TWMS J. App. Eng. Math. V.5, N.1, 2015, pp. 110-117.

CODES CORRECTING KEY ERRORS

PANKAJ KUMAR DAS[1], §

Abstract. The objective of coding theory is to protect a message going through a noisy channel. The nature of errors that cause noisy channel depends on different factors. Accordingly codes are needed to develop to deal with different types of errors. Sharma and Gaur [6] introduced a new kind of error which is termed as 'key error'. This paper presents lower and upper bounds on the number of parity-check digits required for linear codes capable of correcting such errors. An example of such a code is also provided.

Keywords: Parity check matrix, syndrome, standard array, solid burst error.

AMS Subject Classification (2010): 94B05, 94B65.

1. Introduction

With the advancement of information technology, different types of new problems have been coming out. Different types of error patterns are also coming out. Error control coding now is not limited to distant communication only. There are communication channels like the automata or electronics devices which are encountered with specific type of errors for which the coding is required. The error patterns which they produce are all dependent on the characteristics of the device. It is important to carefully study the error patterns that actually occur. This allows correction of all errors that actually occur rather than partial correction or correcting non-errors that results in wasting the capacity of the channel.

Let us consider the keyboard of a computer; it has keys for various numbers and other symbols. Imagine punching a number or an alphabet key on it. While word processing, one may erroneously press a key on one or two positions on either side of the right key, rather than any key on the keyboard. These likely positions will constitute the set of errors for the number or the symbol key pressed. Such errors are already discussed by Sharma and Gaur [6] with respect to S-K metric. We give here a study on such errors with respect to Hamming metric. We call such errors as key errors and they are defined as follows:

Definition 1.1. *A i-key error of length b is a vector such that the i^{th} component is non-zero and the other non-zero components are confined to some b consecutive components in either side of the i^{th} component.*

[1] Department of Mathematics, Shivaji College(University of Delhi), Raja Garden, Delhi-110 027, India. e-mail: pankaj4thapril@yahoo.co.in;

§ Manuscript received: July 05, 2014.
TWMS Journal of Applied and Engineering Mathematics, Vol.5, No.1;

It may be noted that in such error, the entry error i.e., the i^{th} component may be the first position and go to upto the n^{th} position. If the entry error is the first position, then the non-zero components are confined to b consecutive components on the right side of the first position. If the entry error is the second position, then non-zero components are confined to one position of left side of the second position and to the b consecutive components on the right side of the second position. Continue the process such that nonzero components are confined to b consecutive positions of either side of the entry error. Finally if the entry error is the n^{th} position, then non-zero components are confined to the b consecutive components on the left side of the n^{th} component.

For example in a vector of length 6 over a field of 3 elements $GF(3)$, the key errors of length 2 are $(0\underbrace{12}_{2}\ \underbrace{2}_{\text{entry error}}\ \underbrace{12}_{2})$, $(0\underbrace{12}_{2}\ \underbrace{2}_{\text{entry error}}\ \underbrace{02}_{2})$, $(0\underbrace{12}_{2}\ \underbrace{2}_{\text{entry error}}\ \underbrace{00}_{2})$, $(000\underbrace{12}_{2}\ \underbrace{2}_{\text{entry error}})$, $(\underbrace{12}_{2}\ \underbrace{1}_{\text{entry error}}\ \underbrace{12}_{2}0)$, $(\underbrace{2}_{1}\ \underbrace{2}_{\text{entry error}}\ \underbrace{12}_{2}00)$, etc.

The basic purpose of error-correcting codes is to correct errors that are occurred during communication. This is done by adding parity check digits (redundancy) to the information digits. The efficiency of a error correcting code depends on the number of parity check digits. The lesser is the number of parity check symbols in a code, the more is the rate of information of the code. It is not generally possible to give the exact number of redundant/parity check digits for a given error correcting code. But, we can obtain bounds on the number of redundant/parity check digits. This was initiated by Hamming [2] who was concerned with both code constructions and bounds. After that, many other researchers have worked on bounds and this paper is also in that direction.

Das [1] has presented lower and upper bounds on parity check for the codes detecting key errors. Reiger type of bound [4] on codes detecting and simultaneously correcting such errors was also studied. The present paper obtains lower and upper bounds on parity check digits for a linear code capable of correcting such errors.

