나 Optimal normal basis 중에서 한 가지 기저를 선택하여 유한체를 구성한다. 타원곡선 암호 시스템이 적용될 응용에 따라서 기저는 선택되어질 수 있을 것이다. Polynomial basis는 속도가 빠르지만 공간을 많이 차지하고, normal basis는 적은 공간을 사용하지만 역원을 구하는 속도가 느리다.

둘째로 타원곡선에 관해서 살펴보면, 타원곡선은 표수에 따라서 형태가 조금씩 차이가 있다. 체의 표수가 2와 3이 아닌 경우에는 타원 방정식은 다음과 같다.

$$y^2 = x^3 + ax + b \quad (\text{단, } 4a^3 + 27b^2 \neq 0)$$

체의 표수가 2인 경우에는 $y^2 + xy = x^3 + ax^2 b$또는 $y^2 + cy = x^3 + ax + b$이며 체의 표수가 3인 경우에는 $y^2 = x^3 + ax^2 + bx + c$이다.

표 1. 타원곡선의 덧셈과 두 배 연산 공식의 체 연산 횟수

operation	Doubling	Addition
Affine	11+2M+2S+8A	11+2M+S+6A
Projective ($Z_1 \neq 1$)	4M+4S+8A	12M+4S+9A
Projective ($Z_1 = 1$)	4M+4S+8A	8M+5S+9A

좌표계는 Affine coordinates와 Projective coordinates가 있는데 사용하는 좌표계에 따라서 타원곡선 위의 두 점의 덧셈 공식과 두 배 공식의 차이가 있다.

표 1은 두 좌표계에 따라 덧셈 공식과 두 배 공식의 체의 연산 횟수를 나타낸 것이다.

IV. 결 론

공개키 암호시스템은 대규모 네트워크 상에서의 정보 보호에 대해서 대칭키 암호시스템보다 더 효율적이고 관리하기 쉬운 암호 시스템이다. 최근 공개키 암호시스템의 구현에는 세가지 형태가 있다. 첫째 소인수 분해 문제를 이용한 RSA 시스템이고, 둘째는 이산 대수 문제를 이용한 Diffie-Hellman key exchange, ElGamal 시스템이 있다. 셋째 타원곡선 상의 이산대수 문제를 이용한 타원곡선 암호시스템이 있다. 이들 모든 시스템들은 기밀성, 인증, 데이터 무결성, 부인방지 기능을 제공할 수 있다.[7] 이들 세 가지 시스템 중에서 가장 효율적인 시스템은 높은 비트 당 안전도를 나타내고 있는 타원곡선 암호시스템이다. 타원곡선 암호시스템은 작은 크기의 키를 사용하여 다른 공개키 암호시스템과 같은 안전도를 나타내기 때문에 일반적인 응용뿐만 아니라 계산량 혹은 대역폭, 그리고 저장 공간 등의 시스템 환경적인 제약이 있는 특수한 응용에도 적합한 암호시스템이다.

참고문헌

[1] T. A. Berson : "Long key variants of DES," proc. crypto'82, pp. 311-314

[2] "Data Encryption Standard," National Bureau of Standard, Federal Information Processing Standards Publication 46, Jan. 1977

[3] "DES modes of operation," Federal Information Processing Standards Publication 81, National Bureau of Standards US Department of Commerce, 1980

[4] R. L. Rivest, A. Shamir and L. Adleman, "A method of obtaining digital signature and public key cryptosystem", ACM Communication 21 No.2, pp.120-126, 1978

[5] W. Diffie and M. E.Hellman, "New directions in cryptography", IEEE Trans. on Information Theory IT-22 No. 6, pp.644-654, 1976

[6] Certicom whitepaper, "Remarks On The Security of The Elliptic Curve Cryptosystem" pp.2-9, September 1997

[7] 권창영, 양형규, 원동호, "컴퓨터 통신 network를 위한 공개키 암호시스템에 관한 고찰", 한국통신학회 논문지, 제18권/8호, pp.1051-1058