Generalized construction of cyclic (r, t) -locally repairable codes using trace function

Quang Thanh Pham^{a)} and Hideki Yagi^{b)}

Dept. of Communication Engineering and Informatics, The University of Electro-Communications, 1–5–1 Chofugaoka, Chofu, Tokyo 182–8585, Japan

a) phamthanh1312@gmail.com

b) h.yagi@uec.ac.jp

Abstract: With an increasing demand for distributed systems storing big data, locally repairable codes (LRCs) have attracted attentions to recover lost data on failure nodes. An (r, t) -LRC has the property that each coordinate of codewords can be recovered from at most r other coordinates (repair set), and there are at least t disjoint repair sets for each coordinate, where r and t are called locality and availability, respectively. Wang, Zhang, and Lin recently have proposed a construction method of cyclic (r, t) -LRCs that can have high availability with a guaranteed lower bound on the minimum distance. However, due to the lack of flexibility in code design, it is impossible to modify code parameters while keeping the code performance. This letter presents a generalized construction of cyclic (r, t) -LRCs based on the trace function whose high-degree terms are truncated. It is shown that the proposed construction preserves the symbol-repairing performance and therefore provides better flexibility in code design.

Keywords: locally repairable code, cyclic code, availability, trace function

Classification: Fundamental Theories for Communications

References

- [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. DOI:10.1109/TIT.2012.2208937
- [2] D. S. Papailiopoulos and A. G. Dimakis, “Locally repairable codes,” *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 5843–5855, Oct. 2014. DOI:10.1109/TIT.2014.2325570
- [3] A. S. Rawat, D. S. Papailiopoulos, A. G. Dimakis, and S. Vishwanath, “Locality and availability in distributed storage,” *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 681–685, Honolulu, Hawaii, USA, June–July 2014. DOI:10.1109/ISIT.2014.6874919
- [4] P. Huang, E. Yaakobi, H. Uchikawa, and P. H. Siegel, “Binary linear locally repairable codes,” *IEEE Trans. Inf. Theory*, vol. 62, no. 11, pp. 6268–6283, Nov. 2016. DOI:10.1109/TIT.2016.2605119
- [5] A. Wang, Z. Zhang, and D. Lin, “Two classes of (r, t) -locally repairable codes,” *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 445–449, Barcelona, Spain, July 2016. DOI:10.1109/ISIT.2016.7541338

- [6] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North Holland, 1983.
- [7] J. L. Massey, *Threshold Decoding*, MIT Press, 1963.

1 Introduction

Locally repairable codes (LRCs) [1, 2] have recently attracted considerable attention as a promising technology for data protection. An (r, t) -LRC is a code having the following property: each coordinate of codewords can be recovered from at most r other coordinates (called a **repair set**), and there are at least t disjoint repair sets for each coordinate, where r and t are called **locality** and **availability**, respectively [3]. Generally, it requires a large amount of computation to evaluate locality r and availability t for a given linear code. Huang, Yaakobi, Uchikawa, and Siegel [4] showed that a class of cyclic LRCs offers an efficient way to evaluate locality r and availability t . Wang, Zhang, and Lin [5] introduced p -ary cyclic (r, t) -LRCs using the trace function over a finite field [6], where p is a prime. However, for a fixed r , it is impossible to modify code length n and dimension k , which leads to the lack of flexibility in code design.

This letter presents a generalized construction of cyclic (r, t) -LRCs based on the trace function whose high-degree terms are truncated. The generalized construction of cyclic (r, t) -LRCs involves another parameter $\delta \geq 1$, which enables to make the code length and the dimension variable while keeping parameters r , t , and the information rate $\frac{k}{n}$, and thus provides a wider class of cyclic (r, t) -LRCs.