The paper is organized as follows: Section 1 i.e., the Introduction gives brief view of the importance of the study of the paper and basic definition. In Section 2, we obtain lower and upper bounds on the number of parity check digits of a linear code that corrects any key error of length b or less. This is followed by an example of such a code. Section 3 is the conclusion.

In what follows a linear code will be considered as a subspace of the space of all n-tuples over $GF(q)$. The distance between two vectors shall be considered in the Hamming sense.

2. Correction of key errors

We consider the linear codes that are capable of correcting any key errors of length b or less. Firstly, a lower bound on the number of parity-check digits required for such a code is obtained. The proof is based on the technique used in Theorem 4.16, Peterson and Weldon [3].

Theorem 2.1. *Any (n,k) linear code over $GF(q)$ that corrects any key error of length b or less must satisfy*

$$q^{n-k} \geq 1 + \frac{q^{2b+1}-q}{q+1} + (n-2b)(q-1)\left(\frac{q^{2b+1}-q}{q+1}+1\right) + \frac{q^{2b+1}-q^3}{(q+1)^2} + \frac{(q-1)(b+q)}{q+1}.$$

Proof. The result will be proved by counting the number of correctable errors and setting it less than or equal to q^{n-k}.

The number of key errors of length b or less when the entry error of the key errors is from 1^{st} position to b^{th} position is given by

$$\begin{aligned}
&(q-1)q^b\\
&+(q-1)^2q^b+(q-1)q^{b-1}\\
&+(q-1)^2q^{b+1}+(q-1)^2q^{b-1}+(q-1)q^{b-2}\\
&+(q-1)^2q^{b+2}+(q-1)^2q^b+(q-1)^2q^{b-2}+(q-1)q^{b-3}\\
&+.....\\
&+.....\\
&+(q-1)^2q^{2b-2}+(q-1)^2q^{2b-4}+(q-1)^2q^{2b-6}+\cdots+(q-1)^2q^2+(q-1)q,
\end{aligned}$$

which is equal to

$$(q-1)^2\Bigg\{q^b\bigg(\frac{q^2-1}{q^2-1}\bigg)+q^{b-1}\bigg(\frac{(q^2)^2-1}{q^2-1}\bigg)+q^{b-2}\bigg(\frac{(q^2)^3-1}{q^2-1}\bigg)+\cdots+q^3\bigg(\frac{(q^2)^{b-2}-1}{q^2-1}\bigg)+q^2\bigg(\frac{(q^2)^{b-1}-1}{q^2-1}\bigg)\Bigg\}+(q-1)q\bigg(\frac{q^b-1}{q-1}\bigg),$$

which on simplification gives

$$\frac{q^{2b+1}-q}{q+1}. \tag{1}$$

The number of key errors of length b or less when the entry error of the key errors is from $(b+1)^{th}$ position to $(n-b)^{th}$ position is given by

$$(n-2b)\Big\{(q-1)^2q^{2b-1}+(q-1)^2q^{2b-3}+(q-1)^2q^{2b-5}+\cdots+(q-1)^2q+(q-1)\Big\}$$

$$=(n-2b)(q-1)\bigg(\frac{q^{2b+1}-q}{q+1}+1\bigg). \tag{2}$$

If the entry error of the key errors is in the last b positions, the number of key errors of length b or less is

$$\begin{aligned}
&(q-1)^2q^{2b-3}+(q-1)^2q^{2b-5}+(q-1)^2q^{2b-7}+\cdots+(q-1)^2q^3+(q-1)^2q+(q-1)\\
&+(q-1)^2q^{2b-5}+(q-1)^2q^{2b-7}+(q-1)^2q^{2b-9}+\cdots+(q-1)^2q^3+(q-1)^2q+(q-1)\\
&+.....\\
&+.....\\
&+(q-1)^2q^5+(q-1)^2q^3+(q-1)^2q+(q-1)\\
&+(q-1)^2q^3+(q-1)^2q+(q-1)\\
&+(q-1)^2q+(q-1)\\
&+(q-1),
\end{aligned}$$

which is equivalent to

$$(q-1)^2\Bigg\{q\bigg(\frac{(q^2)^{b-1}-1}{q^2-1}\bigg)+q\bigg(\frac{(q^2)^{b-2}-1}{q^2-1}\bigg)+q\bigg(\frac{(q^2)^{b-3}-1}{q^2-1}\bigg)+\cdots+q\bigg(\frac{(q^2)^2-1}{q^2-1}\bigg)+q\bigg(\frac{q^2-1}{q^2-1}\bigg)\Bigg\}+(q-1)b,$$

