Locality of Some Optimal Ternary Codes

Luobin Guo¹, Qiang Fu¹, Gang Chen² and Ruipan Yang³

¹Department of Basic Sciences, Air Force Engineering University, Xi’an, Shaanxi, China
²Integrated Information Service Team (75837), the Southern War Zone, Guangzhou, China
³Air Force Aviation University (95937), Shenyang, Liaoning, China
Email: fuqiangkgd@163.com

Abstract. The locality of a locally repairable code for a distributed storage system is the number of nodes that participate in the repair of failed nodes, which characterizes the repair cost. In this paper, using two ternary codes [14,6,6] and [19,6,9], we first construct optimal codes with locality \( r=2 \). Then by constructing different anticodes using ternary codes with low dimensions, ten optimal codes with locality \( r=2 \) are obtained. Finally, we derive three optimal codes from every optimal code we have constructed and determine their locality \( r=2 \).

1. Introduction

For a distributed storage system (DSS), replication is the simplest and most widespread technique to ensure resilience to node failures. This strategy decreases the storage efficiency. Locally repairable codes (LRCs) are a family of erasure codes that can recover any code symbol by accessing other fixed code symbols. This will enable to recover the data stored in failure nodes locally, by accessing other nodes in the system.

LRC was firstly introduced by Gopalan et al. in [1]. Gopalan et al. introduced basic properties of LRC and proposed an upper bound:

\[
d \leq n - k + 2 - \left\lceil \frac{k}{r} \right\rceil.
\]  

(1.1)

If \( r=k \), the bound (1.1) in [1] becomes the classical Singleton bound. An LRC attaining this bound with equality is an optimal LRC. A new upper bound considering the size of the alphabet of field was established by Cadambe and Mazumdar in [2]

\[
k \leq \min_{r \in \mathbb{N}} \{ tr + k^{(q)}_{\text{opt}}(n - t(r + 1), d) \},
\]  

(1.2)

where \( k^{(q)}_{\text{opt}}(n,d) \) is the largest possible dimension of a code of length \( n \), for a given size \( q \) of a field and a given minimum distance \( d \). A more general bound considering availability \( t \) is presented in [12]. There exist lots of work about LRCs attaining the bound (1.1) and (1.2), such as in [13-16] and references therein. However, these LRCs are all constructed over large finite field \( F_q \), where \( q \) is an exponential function of \( n \) or greater than code length \( n \). In practice, considering fast arithmetic and backward compatibility with existing hardware, LRCs over small field are the best choice. LRCs over \( F_2 \) can be seen in [5-8]. Refs.[9-10] presented the construction of optimal ternary LRCs with...
dimension \( k \leq 5 \). Optimal \([n,6]\) codes with \( n \leq 324 \) are proposed in [11]. Ref. [18] proposed some optimal LRCs over \( F_3 \). In this paper, using two ternary codes with short lengths and the concept of anticodes, we will construct 50 optimal ternary codes with \( n \geq 324 \) and determine their locality \( r=2 \).

This paper is organized as follows: In Section 2, some preliminaries and notations are introduced. In Section 3, we propose the construction of 50 optimal ternary codes and analyze the locality of these optimal ternary codes. The last section is conclusion.

2. Preliminaries and Notations

First let \( F_3 \) be the finite field with 3 elements and \( F_3^n \) be the \( n \)-dimensional row (or column) vector space over \( F_3 \). For \( x=(x_1,x_2,\ldots,x_n), y=(y_1,y_2,\ldots,y_n) \in F_3^n \), the (Hamming) distance between \( x \) and \( y \) is \( d(x,y)=|\{i \mid x_i \neq y_i\}| \). The weight of \( x \) is \( wt(x)=|\{i \mid x_i \neq 0\}|=d(x,0) \).

An \([n,k,d]\) code \( C \) means a \( k \)-dimensional subspace of \( F_3^n \) with minimum distance \( d \), where \( d=\min\{d(x,y) \mid x \neq y, x, y \in C\} =\min\{wt(x) \mid x \in C, x \neq 0\} \). An \([n,k,d]\) code \( C \) is \textit{optimal} if there is no \([n,k,d+1]\) code.

For an \([n,k,d]\) code \( C \), a generator matrix \( G \) is a \( k \times n \) matrix over \( F_3 \), whose \( k \) rows form a basis of \( C \). Define the dual of \( C \) by \( C^\perp =\{x \in F_3^n \mid \mathbf{x} \cdot \mathbf{c} = 0 \forall \mathbf{c} \in C\} \). We say a generator matrix \( H \) of \( C^\perp \) is a parity check matrix of \( C \). We treat a generator matrix \( G \) of a linear code \( C \) as a set of column vectors. If a column vector \( g \) is included in \( G \), denote by \( g \in G \).