2 Generalized construction

In the proposed construction, cyclic (r, t) -LRCs are constructed using the trace function over a finite field with deleted high-degree terms. Let p and e be a prime and a positive integer, respectively. We denote by $\mathbb{F}_p$ the finite field with p elements. Let $\mathbb{F}_p[x]/(x^n - 1)$ be the ring of polynomials over $\mathbb{F}_p$ with degree less than n . For a polynomial $a(x) \in \mathbb{F}_p[x]/(x^n - 1)$, let $\deg(a)$ denote the degree of $a(x)$. Let $\langle a(x) \rangle$ denote the ideal generated by $a(x)$ in the ring $\mathbb{F}_p[x]/(x^n - 1)$. An $[n, k]$ p -ary cyclic code $\mathcal{C} = \langle g(x) \rangle$ is the ideal generated by $g(x)$ in the ring $\mathbb{F}_p[x]/(x^n - 1)$, where $g(x) | x^n - 1$ and $\deg(g) = n - k$. The polynomial $g(x)$ is called the **generator polynomial**, whereas $h(x) \equiv \frac{x^n - 1}{g(x)}$ is the **parity check polynomial** of $\mathcal{C}$ with $\deg(h) = k$.

We use integers m and δ as parameters, where $1 \leq \delta \leq m$. Let $q = p^e$, the code length n is set as $n = q^m - q^{\delta-1}$, and we define the $(\delta - 1)$ -truncated trace function as follows:

$$f_\delta(x) = x + x^q + \cdots + x^{q^{m-\delta}}.$$

When $\delta = 1$, $f_1(x)$ reduces to the ordinary trace function [6], which is a mapping from $\mathbb{F}_{q^m}$ to $\mathbb{F}_q$, and its $\delta - 1$ terms from the highest degree are truncated to give $f_\delta(x)$. In the proposed construction, the parity check polynomial and the generator polynomial are defined as follows:

$$h(x) = \frac{f_{\delta}(x)^{q^{\delta-1}}}{x^{q^{\delta-1}}}, \quad g(x) = \frac{x^n - 1}{h(x)} = f_{\delta}(x)^{q^{\delta} - q^{\delta-1}} - 1.$$

It is easy to verify that $k = \deg(h) = q^{m-1} - q^{\delta-1}$, and thus the obtained code $\mathcal{C}$ is an $[n = q^m - q^{\delta-1}, k = q^{m-1} - q^{\delta-1}]$ p -ary cyclic code. When $\delta = 1$, the constructed code $\mathcal{C}$ is the cyclic LRC given by the construction of Wang et al. [5].

The following theorem shows the parameters of the (r, t) -LRCs obtained by the proposed construction.

Theorem 1 *The constructed $[n = q^m - q^{\delta-1}, k = q^{m-1} - q^{\delta-1}]$ p -ary cyclic code $\mathcal{C}$ is an (r, t) -LRC with $r = m - \delta$ and $t = e(m - \delta + 1)$.*

(Proof) The proof proceeds along the same line as in the proof of [5, Theorem 3]. Consider the following e polynomials in $\langle h(x) \rangle$:

$$\begin{aligned} f_{\delta}(x)^{q^{\delta-1}} &= x^{q^{\delta-1}} + x^{q^{\delta}} + \cdots + x^{q^{m-1}}, \\ f_{\delta}(x)^{pq^{\delta-1}} &= x^{pq^{\delta-1}} + x^{pq^{\delta}} + \cdots + x^{pq^{m-1}}, \\ &\vdots \\ f_{\delta}(x)^{p^{e-1}q^{\delta-1}} &= x^{p^{e-1}q^{\delta-1}} + x^{p^{e-1}q^{\delta}} + \cdots + x^{p^{e-1}q^{m-1}}. \end{aligned}$$

By cyclically shifting each of these polynomials, we obtain $m - \delta + 1$ polynomials containing the term $x^{q^{\delta-1}}$. Therefore, there are in total $t = e(m - \delta + 1)$ polynomials containing the term $x^{q^{\delta-1}}$, and these t polynomials are expressed in the form of

$$\{f_{\delta}(x)^{p^i q^{\delta-1}} x^{-p^i q^j + q^{\delta-1}}\}_{0 \leq i \leq e-1, \delta-1 \leq j \leq m-1}, \quad (1)$$

where the power of each term is taken to modulo n . Using the method developed in [5, Proof of Theorem 3], we can prove that $x^{q^{\delta-1}}$ is the only common term between any pair of these t polynomials.