which again on simplification gives

$$\frac{q^{2b+1}-q^3}{(q+1)^2}-\left(\frac{q-1}{q+1}\right)q(b-1)+(q-1)b$$

$$=\frac{q^{2b+1}-q^3}{(q+1)^2}+\frac{(q-1)(b+q)}{q+1}. \tag{3}$$

Therefore, the total number of key errors of length b or less is given by

$$expr.(1)+expr.(2)+expr.(3)$$
$$=\frac{q^{2b+1}-q}{q+1}+(n-2b)(q-1)\left(\frac{q^{2b+1}-q}{q+1}+1\right)+\frac{q^{2b+1}-q^3}{(q+1)^2}+\frac{(q-1)(b+q)}{q+1}.$$

For correction, all these vectors must belong to different cosets. The total number of cosets available is q^{n-k} . Therefore, we must have

$$q^{n-k}\geq 1+\frac{q^{2b+1}-q}{q+1}+(n-2b)(q-1)\left(\frac{q^{2b+1}-q}{q+1}+1\right)+\frac{q^{2b+1}-q^3}{(q+1)^2}+\frac{(q-1)(b+q)}{q+1}.$$

□

Now the following theorem gives an *upper* bound on the number of check digits required for the construction of a linear code considered in Theorem 2.1. This bound assures the existence of a linear code that can correct all key errors of length b or less. The proof is based on the well known technique used in Varshomov-Gilbert Sacks bound by constructing a parity check matrix for such a code (refer Sacks [5], also Theorem 4.7 Peterson and Weldon [3]). The procedure involves suitable modifications of the technique used in deriving Varshamov-Gilbert-Sacks bound.

Theorem 2.2. *There shall always exist an (n,k) linear code over $GF(q)$ that corrects any key error of length b or less $(n>4b+2)$ provided that*

$$q^{n-k}>\left(\frac{q^{2b+1}+1}{q+1}\right)\left\{1+\frac{q^{2b+1}-q}{q+1}+(n-1-4b)(q-1)\left(\frac{q^{2b+1}-q}{q+1}+1\right)\right.$$
$$\left.+\frac{q^{2b+1}-q^3}{(q+1)^2}+\frac{(q-1)(b+q)}{q+1}\right\}.$$

Proof. The existence of such a code will be proved by constructing an $(n-k)\times n$ parity check matrix H for the desired code as follows:
Select any non zero $(n-k)$-tuple as the first column h_1 of the matrix H. After having selected the first $j-1$ columns $h_1,h_2,\ldots,h_{j-1}$ appropriately, we lay down the condition to add j^{th} column such that

$$\begin{aligned} h_j \neq\ & (u_{j-1}h_{j-1}+u_{j-2}h_{j-2}+\cdots+u_{j-2b}h_{j-2b}) \\ & +(v_ih_i+v_{i+1}h_{i+1}+\cdots+v_{i+2b}h_{i+2b}), \end{aligned} \tag{4}$$

where $u_i,v_i\in GF(q)$; $i+2b<j-2b$; if the coefficient u_i (or v_i) is non zero, then the other non zero coefficients of u_i (or v_i) are confined to b consecutive positions of either side of u_i (or v_i), also if $u_i=0$ for $(j-1\leq i\leq j-b)$, then $u_i=0\ \forall\ i=j-1$ to $j-2b$.

This condition ensures that there shall not be a code vector which can be expressed as sum (difference) of key errors of length b or less each. Thus, the codes so constructed will be able to correct such errors.

We now enumerate all possible linear combinations on the R.H.S. of (4):