Locally repairable codes (LRCs) are a class of codes designed for the local correction of erasure [1]. For \( c=(c_1,c_2,\ldots,c_r) \in C \), the \( i \)-th code symbol \( c_i \) is said to have \textit{locality} \( r \) if \( c_i \) can be recovered by accessing at most \( r \) other code symbols.

Definition 2.1. A code \( C=[n,k,d] \) is said to have \textit{locality} \( r \) if all its symbols have locality \( r \).

Let \([n]\) be the set \([1,2,\ldots,n]\). The locality of \( C=[n,k,d] \) can be judged by its generator matrix as follows:

Lemma 2.1. [4] Let \( G=(g_1,g_2,\ldots,g_n) \) be a generator matrix of \( C=[n,k,d] \). For each \( i \in [n] \), there is a set \( A_i \subseteq [n] \setminus \{i\} \) of size at most \( r \), such that \( g_i \) is a linear combination of the columns \( g_j \) for \( j \in A_i \), then \( C \) has locality \( r \).

Let \( 1_0=(1,1,\ldots,1) \) and \( 0_0=(0,0,\ldots,0) \) be the all one vector and the all zero vector of length \( n \), respectively. In order to present our construction of LRCs, we give the construction of Simplex \([(3^k-1)/2,k,3^k-1] \) codes as follows:

Let

\[
S_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad S_3 = \begin{pmatrix} S_2 & 0_2 & S_2 & S_2 \\ 0_4 & 1 & 1_4 & 2_4 \end{pmatrix}
\]

and recursively, we have

\[
S_{k+1} = \begin{pmatrix} S_k & 0_k & S_k & S_k \\ 0_{N_k} & 1 & 1_{N_k} & 2_{N_k} \end{pmatrix},
\]

where \( N_k = (3^k-1)/2 \).

We rewrite Lemma 2.2 in [9] as follows:

Lemma 2.2. Let \( 1 \leq i \leq N_k \) and \( M = \begin{pmatrix} S_k \\ 0 \end{pmatrix} \). Then \( G_{m+1} = \begin{pmatrix} \alpha_i & \cdots & \alpha_i \\ 0 & \cdots & 0 \end{pmatrix} \) generates a linear code with \( r=2 \) where \( \alpha_i \in F_3 \).
Proof. For $\forall \alpha$, with $1 \leq i \leq s$, we have\[
\begin{pmatrix}
\alpha
\end{pmatrix}_0 + x_i \begin{pmatrix}
\alpha_i
\end{pmatrix}_1 + 2 \begin{pmatrix}
\alpha
\end{pmatrix}_x = 0_{\alpha+1}. For \beta \in S_k$ and $\begin{pmatrix}
\beta
\end{pmatrix}_y \notin M$, then we have $\begin{pmatrix}
0
\end{pmatrix}_x + 2 \begin{pmatrix}
\beta
\end{pmatrix}_x = 0_{\alpha+1}$, hence $G_{\alpha+1}$ generates a linear code with $r=2$.

Before constructing more optimal ternary codes, we present some notations:

\[
f_2(i, j, k) = \begin{pmatrix}
iS_2
jS_2
kS_2
\end{pmatrix}, f_4(i, j) = \begin{pmatrix}
iS_3
jS_3
\end{pmatrix} \text{ for } i, j, k \in \{0, 1, 2\}.
\]

For two matrices $A$ and $B$, $A \setminus B$ means deleting the columns in $B$ from $A$.

3. Construction and Locality of Some Optimal $[n, 6]_3$ Codes

3.1. Construction And Locality of Two Optimal Codes from Codes with Short Length

In the following, we will construct optimal $[406, 6, 290]_3$ and $[447, 6, 297]_3$ codes, respectively.

Let

\[
B_1 = \begin{pmatrix}
11000001020110 \\
001000110202 \\
0001000102122 \\
00001001100111 \\
00000102211200 \\
00000010022111
\end{pmatrix}, B_2 = \begin{pmatrix}
0111111011111011110111 \\
101222101122102101 \\
2220121111201210110 \\
1101112112120022011 \\
1212201211021201101 \\
2221022120121120110
\end{pmatrix}.
\]

Matrix $B_1$ generates an optimal $C_{406} = [14, 6, 6]_3$ code with weight polynomial $1 + 102z^6 + 440z^9 + 186z^{12}$. Choose the monic words in $C_{406}$ with weight 9 and all words with weight 12, denote these row vectors by $B_{406,14}$, the 2,..., 7 columns in $B_{406,14}$ form a $406 \times 6$ matrix, the transpose of this matrix generates an optimal $[406, 6, 270]_3$ code with weight polynomial $270 + 297 + 324$. Matrix $B_2$ generates an optimal $C_{447} = [19, 6, 9]_3$ code with weight polynomial $W_t = 1 + 86z^9 + 402z^{12} + 228z^{15} + 12z^{18}$. Choose the monic words in $C_{447}$ with weight 12, 15 and monic words with weight 18 three times, denote the first 6 entries of these words by $B_{447,6}$. The transpose of $B_{447,6}$ generates an optimal $[447, 6, 297]_3$ code with weight polynomial $W_t = 1 + 690z^{297} + 38z^{324}$.

