부호기반 비동기식 스트림암호

이용우*, 노종선*

*서울대학교 전기정보공학부 뉴미디어통신공동연구소
*yongwool@ccl.snu.ac.kr, *jsno@snu.ac.kr

Code Based Asynchronous Stream Cipher

Yongwoo Lee*, Jong-Seon No*

*INMC, Department of ECE, Seoul National Univ.

요 약

본 논문에서는 부호기반 암호 시스템에서 코드 워드(codeword)에 더해지는 임의 오류(random error)를 키처럼 사용하는 방법을 제안한다. 이 방법을 이용하면 오류 정정 한계(error correction capability)를 넘는 큰 해밍 웨이트(Hamming weight)의 임의 오류가 더해진 암호문을 복호화할 수 있게 된다. 따라서 적은 오류 정정 한계를 가지는 부호를 이용하더라도 information set decoding 공격에 매우 안전한 부호 기반 암호를 설계할 수 있게 된다.

Ⅰ. 서 론

고파 부호(Goppa code)의 복호화 문제의 어려움에 기반을 둔 McEliece 암호시스템[1]이 제안되고 난 후, 이에 대한 다양한 공격 방법이 제안되었다. 그중 가장 유효한 것은 information set decoding 이다. Information set decoding 은[2] 암호문 중에서 오류가 포함되지 않은 부분을 찾아내어 복원하는 방법이다. 따라서 암호문에 더하는 임의 오류는 오류 정정 한계만큼의 해밍 웨이트를 가지는 것을 사용한다.

QC-LDPC 혹은 QC-MDPC 를 기반으로 하는 McEliece 암호의 변형도 발표되었는데, 여기에서 사용하는 부호의 특성상 최대로 정정할 수 있는 오류의 크기를 특정할 수가 없다. 따라서 높은 확률로 정정이 가능한 해밍 무게(t)를 가지는 오류를 암호화에 사용한다. 이때 t를 작게 하면 복호화 가능할 확률이 올라가지만, information set decoding 은 그만큼 쉬워지게 된다. 이뿐 아니라, 오류 정정 능력이 적은 부호의 경우 McEliece 암호 시스템에 적용이 어렵다.

본 논문에서는 이러한 오류의 한계를 해결하고자 McEliece 암호시스템에 사용되는 오류를 누적하여 사용하는 방법을 제안한다. 이 방법을 사용하면 오류의 해밍 무게가 평균적으로 코드 워드 길이의 절반이 되므로 information set decoding 을 적용할 수 없다. 기존에 사용된 오류들이 일종의 키처럼 사용되기 때문에, 비 동기식 스트림 암호의 형태를 띤다. 그러나 추가적인 키 교환 과정이 필요 없으며, 코드 워드의 길이만큼의 공간이 필요하다.

Ⅱ. McEliece 암호 시스템

McEliece 암호시스템은 다음과 같이 이루어진다. Bob 은 t 개의 오류를 정정할 수 있는 (n,k)고파 부호의 생성 행렬(generator matrix) $\mathbf{G}$ 와, $k \times k$ 가역행렬 $\mathbf{S}$, 그리고 $n \times n$ 순열 행렬 $\mathbf{P}$를 생성하여 비밀 키로 갖는다.

Bob 은 $\mathbf{G_{pub}} = \mathbf{SGP}$를 계산하고, t와 함께 공개 키로써 공개한다.

A. 암호화

Alice 는 다음의 방법으로 k 비트 메시지 m 을 암호화하여 Bob 에게 전송한다.

① 크기가 n, 해밍 웨이트가 t 인 임의 오류 e 를 생성한다.

② $\mathbf{c} = \mathbf{mG_{pub}} \oplus \mathbf{e}$를 구한다. c 가 암호문이다.

B. 복호화

Bob 은 Alice 로부터 받은 암호문 c를 아래와 같은 방법으로 복원할 수 있다.

① 암호문의 오른쪽에 $\mathbf{P}$ 의 역행렬을 곱하여 $\mathbf{cP^{-1}} = (\mathbf{mS})\mathbf{G} + \mathbf{eP^{-1}}$을 계산한다.

② G 에 대한 복호화(decoding) 알고리즘을 이용하여 $\mathbf{eP^{-1}}$을 제거한다.

③ 구한 $\mathbf{mS}$ 의 왼쪽에 $\mathbf{S}$ 의 역행렬을 곱해서 메시지 m 을 구한다.

Ⅲ. Information Set Decoding

Algorithm 1. *Information Set Decoding*

$\mathbf{c} \leftarrow$ *ciphertext*
$\mathbf{pos} \leftarrow$ *a set of* k *positions among 0~n-1*

$\mathbf{c}' \leftarrow$ del i − th bit *from* $\mathbf{c}$ ***for*** i ***not in*** $\mathbf{pos}$
$\mathbf{G}'_{\mathbf{pub}} \leftarrow$ del i − th column *from* $\mathbf{c}$ ***for*** i ***not in*** $\mathbf{pos}$

if $\mathbf{G}'_{\mathbf{pub}}$ is invertible:
$\mathbf{m} = \mathbf{c}'{\mathbf{G}'_{\mathbf{pub}}}^{-1}$

Information Set Decoding 은 오랜 기간 연구되어 다양한 변형이 있지만, 기본적인 방법은 ***Algorithm 1***.과 같다.

