Goppa 부호와 Patterson 디코딩에 관한 연구

전창열, 최장혁, 김동찬
국민대학교 정보보안암호수학과

{chjeon96, jhdh0813, dckim}@kookmin.ac.kr

A Study on Goppa Code and Patterson Decoding

Changyeol Jeon, Janghyuck Choi, Dong-Chan Kim
Kookmin University

요 약

NIST 는 2024 년 표준으로 제정할 양자 내성 암호 공모전을 진행하고 있다. 현재 3 라운드까지 진행되었고, 데이터암호화/키체결 알고리듬 분야에는 4 종의 암호가 최종후보로 선정되었다. 그 중 NTS-KEM 과 결합된 Classic McEliece 는 이진 Goppa 부호를 사용한다. 이진 Goppa 부호는 Patterson 디코딩 알고리듬을 통해 효율적인 복호화가 가능하다는 장점이 있다. 본 논문에서는 Goppa 부호의 정의와 이진 Goppa 부호의 복호화 알고리듬인 Patterson 디코딩에 대해 소개한다.

Ⅰ. 서론

현재 여러 국가와 기업이 양자 컴퓨터 개발에 많은 투자를 하고 있다. 기존의 이산대수, 인수분해 기반 암호 알고리듬은 지수 시간 복잡도를 가진다. 하지만 양자 컴퓨팅 환경에서는 Shor 의 소인수분해 알고리듬과 Grover 의 검색 알고리듬에 의해 다항식 시간 복잡도를 가지게 되어 더 이상 기존 기대 안전성을 보장할 수 없게 된다[2,5]. 이러한 이유로 양자 내성 암호의 연구가 활발히 진행되고 있다.

NIST 는 2021 년까지 양자 내성 암호 표준 제정 사업을 진행한다. 현재 3 라운드까지 진행되었고, 최종적으로 선정된 암호는 2024 년부터 국제 표준으로 사용된다. 데이터암호화/키체결 알고리듬 분야에는 4 종의 암호가 최종후보로 선정되었다. 그 중 NTS-KEM 과 결합된 Classic McEliece 는 1970 년 V. D. Goppa 가 제안한 Goppa 부호를 사용한다[1,3]. Goppa 부호 중 이진유한체에서 정의한 Goppa 부호는 Patterson 디코딩 알고리듬을 통해 효율적인 복호화가 가능하다는 장점이 있다[4].

본 논문의 구성은 다음과 같다. Ⅱ장에서 기호를 정의한다. Ⅲ장에서는 Goppa 부호, Ⅳ장에서는 이진 Goppa 부호를 소개한다. Ⅴ장에서는 이진 Goppa 부호의 효율적인 디코딩 알고리듬인 Patterson 디코딩 알고리듬에 대해 설명한다.

Ⅱ. 기호 및 정의

본 논문은 다음의 기호를 사용한다.

- q 소수의 거듭제곱
- $\mathbb{F}_{q^m}$ q^m개의 원소를 갖는 유한체, $m \geq 1$
- $\mathbb{F}_q^n$ $\mathbb{F}_q$에서 정의된 n차원 벡터공간.
- $supp(c)$ 벡터 c의 성분 중 0이 아닌 성분의 위치 집합
- $[n]$ $\{0,1,\dots,n-1\}$
- $\sqrt{K}$ $\mathbb{F}_{2^m}$ 에서 $K(\in \mathbb{F}_{2^m})$의 제곱근
- $\lfloor a \rfloor$ 실수 a에 대하여 a보다 크지 않은 최대 정수

Ⅲ. Goppa 부호

Goppa 부호는 다음 두 인스턴스를 사용한다.

- Support 집합 $L\,(\subseteq \mathbb{F}_{q^m})$: 서로 다른 $n(\leq q^m)$개의 원소 $\{\alpha_0,\dots,\alpha_{n-1}\}$를 갖는 집합.
- Goppa 다항식 $g(X) \in \mathbb{F}_{q^m}[X]$: L의 모든 원소를 근으로 갖지 않는 t차 다항식.

