순환 부호를 이용한 6 보다 큰 해밍 거리의 이진 부분접속 복구 부호의 새로운 설계 기법

김찬기, 노종선

서울대학교 전기정보공학부

A New Construction of Binary Locally Repairable Code with Hamming Distance Larger Than 6 Using Cyclic Codes

Chanki Kim and Jong-Seon No

Department of Electrical and Computer Engineering, Seoul National University

carisis@ccl.snu.ac.kr, jsno@snu.ac.kr

요 약

본 논문은 최근 제시된 기법을 변형하여 이진 부분접속 복구 부호 중 순환 부호를 활용하여 6 보다 큰 해밍 거리를 설계할 수 있는 새로운 방법을 제시한다. 본 방법을 통해 설계된 부분 접속 복구 부호 중 몇몇은 이론적 상한에 가까운 값이 나타남을 보인다.

1. 서론

부분접속 복구 부호(locally repairable code, LRC)는 분산 저장 시스템(distributed storage system, DSS)을 위한 부호 중 하나로 논문 [1]에서 제시된 이래로 많은 연구가 있었다. 최근 이진 부분접속 복구 부호가 낮은 설계 복잡도의 장점으로 최근 연구가 활발히 이루어지고 있으며, 특히 해밍 거리가 4 혹은 6 정도의 경우에는 최적 혹은 최적에 가까운 여러 설계 기법들이 제안되었다. [2]-[4]. 그 중 논문 [4]는 최근, 순환 부호를 이용하여 다양한 길이의 최적 이진 부호를 만들 수 있는 여러 경우를 제시하였으며, 이 중 특수한 경우에 대해서는 논문 [5],[6]에서 유도된 상한(upper bound) 값을 만족하였다. 본 논문에서는 논문 [4]의 방법을 활용하여, 많은 연구가 필요한 6 보다 큰 해밍 거리인 부분 접속 복구 부호에 대해서 최적에 가까운 설계가 가능함을 보여주고자 한다.

2. 본론

A. 부분접속 복구 부호의 정의 및 최적성

(n, k, d, r) 부분접속 복구 부호는 부호 길이 n, 메시지 길이 k, 해밍 거리 d, 부분 접속수 r의 파라미터로 표현된다. 해당 부분접속 복구 부호는 해당 부호의 쌍대 부호(dual code) $C^{\perp}$의 최소 해밍 거리인 $d^{\perp}$가 일정 부분 접속수 $r+1$값보다 작거나 같아야 하며, 모든 부호 원소가 속한 위치에 위의 조건을 만족하는 쌍대 부호의 부호어(codeword)가 존재해야 한다. 부분접속 복구 부호의 최적성을 보이는 방법은 고정된 파라미터 중 하나의 파라미터의 상한 혹은 하한을 보이는 것으로서, 본 논문에서는 메시지 길이 k에 대한 상한과의 가까움을 보이는 것으로 보이고자 한다.

B. 제안하는 설계 기법

본 논문에서 제안하는 설계법은 아래와 같다.

설계 1. ($r=2$ 및 높은 해밍 거리를 위한 부분접속 복구 부호 설계) 이진 체 상에서 홀수인 d'에 대해 $(n, k, d) = (2m+1, m+1, d')$이며 order 가 $2m$인 원시 원소(primitive element) α의 최소 다항식(minimal polynomial)이며 degree 가 m인 생성 다항식(generator polynomial)로 갖는 순환 부호 C'가 존재한다고 가정하자. 그 경우 첫번째 정보 비트를 shortening 하여 $(n, k) = (2m, m)$를 만들 수 있다. 그 이후 앞서 생성된 부호를 외부 부호로 패리티 검사 다항식(parity check polynomial)이 $x^{2m}+x^{m}+1$ 인 $(n, k) = (3m, 2m)$ 순환 부호를 내부 부호로 연접하여 최종적으로 $(n, k) = (3m, m)$을 따르는 부호 C를 만든다. 그 부호의 부분 접속수 r은 2이며, 해밍 거리 d는 $d \geq d' + 3$이다.

증명) 먼저 최종적인 부호가 $r=2$인 이유는 외부 부호인 $x^{2m}+x^{m}+1$에 의해 최종적인 패리티검사 행렬이 아래와 같이 나타나기 때문이다.

$$H = \begin{bmatrix} I_{C'} & P_{C'} & O \\ I_m & I_m & I_m \end{bmatrix}$$

아래 행에 존재하는 3 개의 $(m \times m)$ 항등 행렬 I_m

에 의해 모든 위치의 쌍대부호의 해밍 거리가 3 인 부호어가 존재함을 만족함을 확인할 수 있다.

또한 해밍 거리가 $d \geq d' + 3$임을 증명하자. 이진 부호에서 H의 아래 행을 모두 합한 선형 결합한 $[1,1,\ldots,1]$에 의해 d은 짝수밖에 존재할 수 없다. 따라서 짝수 경우인 해밍 거리가 $d' + 1$의 무존재성만 증명하면 된다. H의 제약조건을 만족시키며 해밍 거리가 $d' + 1$인 경우는 아래의 두 경우만 존재한다.

1) 전체 부호원소 중 앞 $2m$ 비트에서 0 이 아닌 원소 수가 d', 뒤 m 비트에서는 1 인 경우

2) 앞 $2m$ 비트에 위치한 비트에서 0 이 아닌 원소 수가 $d' + 1$인 경우

1)의 경우, 외부 부호의 부호어 다항식 $c(x) = (1 + x^m)p(x) + x^r$에 대해 $\mathrm{wt}(p(x)) = (d' - 1)/2$ 이며, $r, \deg(p(x)) < m$로 표현된다. 여기서 내부 부호의 해(zero)인 α^2에 대해 신드롬 값인 $c(\alpha^2) = \alpha^{2r} \neq 0$이므로 이 경우는 부호어가 존재하지 않는다.