Observing the generator matrix of optimal $[447, 6, 297]_3$ code, we know it has the following form

\[
G_{6,447} = \begin{pmatrix}
S_x \\
S_y \\
N
\end{pmatrix}, \quad (3.1)
\]

where the entries of row vectors $x$ and $y$ are all non-zero. According to Lemma 2.1, we know optimal $[447, 6, 297]_3$ code has locality $r=2$. Let
From Lemma 2.1, we know matrix \( S \setminus L \) generates a ternary linear code with \( r = 2 \). We can rewrite the generator matrix of optimal \([406,6,270]\) code as the following form

\[
G_{6,406} = \begin{pmatrix}
S_x \\
S_y
\end{pmatrix} \left( \begin{array}{c|c}
L & N
\end{array} \right),
\]

(3.2)

where \( x \) and \( y \) are vectors and the entries are all non-zero. Hence, according to Lemma 2.1, this optimal \([406,6,270]\) code has locality \( r = 2 \).

3.2. Optimal Codes from \([477,6,297]\) and \([406,6,270]\) Codes

We will construct ternary optimal codes from anticodes. A ternary linear anticode \( \mathcal{A} \) of length \( n \) and maximum distance \( \delta \) is a set of codewords in \( F_3^n \) such that the Hamming distance between any pair of codewords is less than or equal to \( \delta \). The generator matrix \( G_\mathcal{A} \) of \( \mathcal{A} \) is a \( k \times n \) ternary matrix such that all the \( 3^k \) combinations of its rows form the codewords of the anticode. First we need to construct the generator matrices of anticodes which can be used to construct optimal ternary codes.

Construct two anticodes \( \mathcal{A}_1 \) and \( \mathcal{A}_2 \), which have generator matrices

\[
G_{\mathcal{A}_1} = \begin{pmatrix}
0_{2,4} \\
S_2 \\
1012 \\
1012
\end{pmatrix}, \quad G_{\mathcal{A}_2} = \begin{pmatrix}
S_2 \\
1012 \\
1012 \\
1012
\end{pmatrix}.
\]

We can derive many optimal ternary codes from anticodes \( \mathcal{A}_1, \mathcal{A}_2, f_2(i,j,k) \) and \( f_3(i,j) \). The matrix obtained by deleting columns in \( \mathcal{A}_1 \) from \( G_{6,447} \) is a generator matrix of optimal \([443,6,294]\) code. Because the generator matrix of optimal \([447,6,297]\) code has form (3.1), the generator matrix of optimal \([443,6,294]\) code satisfies the condition in Lemma 2.1, this optimal \([443,6,294]\) code has locality \( r = 2 \).

Deleting the columns in \([G_{\mathcal{A}_1} | G_{\mathcal{A}_2}]\) from \( G_{6,447} \), an optimal \([439,6,291]\) code with \( r = 2 \) is obtained.

In the following, we will construct 8 optimal ternary codes, and then 3 optimal codes can be derived from any of these 8 optimal codes. All of these optimal ternary have locality \( r = 2 \).

Let

\[
G_\mathcal{A} = \begin{pmatrix}
101111101010 \\
2022111010101 \\
0112012001122 \\
000011222222 \\
1011111010101 \\
2022111010101
\end{pmatrix}.
\]
Matrices $G_{6,447} \setminus G_{A}, G_{6,447} \setminus (G_{A} | f_{2}(1,1,2))$ and $G_{6,447} \setminus (G_{A} | f_{2}(1,1,2) | f_{2}(1,2,2))$ generate three optimal [434,6,288], [430,6,285], and [426,6,282] codes.

Construct

\[
G_{6,406} \setminus f_{2}(1,1,1), G_{6,402} \setminus f_{2}(1,2,1), G_{6,408} \setminus f_{3}(1,1),
\]

\[
G_{6,406} \setminus f_{2}(1,1,1) | f_{2}(1,1,1), G_{6,406} \setminus f_{3}(1,1) | f_{2}(1,1,2),
\]

these five matrices generate five optimal [402,6,267], [398,6,264], [393,6,261], [389,6,258], and [385,6,255] codes.