정의 1. Support 집합 L과 Goppa 다항식 $g(X)$로 Goppa 부호 $\mathcal{C}$를 다음과 같이 정의한다.

$$\mathcal{C} := \{c = (c_0,\dots,c_{n-1}) \in \mathbb{F}_q^n : \sum_{j=0}^{n-1} c_j (X-\alpha_j)^{-1} \equiv 0 \ (mod\ g(X))\}.$$

이때 $(X-\alpha_j)^{-1}$는 modulo $g(X)$에 대한 $X-\alpha_j$의 역원으로, $t-1$차 다항식이다.

정리 1. $\mathcal{C}$의 최소해밍거리 d와 $g(X)$의 차수 t는 다음을 만족한다.

$$d \geq t+1.$$

증명. 영벡터가 아닌 임의의 벡터 $c = (c_0,\dots,c_{n-1}) \in \mathcal{C}$는 다음이 성립한다.

$$\begin{aligned} 0 &\equiv \sum_{j\in[n]} c_j (X-\alpha_j)^{-1} \equiv \sum_{j\in supp(c)} c_j (X-\alpha_j)^{-1} \\ &\equiv \left(\prod_{i\in supp(c)} (X-\alpha_i)\right)\left(\sum_{j\in supp(c)} c_j (X-\alpha_j)^{-1}\right) \\ &\equiv \sum_{j\in supp(c)} c_j \left(\prod_{i\in supp(c)} (X-\alpha_i)\right)(X-\alpha_j)^{-1} \\ &\equiv \sum_{j\in supp(c)} c_j \left(\prod_{i\in supp(c),\ i\neq j} (X-\alpha_i)\right) (mod\ g(X)). \end{aligned}$$

다항식 $c_j\left(\prod_{i\in supp(c),\ i\neq j}(X-\alpha_i)\right)$는 $|supp(c)|-1$차이다. 만약 $|supp(c)|-1 < t$이면, 모든 c_j가 0이 되어야 하므로 모순이다. 따라서, $|supp(c)|-1 \geq t$를 만족해야 한다. c는 임의의 벡터이고, $|supp(c)|$는 벡터 c의 해밍거리이므로 d가 $t+1$이상이 되어야 한다. □

따라서 적절한 t를 선택하여 원하는 d를 가지는 부호를 생성할 수 있다.

Ⅳ. 이진 Goppa 부호

정의 2. $\mathcal{C}$가 표수 2인 유한체 $\mathbb{F}$에서 정의되고, $g(X) \in \mathbb{F}_{2^m}[X]$가 분해가능(separable)[1] 하다면 $\mathcal{C}$를 이진 Goppa 부호라고 정의한다.

유한체에서 모든 기약다항식은 분해가능하다. 따라서 일반적으로 이진 Goppa 부호는 기약다항식을 Goppa 다항식으로 사용한다.

이진 Goppa 부호의 최소해밍거리 범위는 유한체 미분을 사용하여 계산할 수 있다. 유한체에서의 미분은 다음과 같이 정의한다. 체 $\mathbb{F}$기반 다항식환 $\mathbb{F}[X]$의 n차 다항식 원소 $f(X) = a_0 + a_1X + \cdots + a_nX^n$ 가 존재할 때, $\frac{d}{dx}f(X)$는 $a_1 + 2a_2X + \cdots + na_nX^{n-1}$로 정의하고, $f'(X)$로 표기한다.

이때, 표수가 2 인 유한체에서는 $f'(X)$의 홀수 차수의 계수가 짝수이므로, 짝수 차수 항으로만 이루어지게 된다. 또한 $\mathbb{F} = \mathbb{F}_{2^m}$ 인 경우에는 $f'(X) = a_0 + a_2X^2 + a_4X^4 + \cdots + a_{2l}X^{2l} = \left(\sqrt{a_0} + \sqrt{a_2}X^1 + \cdots + \sqrt{a_{2l}}X^l\right)^2$과 같은 완전 제곱 식 형태로 표현이 가능하다.