The coefficient u_i on R.H.S. of (4) are such that if $u_i = 0$ for $j-1 \le i \le j-b$, then $u_i = 0$ for $j-b-1 \le i \le j-2b$ and if u_i is non zero coefficient, then the other non zero coefficients of u_i are confined to b consecutive positions of either side of u_i. The number of such coefficient u_i, including the zero vector, is given by

$$\begin{aligned}
&(q-1)q^b \\
&+(q-1)^2q^b + (q-1)q^{b-1} \\
&+(q-1)^2q^{b+1} + (q-1)^2q^{b-1} + (q-1)q^{b-2} \\
&+(q-1)^2q^{b+2} + (q-1)^2q^b + (q-1)^2q^{b-2} + (q-1)q^{b-3} \\
&+..... \\
&+..... \\
&+(q-1)^2q^{2b-2} + (q-1)^2q^{2b-4} + (q-1)^2q^{2b-6} + \cdots + (q-1)^2q^2 + (q-1)q \\
&+1,
\end{aligned}$$

which on simplification gives

$$\frac{q^{2b+1}-q}{q+1} + 1 = \frac{q^{2b+1}+1}{q+1}. \tag{5}$$

To enumerate the coefficients v_i is equivalent to the number of key errors of length b or less in a $(j-1-2b)$-tuple. The number of key errors of length b or less in a $(j-1-2b)$-tuple, including the zero vector, is given by (refer Theorem 2.1)

$$1 + \frac{q^{2b+1}-q}{q+1} + (j-1-4b)(q-1)\left(\frac{q^{2b+1}-q}{q+1}+1\right) + \frac{q^{2b+1}-q^3}{(q+1)^2} + \frac{(q-1)(b+q)}{q+1}. \tag{6}$$

Thus, the total number of linear combinations on the R.H.S. of (4) equals

$$\begin{aligned}
&expr.(5) \times expr.(6) \\
=&\left(\frac{q^{2b+1}+1}{q+1}\right)\left\{1 + \frac{q^{2b+1}-q}{q+1} + (j-1-4b)(q-1)\left(\frac{q^{2b+1}-q}{q+1}+1\right)\right. \\
&\left. + \frac{q^{2b+1}-q^3}{(q+1)^2} + \frac{(q-1)(b+q)}{q+1}\right\}.
\end{aligned} \tag{7}$$

Since these many linear combinations can not be equal to h_j and at worst all the linear combinations computed in (7) might yield a distinct sum, therefore in view of the fact that the total possible number of $(n-k)$-tuples is q^{n-k} , the j^{th} column h_j can be added to H provided that

$$q^{n-k} > expr.(7). \tag{8}$$

To obtain a code of length n, we replace j by n in the above inequality and the inequality (8) becomes

$$\begin{aligned}
q^{n-k} >&\left(\frac{q^{2b+1}+1}{q+1}\right)\left\{1 + \frac{q^{2b+1}-q}{q+1} + (n-1-4b)(q-1)\left(\frac{q^{2b+1}-q}{q+1}+1\right)\right. \\
&\left. + \frac{q^{2b+1}-q^3}{(q+1)^2} + \frac{(q-1)(b+q)}{q+1}\right\}.
\end{aligned} \tag{9}$$

□

Alternate Form 2.1 If N is the largest value of n satisfying the inequality (9), then by replacing n by $N+1$, the inequality in (9) gets reversed and we get

$$q^{n-k} \leq \left(\frac{q^{2b+1}+1}{q+1}\right)\left\{1+\frac{q^{2b+1}-q}{q+1}+(N-4b)(q-1)\left(\frac{q^{2b+1}-q}{q+1}+1\right)\right.$$
$$\left.+\frac{q^{2b+1}-q^3}{(q+1)^2}+\frac{(q-1)(b+q)}{q+1}\right\}. \tag{10}$$

It should be noted that we don't need to replace n in terms of N on the L.H.S. of above inequality since L.H.S. represents the number of cosets/vectors of length $n-k$ which remains the same.

Alternate Form 2.2 If B is the largest value of b satisfying the inequality (9), then for $b = B+1$, the inequality in (9) gets reversed and we get

$$q^{n-k} \leq \left(\frac{q^{2B+3}+1}{q+1}\right)\left\{\frac{q^{2B+3}-q}{q+1}+(n-5-4B)(q-1)\left(\frac{q^{2B+3}-q}{q+1}+1\right)\right.$$
$$\left.+\frac{q^{2B+3}-q^3}{(q+1)^2}+\frac{(q-1)B}{q+1}+q\right\}. \tag{11}$$

Example 2.1. For a $(11, 2)$ linear code over $GF(2)$, we construct the following 9×11 parity check matrix H, according to the synthesis procedure given in the proof of Theorem 2.2 by taking $q = 2$,$b = 2$ and $n = 11$.