These t polynomials in $\langle h(x) \rangle$ form t parity check equations of $\mathcal{C}$, each of which has $m - \delta + 1$ terms with the common term $x^{q^{\delta-1}}$. Therefore, the locality and the availability for the term $x^{q^{\delta-1}}$ are $r = m - \delta$ and $t = e(m - \delta + 1)$, respectively. Because of the symmetry of the argument, $\mathcal{C}$ is an (r, t) -LRC with $r = m - \delta$ and $t = e(m - \delta + 1)$. ■

Example 1 Suppose $\delta = 2$, then we have the code length $n = q^m - q$, the $(\delta - 1)$ -truncated trace function $f_2(x) = x + x^q + \cdots + x^{q^{m-2}}$, locality $r = m - 2$, and availability $t = e(m - 1)$.

Choosing $m = 4$, $p = e = 2$, $q = 4$ and $f_2(x) = x + x^4 + x^{16}$, a $[252, 60]$ binary cyclic code $\mathcal{C}$ is obtained. Consider the following $e = 2$ polynomials in $\langle h(x) \rangle$:

$$\begin{aligned} f(x)^4 &= x^4 + x^{16} + x^{64}, \\ f(x)^{2 \cdot 4} &= x^8 + x^{32} + x^{128}. \end{aligned}$$

By cyclically shifting each of these polynomials, we obtain in total the following $e(m - \delta + 1) = 6$ polynomials containing the term x^4 as shown in (1):

$$\begin{aligned} f(x)^4 &= \underline{x^4} + x^{16} + x^{64}, & x^{-12} f(x)^4 &= \underline{x^4} + x^{52} + x^{244}, \\ x^{-60} f(x)^4 &= \underline{x^4} + x^{196} + x^{208}, & x^{-4} f(x)^8 &= \underline{x^4} + x^{28} + x^{124}, \\ x^{-28} f(x)^8 &= \underline{x^4} + x^{152} + x^{232}, & x^{-124} f(x)^8 &= \underline{x^4} + x^{132} + x^{160}. \end{aligned}$$

Therefore, this code has locality $r = m - 2 = 2$ and availability $t = e(m - 1) = 6$. The cyclic code $\mathcal{C}$ is a $(2, 6)$ -LRC. ■

In addition to the locality, the availability, and the information rate, it is preferable to have a large minimum distance. We analyze the minimum distance of the constructed codes.

Theorem 2 *The minimum distance d of the constructed (r, t) -LRC $\mathcal{C}$ satisfies $d \geq \max\{q + 1, t + 1\}$.*

(Proof) To prove $d \geq t + 1$, note that there are $t = e(m - \delta + 1)$ polynomials that contain the common term $x^{q^{\delta-1}}$ as shown in (1). Each of these polynomials corresponds to a row in a parity check matrix of a **one-step majority-logic decodable code** [4]. It is well-known that the minimum distance d of such a majority-logic decodable code is at least $t + 1$ [7].

To prove $d \geq q + 1$, we evaluate the number of consecutive zeros of the generator polynomial $g(x)$ and then apply the BCH bound [6]. We interpret $x^n - 1$ as a polynomial over $\mathbb{F}_{q^w}$, where $w \equiv m - \delta + 1$, and we notice that

$$x^n - 1 = x^{q^m - q^{\delta-1}} - 1 = (x^{q^{m-\delta+1}-1} - 1)^{q^{\delta-1}}. \quad (2)$$