Collecting the results above, we have constructed 12 optimal ternary codes with locality $r=2$. Next, we will discuss how to derive more optimal codes from these 12 optimal codes. The optimal [443,6,294] code is constructed by deleting four column vectors in $A$ from generator matrix of [447,6,297] code. According to (3.1) and [17], it's not difficult to know that deleting columns in $G_{A}$ from $G_{6,447}$ one by one, we can construct optimal [446,6,296], [445,6,295], [444,6,294], and [443,6,294] codes, the first three optimal codes still have locality $r=2$.

In Section 3.1 and Section 3.2, we have constructed 12 optimal codes with parameters:

- [447,6,297], [443,6,294], [439,6,291], [434,6,288], [430,6,285], [426,6,282], [406,6,270], [402,6,267], [398,6,264], [393,6,261], [389,6,258], and [385,6,255].

Similar to discussion of optimal codes derived from optimal [447,6,297] code, from other 11 optimal codes, we can construct 33 optimal codes with parameters:

- [442,6,293], [441,6,292], [440,6,291], [438,6,290], [437,6,289], [436,6,288], [433,6,287], [432,6,286], [431,6,285], [429,6,284], [428,6,283], [427,6,282], [425,6,281], [424,6,280], [423,6,279], [405,6,269], [404,6,268], [402,6,267], [401,6,266], [400,6,265], [399,6,264], [397,6,263], [396,6,262], [395,6,261], [392,6,260], [391,6,259], [390,6,258], [388,6,257], [387,6,256], [386,6,255], [385,6,255], [386,6,254], [387,6,254], [388,6,254], [389,6,254].

Optimal [435,6,288], and [394,6,261], codes can be derived from optimal [434,6,288], and [393,6,261] code, respectively. The generator matrices of these optimal codes satisfy the condition in Lemma 2.1, hence all these optimal codes have locality $r=2$.

4. Conclusion

In this paper, we have constructed 50 optimal ternary linear codes with dimension 6 and analyzed their locality. Using two ternary codes with short lengths and different anticodes, we constructed 50 optimal ternary codes. By analyzing the constitution of generator matrices of these optimal codes, we determined their locality $r=2$, which have potential application prospect. In our ongoing work, we will focus on constructing more optimal ternary codes with length $n > 324$.

5. Acknowledgments

This work is supported by National Natural Science Foundation of China under Grant No.11471011 and Natural Science Foundation of Shaanxi under Grant No.2017JQ1032.

6. References

[1] Gopalan P, Huang C, Simitci H and Yekhanin S 2012 IEEE Trans. Inf. Theory 58 6925-6934.
[2] Cadambe V and Mazumdar A 2013 Proc. IEEE Symp. Netw. Coding 1-5.
[3] Baumert L D and McEliece R J 1973 IEEE Trans. Inf. Theory 9 134-135.
[4] Huang P, Yaakobi E, Uchikawa H and Siegel P H 2016 IEEE Trans. Inf. Theory 62 6268-6283.
[5] Silverstein N and Zeh A 2015 IEEE International Symposium on Inf. Theory 1247-1251.
[6] Gopraj S and Calderbank R 2014 IEEE International Symposium on Inf. Theory 676-680.
[7] Fu Q, Li R, Guo L and Lv L 2017 Finite Fields and Their Applications 48 371-394.
[8] Zeh A and Yaakobi E 2015 IEEE International Symposium on Inf. Theory 1247-1251.
[9] Yang R, Li R, Guo L, Fu Q and Rao Y 2017 Procedia Computer Science 164-169.
[10] Yang R, Li R, Guo L and Fu Q 2017 Journal of Air Force Engineering University (Natural Science Edition) 18 105-111.
[11] Yang R, Li R, Guo L and Fu Q 2017 IEICE Trans. Fundamentals. E100-A 2172-2175.
[12] Rawat A, Papailiopoulos D, Dimakis A and Vishwanath S 2014 2014 IEEE International Symposium on Inf. Theory 681-685.
[13] Tamo I, Papailiopoulos D and Dimakis A 2013 IEEE International Symposium on Inf. Theory 1814-1818.
[14] Tamo I and Barg A 2014 IEEE Trans. Inf. Theory 60 4661-4676.
[15] Prakash N, Kamath G, Lalitha V and Kumar P 2012 IEEE International Symposium on Inf. Theory 2776-2780.
[16] Song W, Dau S, Yuen C and Li T 2014 IEEE Journal on Selected Areas in Communications 32 1019-1036.
[17] Grassl M Bounds On the minimum distance of linear codes. http:\\www.codetables.de.
[18] Song Q, Li R, Fu Q and Yang R 2018 Journal of Air Force Engineering University (Natural Science Edition) 19 100-105.