$$H = \begin{bmatrix} 1&0&0&0&0&0&0&0&1&0&0 \\ 1&1&0&0&0&0&0&0&1&1&0 \\ 0&1&1&0&0&0&0&0&1&1&1 \\ 0&0&1&1&0&0&0&0&1&1&1 \\ 0&0&0&1&1&0&0&0&1&1&0 \\ 0&0&0&0&1&1&0&0&1&1&1 \\ 0&0&0&0&0&1&1&0&1&1&1 \\ 0&0&0&0&0&0&1&1&0&1&1 \\ 0&0&0&0&0&0&0&1&0&0&1 \end{bmatrix}$$

It can be seen from Table 2.1 that the syndromes of any key error of length 2 or less are all nonzero and distinct. This shows that the code that is the null space of this matrix can correct any key error of length 2 or less.

Table 2.1
Error pattern - syndromes table

Error-patterns	Syndromes	Error-patterns	Syndromes
10000000000	110000000	00000111000	000001001
11000000000	101000000	00010110000	000111010
10100000000	111100000	00010101000	000111111
11100000000	100100000	00010111000	000111001
01000000000	011000000	00001101000	000010111
01100000000	010100000	00001111000	000010001
01010000000	011110000	00011101000	000100111
01110000000	010010000	00011111000	000100001

Cont...

Error-patterns	Syndromes	Error-patterns	Syndromes
11010000000	101110000	00000010000	000000110
11110000000	100010000	00000011000	000000101
00100000000	001100000	00000010100	111111010
00110000000	001010000	00000011100	111111001
00101000000	001111000	00001011000	000011101
00111000000	001001000	00001010100	111100010
10110000000	111010000	00001011100	111100001
10101000000	111111000	00000110100	111110110
10111000000	111001000	00000111100	111110101
01101000000	010111000	00001110100	111101110
01111000000	010001000	00001111100	111101101
11101000000	100111000	00000001000	000000011
11111000000	100001000	00000001100	111111111
00010000000	000110000	00000001010	011111101
00011000000	000101000	00000001110	100000001
00010100000	000111100	00000101100	111110011
00011100000	000100100	00000101010	011110001
01011000000	011101000	00000101110	100001101
01010100000	011111100	00000011010	011111011
01011100000	011100100	00000011110	100000111
00110100000	001011100	00000111010	011110111
00111100000	001000100	00000111110	100001011
01110100000	010011100	00000000100	111111100
01111100000	010000100	00000000110	100000010
00001000000	000011000	00000000101	110010011
00001100000	000010100	00000000111	101101101
00001010000	000011110	00000010110	100000100
00001110000	000010010	00000010101	110010101
00101100000	001110100	00000010111	101101011
00101010000	001111110	00000001101	110010000
00101110000	001110010	00000001111	101101110
00011010000	000101110	00000011101	110010110
00011110000	000100010	00000011111	101101000
00111010000	001001110	00000000010	011111110
00111110000	001000010	00000000011	010010001
00000100000	000001100	00000001011	010010010
00000110000	000001010	00000000001	001101111
00000101000	000001111		

3. Conclusion

The paper presents lower and upper bound for a code correcting key errors of length b or less. The equality of the inequality in the statement of Theorem 2.1 (Lower bound) gives us the optimal case and the corresponding codes. These codes are termed as optimal/perfect codes as they correct key errors of length b or less and no others. These codes are useful in communication due to their high rate of information. To study such optimal codes remains a further study.

References

[1] Das, P. K., (2014), Codes on Key Errors, Cybernetics and Information Technologies, 14(2), pp. 31-37.
[2] Hamming, R. W., (1950), Error-detecting and error-correcting codes, Bell System Technical Journal, 29, pp. 147-160.
[3] Peterson, W. W. and Weldon(Jr.), E. J., (1972), Error-Correcting Codes, 2nd edition, The MIT Press, Mass.
[4] Reiger, S. H., (1960), Codes for the Correction of Clustered Errors, IRE Trans. Inform. Theory, IT-6, pp. 16-21.
[5] Sacks, G. E., (1958), Multiple error correction by means of parity-checks, IRE Trans. Inform. Theory, IT-4, pp. 145-147.
[6] Sharma, B. D. and Gaur, A., (2013), Codes Correcting Limited Patterns of Random Errors Using S-K Metric, Cybernetics and Information Technologies, 13(1), pp. 34-45.

Pankaj Kumar Das for the photography and short autobiography, see TWMS J. App. Eng. Math., V.3, N.2.