We seek zeros of the polynomial $l_\delta(x) \equiv x^{q^{m-\delta+1}-1} - 1$. Let $n_0 \equiv q^{m-\delta+1} - 1 = q^w - 1$. Then, it is well-known that $l_\delta(x)$ can be factorized as

$$l_\delta(x) = x^{n_0} - 1 = (x - \alpha^0)(x - \alpha^1) \dots (x - \alpha^{q^{m-\delta+1}-2}) \quad (3)$$

with some $\alpha \in \mathbb{F}_{q^w}$, where $\alpha^0, \alpha^1, \dots, \alpha^{q^{m-\delta+1}-2}$ are all distinct [6]. Therefore, it follows from (2) and (3) that the polynomial $x^n - 1$ has $n_0 = q^{m-\delta+1} - 1$ consecutive zeros with multiplicity $q^{\delta-1}$. This means that $h(x)$ has $\frac{k}{q^{\delta-1}}$ zeroes because $h(x) = (\frac{f_\delta(x)}{x})^{q^{\delta-1}}$ and $k = \deg(h(x))$. Since $x^n - 1 = g(x)h(x)$, this in turn indicates that the generator polynomial $g(x)$ has $n_0 - \frac{k}{q^{\delta-1}}$ zeroes. Then by the pigeonhole principle, $g(x)$ has at least $\lceil \frac{n_0 - \frac{k}{q^{\delta-1}}}{\frac{k}{q^{\delta-1}}} \rceil = \lceil \frac{n-k}{k} \rceil = q$ consecutive zeros. Therefore, the inequality $d \geq q + 1$ follows directly from the BCH bound. ■

3 Concluding remarks

In this letter, we have proposed a construction method for cyclic (r, t) -LRCs based on the $(\delta - 1)$ -truncated trace function $f_\delta(x)$, which can be obtained by truncating $(\delta - 1)$ terms of the ordinary trace function $f_1(x)$ from the highest degree. Theorem 1 indicates that the information rate of the constructed code is evaluated as

$$\frac{k}{n} = \frac{q^{m-\delta} - 1}{q^{m-\delta+1} - 1} = \frac{q^r - 1}{q^{r+1} - 1}.$$

For a fixed r , the information rate no longer depends on δ and m . Making two parameters δ and m variable while fixing $r = m - \delta$, we can modify the code length n and the dimension k while the information rate and (r, t) still remain.

When $\delta = 1$, the constructed cyclic (r, t) -LRCs reduce to the code given by Wang et al. [5]. A comparison between two constructions is shown in Table I.

Table I. A comparison of two code constructions

	Construction by Wang et al. [5]	Proposed Construction
Locality r	$m - 1$	$m - 1$
Availability t	em	em
Code length n	$q^m - 1$	$q^{m+\delta-1} - q^{\delta-1}$
Dimension k	$q^{m-1} - 1$	$q^{m+\delta-2} - q^{\delta-1}$
Information rate	$\frac{q^{m-1}-1}{q^m-1}$	$\frac{q^{m-1}-1}{q^m-1}$

Here, we use $m + \delta - 1$ in place of m in the proposed generalized construction so that locality r is set to be equal in both constructions for a fair comparison. From Table I, we see that the proposed generalized construction of cyclic (r, t) -LRCs keeps r , t , and the information rate from the code given by Wang et al. [5]. Therefore, the proposed construction of cyclic (r, t) -LRCs has a better flexibility in code design while keeping the symbol-repairing performance. It can achieve high availability with a guaranteed lower bound on the minimum distance shown in Theorem 2. The obtained code is a one-step majority-logic decodable code, whose usefulness as an (r, t) LRC has already been pointed out in [4]. However, unlike the construction method developed in [4], the parity check matrix of the obtained code does not have a tensor product structure.

As is shown in Table I, letting $\delta \geq 2$, we obtain an (r, t) -LRC with the code length $q^{\delta-1}$ times as large as the code for $\delta = 1$. However, Theorem 2 only states that the lower bounds on the minimum distance are the same in both cases. Investigating if the case $\delta > 2$ actually gives a larger minimum distance remains as a future study.