정리 2. 이진 Goppa 부호의 최소해밍거리 d는 다음을 만족한다.

$$d \geq 2t + 1.$$

증명. 정리 1 에서 사용한 식을 이용한다. 이진 유한체에서 연산이 이루어 지고, $j \in supp(c)$ 이므로, c_j는 1 이다.

$$0 \equiv \sum_{j \in supp(c)} \prod_{\substack{i \in supp(c) \\ i \neq j}} (X - \alpha_i) \pmod{g(X)}.$$

상기 식은 $\frac{d}{dx}\prod_{i \in supp(c)}(X - \alpha_i)$와 같다. 동시에 이는 표수가 2 인 체에서 이루어 지는 미분이므로 완전 제곱 식 형태로 표현이 가능하다.

$$0 \equiv \frac{d}{dx}\prod_{i \in supp(c)} (X - \alpha_i) \equiv \sum_{i=0}^{l} b_iX^{2i} \equiv \left(\sum_{i=0}^{l} \sqrt{b_i}X^i\right)^2 \pmod{g(X)}.$$

이때, l은 $\left\lfloor \frac{|supp(c)-1|}{2} \right\rfloor$이다. l이 $g(X)$의 차수인 t보다 작으면, 모든 계수 $\sqrt{b_i}\ (i = \{0, 1, \ldots, l\})$가 0 이어야 한다. 따라서 l은 t 이상이어야 한다. $|supp(c)| - 1 \geq 2l$과 $l \geq t$를 통해 $|supp(c)| \geq 2t + 1$을 얻을 수 있다. $|supp(c)|$는 해밍거리이므로, $d \geq 2t + 1$가 성립한다는 것을 알 수 있다. □

Ⅴ. Patterson 디코딩 알고리듬

1975 년 Patterson 은 이진 Goppa 부호의 디코딩 알고리듬을 제안하였다. 이를 Patterson 디코딩 알고리듬이라고 한다.

Patterson 디코딩은 다음을 정의한다.

- 오류 위치 집합 E: 오류벡터 $\{e_0, \ldots, e_{n-1}\}$의 성분 중 0이 아닌 성분의 위치 집합
- 오류 위치 다항식 $\sigma(X)$: $\prod_{j \in E}(X - \alpha_j)$
- 오류 계산 다항식 $\omega(X)$: $\sum_{i \in E} e_i \prod_{j \in E \setminus \{i\}}(X - \alpha_j)$
- 신드롬 $s(X)$: $\sum_{j=0}^{n-1} y_j(X - \alpha_j)^{-1}\left(= \sum_{j \in E} e_j(X - \alpha_j)^{-1}\right)$, $y = \{y_0, \ldots, y_{n-1}\} \in \mathbb{F}_q^n$

정리 3. $\sigma(X)$, $\omega(X)$, $s(X)$는 다음을 만족한다.

$$\sigma(X)s(X) \equiv \omega(X) \pmod{g(X)}.$$

증명.

$$\begin{aligned}\sigma(X)s(X) &\equiv \left(\prod_{i \in E}(X - \alpha_i)\right)\left(\sum_{j \in E} e_j(X - \alpha_j)^{-1}\right) \\ &\equiv \sum_{j \in E} e_j\left(\prod_{i \in E}(X - \alpha_i)\right)(X - \alpha_j)^{-1} \\ &\equiv \sum_{j \in E} e_j\left(\prod_{i \in E \setminus \{j\}}(X - \alpha_i)\right) = \omega(X) \pmod{g(X)}.\end{aligned}$$ □

이때 이진 Goppa 부호는 $j \in E$ 인 경우 e_j가 1 이므로 $\omega(X) = \sigma'(X)$를 만족한다.