더해진 오류의 해밍 웨이트를 t라고 할 때, 위 방법의 성공 확률은 $\frac{\binom{n-t}{k}}{\binom{n}{k}}$으로, t가 줄어들면 공격 성공 확률이 크게 증가한다. 반대로 큰 해밍 웨이트를 가진 오류를 더하면 information set decoding 으로부터 안전한 부호 기반 암호를 설계할 수 있다.

Ⅳ. 부호 기반의 비동기식 스트림 암호

본 논문에서 제안하는 암호 시스템은 다음과 같은 방법으로 암호화와 복호화를 할 수 있다.

A. 암호화

길이가 n인 영벡터를 생성하여 $e_{cum\,ul}$이라 한다.

① 길이가 n, 해밍 웨이트가 t 인 임의의 오류 e 를 생성한다. $e_{cum,a} = e_{cum,a} + e$를 계산한다.

② $c = m\mathbf{G_{pub}} \oplus e_{cum.a}$를 구한다. c 가 암호문이다.

B. 복호화

길이가 n인 영벡터를 생성하여 $e_{cum,b}$이라 한다.

① $c' = c \oplus e_{cum,b}$을 계산하고, 오른쪽에 $\mathbf{P}$의 역행렬을 곱하여 $c'\mathbf{P^{-1}} = (m\mathbf{S})\mathbf{G} + (e_{cum.a} \oplus e_{cum.b})\mathbf{P^{-1}}$ 을 계산한다.

② G에 대한 복호화(decoding) 알고리즘을 이용하여 $(e_{cum.a} \oplus e_{cum.b})\mathbf{P^{-1}}$을 제거한다.

③ 구한 $m\mathbf{S}$의 왼쪽에 $\mathbf{S}$의 역행렬을 곱해서 메시지 m을 구한다.

④ $e_{cum.b} = e_{cum.a}$로 업데이트 한다.

상기한 방법을 이용하면 $(e_{cum.a} \oplus e_{cum.b})$ 는 ①에서 생성한 오류 e와 같기 때문에, 정정이 가능한 오류이다. 그러나, 몇 번의 통신을 거치고 나면 $e_{cum.a}$는 큰 헤밍 웨이트를 가지는 오류가 되기 때문에 information set decoding 을 적용하기가 어렵다.

그림 1.은 통신이 지속됨에 따라서 $e_{cum.a}$ 의 해밍 웨이트가 변하는 양상을 실험한 결과이다. [1]의 파라미터를 기준으로 실험하였다. 그 결과, 해밍 무게는 $\frac{n}{2}$ 에 수렴하고, 짧은 시간 안에 $n-k$ 를 넘어서 information set decoding 의 적용이 불가능하게 됨을 알 수 있다.

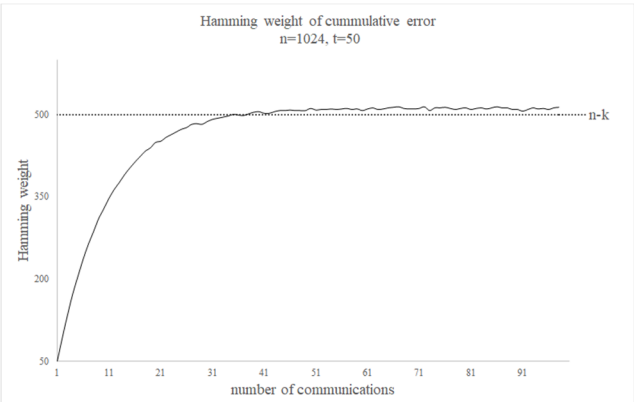

그림 1. 통신 횟수에 따른 오류의 해밍 웨이트

Ⅴ. 결론

본 논문에서는 McEliece 암호를 information set decoding 에 저항하도록 개선하는 방법을 제안한다. 이 방법은 추가적인 키 교환이 필요 없기 때문에 기존 암호 시스템과 크게 다르지 않다. 그러나 n -bit 만큼의 추가적인 공간이 필요하고, 오류를 계산해야 하므로 복호화 알고리즘에 따라 추가적인 연산을 필요로 할수도 있다. 또한, 스트림 암호이기 때문에 패킷 손실이 일어나면 오류의 정정이 불가능해져 새롭게 동기화 과정이 필요하다는 한계가 있다.

ACKNOWLEDGMENT

이 논문은 2017 년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (R-20160229-002941, IoT 및 클라우드 컴퓨팅을 위한 경량 포스트 양자 암호 시스템 연구)

참 고 문 헌

[1] R. J. McEliece, "A public-key cryptosystem based on algebraic coding theory," *Coding Thv,* vol. 4244, pp. 114-116, 1978.

[2] P. J. Lee and E. F. Brickell, "An observation on the security of McEliece's public-key cryptosystem," in *Eurocrypt*, 1988, vol. 88, no. 28, pp. 275-280: Springer.