2) 의 경우 외부 부호의 부호어 다항식은 $c(x) = (1 + x^m)p(x)$, $\mathrm{wt}(p(x)) = (d' + 1)/2$ 이며, $\deg(p(x)) < m$로 표현된다. 이 경우 내부 부호인 해 α에 대해, $c(\alpha) = (1 + \alpha^m)p(\alpha)$이며, $(1 + \alpha^m) \neq 0$임은 자명하다. 또한 $\deg(p(x)) < m$인 $p(x)$에서 α가 해인 경우, 앞서 가정한 내부 부호의 최소 다항식인 가정과 모순되므로 $p(\alpha) \neq 0$이다. 따라서 $c(\alpha) \neq 0$ 이므로, 이 경우에도 부호어가 존재하지 않는다. 가능한 모든 두 가지 경우에 대해 신드롬이 0 이 아니므로, 따라서 해밍 거리가 $d' + 1$인 부호는 존재하지 않는다. 따라서 최종적인 해밍 거리는 $\mathrm{d}' + 3$ 이상이 되어야 한다. □

C. 최적성 분석

제안된 부호 중 몇몇의 최적성을 분석하기 위하여 두 가지 상한이 사용된다.

상한 1. (C-M 상한 [5]) (n, k, d, r)인 부호에 대해 다음 부등식을 만족한다.

$$k \leq \min_{x \in Z^+} \left\{ xr + k_{opt}^{(q)}(n - x(r+1), d) \right\}$$

$k_{opt}^{(q)}(a, b)$는 이진 체 상에서 부호길이 a 및 해밍 거리 d인 부호의 최대 메시지 길이이다.

상한 2. (L-space 상한 [6]) $(n, k, d \geq 9, r = 2)$ 이며 disjoint repair group 인 부호에 대해 다음 부등식을 만족한다.

$$k \leq \frac{2n}{3} - \log_2\left(1 + n(n - r)\right)$$

설계 1 중 두 가지 예시에 대한 부호에서 앞서 소개한 두 상한을 활용하여 최적성 분석을 수행하고자 한다.

예시 1. $(n, k, d) = (17,9,5)$ 이며 생성 다항식이 $g(x) = x^8 + x^5 + x^4 + x^3 + 1$인 순환 부호는 설계 1 에서 제시한 조건을 만족하며, 따라서 외부 부호로 이용하여 설계 1 로 수행한 경우, 최종적인 부분 접속 부호는 $(n, k, d, r) = (24,8,8,2)$가 된다. 상한 1 에서 나오는 메시지 길이는 $k \leq 10$이므로 제안한 설계는 상한 대비 2 의 차이가 발생한다.

예시 2. $(n, k, d) = (23,12,7)$ 인 이진 Golay 부호는 순환 부호이며 또한 설계 1 의 조건을 만족한다. 이 부호를 외부 부호로 이용하여 설계 1 로 수행한 경우, 최종적인 부분 접속 부호는 $(n, k, d, r) = (33,11,10,2)$이다. 상한 1 에서 나오는 값은 $k \leq 13$이나 상한 2 에서 나오는 메시지 길이 값은 $k \leq 12$가 되어 제안된 설계는 상한 대비 1 의 차이가 발생한다.

3. 결론

본 논문에서는 논문 [4]의 방법을 활용하여, 많은 연구가 필요한 6 보다 큰 해밍 거리인 부분 접속 복구 부호 설계 방법을 제안하였고, 또한 제안된 부호가 기존 상한 대비 가까운 값을 보여줌을 보였다. 차후 연구로서 제안된 방법의 조건의 성립하는 외부 부호의 조건이나 일반적인 경우의 최적성 분석이 이루어질 예정이다. 그 외 앞서 제안된 것과는 다른 방법을 통해서도 높은 값의 해밍 거리에서 부분 접속 복구 기법의 설계 방법의 연구가 수행될 예정이다.

ACKNOWLEDGEMENT

이 논문은 2018 년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. NRF-2016R1A2B2012960).

4. 참고 문헌

[1] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, "On the locality of codeword symbols," *IEEE Trans. Inf. Theory*, vol. 58, no.11 pp.6925-6934, Nov. 2012.

[2] S. Goparaju and R. Calderbank, "Binary cyclic codes that are locally repairable," *in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Honolulu,* 2014. pp. 676-680.

[3] M. Y. Nam and H. Y. Song, "Binary locally repairable codes with minimum distance at least 6 based on partial *t*-spreads," *IEEE comm. Lett.* Vol. 21, no. 8, pp. 1683-1686, Apr. 2017.

[4] C. Kim and J.-S. No, "New constructions of binary and ternary locally repairable codes using cyclic codes," *IEEE comm. Lett.* Vol. 22, no. 2, pp. 228-231, Feb. 2018.

[5] V. Cadambe and A. Mazumdar, "Bounds on the size of locally recoverable codes," *IEEE Trans. Inf. Theory*, vol.61, no.11, pp. 5785-5794, Nov. 2015.

[6] A. Wang, Z. Zhang, and D. Lin, "Bounds and constructions for linear locally repairable codes over binary fields," *in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Aachen,* 2017. pp. 171-175.