이진 Goppa 부호는 최대 t개의 오류를 정정할 수 있으므로 $\sigma(X)$를 다음과 같이 t차 이하 다항식으로 표현 가능하다.

$$\sigma(X) = \sigma_0 + \sigma_1X + \cdots + \sigma_tX^t, \qquad \sigma_j \in \mathbb{F}_{2^m}.$$

이는 홀수 차수 항과 짝수 차수 항으로 분리하여 표현이 가능하다. 또한, Patterson 디코딩은 $\mathbb{F}_{2^m}$에서 수행하므로 각각을 완전 제곱 식 형태로 표현이 가능하다.

$$\sum_{k=0}^{i} \sigma_{2k}X^{2k} = \left(\sum_{k=0}^{i} \sqrt{\sigma_{2k}}X^k\right)^2 \quad , i = \left\lfloor \frac{t}{2} \right\rfloor \quad ,$$

$$\sum_{k=0}^{i} \sigma_{2k+1}X^{2k+1} = \left(\sum_{k=0}^{i} \sqrt{\sigma_{2k+1}}X^k\right)^2 X \quad , i = \left\lfloor \frac{t-1}{2} \right\rfloor.$$

즉, $\sigma(X) = A(X)^2 + B(X)^2X$ 형태이다. 이후, $\sigma(X)$를 미분하였을 때, 계수가 짝수인 항은 소거된다.

$$\sigma'(X) = 2A(X)A'(X) + 2B(X)B'(X) + B(X)^2 = B(X)^2.$$

해당 식을 $\sigma(X)s(X) \equiv \omega(X) = \sigma'(X) \pmod{g(X)}$에 대입하면 $(A(X)^2 + B(X)^2X)s(X) = B(X)^2 \pmod{g(X)}$ 이다. $A(X)^2$ 에 관하여 식을 정리하면 다음과 같다.

$$(s^{-1}(X) + X)B(X)^2 \equiv A(X)^2 \pmod{g(X)}.$$

이때 $\mathbb{F}_{2^m}[X]/(g(X))$ 에서는 제곱근이 항상 존재하므로, $\sqrt{(s^{-1}(X) + X)}B(X) \equiv A(X) \pmod{g(X)}$ 로 표현이 가능하다. $g(x)$는 기약다항식이므로, $\sqrt{(s^{-1}(X) + X)}$와의 확장 유클리드 알고리듬을 통해 $A(X)$와 $B(X)$를 구할 수 있다. 동작과정에서 $A(X)$와 $B(X)$의 차수가 각각 $\left\lfloor \frac{t}{2} \right\rfloor$와 $\left\lfloor \frac{t-1}{2} \right\rfloor$를 넘지 않는 경우 그 즉시 확장 유클리드 알고리듬을 종료하고, 해당 다항식을 반환한다. 이를 통해 $\sigma(X)$를 구한다. 이때 $\sigma(\alpha_j) = 0\ (j \in [n])$를 만족하는 j를 구할 수 있다. 해당하는 j는 E의 원소가 되므로 오류를 정정하는 것이 가능하다.

α_j를 구하는 방법은 인수분해와 전수조사가 있다. 하지만 인수분해는 알고리듬 동작 시간 정보 등에 의해 부채널 공격에 노출될 수 있다. 따라서 전수조사 방법을 사용한다.

Ⅵ. 결론

본 논문에서는 McEliece 암호의 기반 부호인 Goppa 부호와 Patterson 디코딩 알고리듬을 소개하였다. 향후 Goppa 부호와 Patterson 디코딩 알고리듬의 효율적 구현 기법에 관한 연구를 진행할 예정이다.

참 고 문 헌

[1] V. D. Goppa, "A new class of linear correcting codes," Probl. Peredach. Inform., vol. 6, pp. 24–30, Sept. 1970

[2] L. K. Grover, "A fast quantum mechanical algorithm for database search," Proceedings, 28th Annual ACM Symposium on the Theory of Computing, pp. 212, May 1996

[3] NIST, "Post Quantum Cryptography – Round 3 Submissions," July 22, 2020.

[4] N. Patterson, "The algebraic decoding of Goppa codes," IEEE Trans. Inf. Theor., 21(2): 203–207, Sept. 2006

[5] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," IEEE Comput. Soc., pp.124–134, Nov. 1994

1) 분해가능(separable): 체 $\mathbb{F}$ 에 대해 $f(X) \in \mathbb{F}[X]$가 $f(X)$에 대한 분해체 상에서 서로 다른 해를 가지는 경우.