BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **14**, No 2

Sofia • 2014

Print ISSN: 1311-9702; Online ISSN: 1314-4081
DOI: 10.2478/cait-2014-0017

Codes on Key Errors

P. K. Das

Department of Mathematics, Shivaji College (University of Delhi),
Raja Garden, Delhi-110027, India
Email: pankaj4thapril@yahoo.co.in

Abstract: *Coding theory has started with the intention of detection and correction of errors which have occurred during communication. Different types of errors are produced by different types of communication channels and accordingly codes are developed to deal with them. In 2013 Sharma and Gaur introduced a new kind of an error which will be termed "key error". This paper obtains the lower and upper bounds on the number of parity-check digits required for linear codes capable for detecting such errors. Illustration of such a code is provided. Codes capable of simultaneous detection and correction of such errors have also been considered.*

Keywords: *Parity check digit, parity check matrix, syndrome, standard array.*

1. Introduction

With the advancement of information technology, different types of new problems have appeared. Since error control coding is now not limited to distant communication only, mathematics is needed which can suitably match the characteristics of the device for which the coding is required. The communication channels, like the automata or the electronic devices, have varying characteristics. The errors patterns which they produce have different characteristics. We need to consider only those patterns that need to be detected/corrected rather than the wasted capacity of detecting/correcting non-errors by default.

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 strike a key on one or two positions on either side of the correct key, rather than any key on the keyboard. These positions will

constitute the set of errors for the number or the symbol key pressed. S h a r m a and G a u r [11] have discussed such errors and done a study on such errors with respect to S-K metric. We call such errors *key errors* and they are defined as follows:

Definition 1. An *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 immediate *b* consecutive components in either side of the *i*-th component.

It may be noted that in a vector of length *n*, for such error, the entry error, i.e., *i*-th component may be the first position and go up to the *n*-th position. If the error entry is the first position, then the non-zero components are confined to immediate *b* consecutive components on the right side of the first position. If the error entry is the second position, then non-zero components are confined to one position on the left side of the second position and to the immediate *b* consecutive components on the right side of the second position. In the same way, if the error entry is the *n*-th position, then the non-zero components are confined to the immediate *b* consecutive components on the left side of *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

$$\underset{2}{0}\underset{2}{1}\underset{2}{2}\underset{2}{1}\underset{2}{2}, \underset{2}{0}\underset{2}{1}\underset{2}{2}\underset{2}{0}\underset{2}{2}, \underset{2}{0}\underset{2}{1}\underset{2}{2}\underset{2}{0}\underset{2}{0}, \underset{2}{0}\underset{2}{0}\underset{2}{0}\underset{2}{1}\underset{2}{2}, \underset{2}{0}\underset{2}{0}\underset{2}{0}\underset{2}{1}\underset{2}{2}, \underset{2}{1}\underset{2}{2}\underset{2}{1}\underset{2}{1}\underset{2}{2}, \underset{2}{0}\underset{2}{2}\underset{2}{1}\underset{2}{2}, \underset{1}{2}\underset{2}{1}\underset{2}{2}, \text{etc.}$$

It is important to know the ultimate capabilities and limitations of error correcting codes. This information, along with the knowledge of what is practically achievable, indicates which problems are virtually solved and which need further work. This was initiated by H a m m i n g [7] who was concerned with both code constructions and bounds. The bounds on the number of parity check symbols determine the efficiency of a code. The less the number of parity check symbols in a code is, the greater the rate of the code information is.

The paper presents a study on the bounds of linear codes detecting key errors. Detecting error is one of the very important studies to researchers. For more details on such studies, one may refer to [1-7].

The paper is organized as follows:

Section 1 gives a brief review of the importance of the study of the paper, basic background and definition. In Section 2 we obtain the lower and upper bounds on the number of parity check digits of linear codes that detect any key error of length *b* or less. This is followed by an illustration of such a code in Section 3. Section 4 presents a bound on parity check digits for codes capable of detecting and simultaneously correcting such errors. Section 5 gives the conclusion.

Further on a linear code will be considered as a subspace of the space of all *n*-tuples over GF(*q*). The distance between two vectors will be considered in the sense of Hamming.

2. Codes detecting key errors

We consider linear codes over GF(*q*) that are capable of detecting any key error of length *b* or less. The error patterns to be detected must not be a codeword. In other words, we consider codes that have no key error of length *b* or less as a code word.

Firstly, we obtain a lower bound over the number of parity-check digits required for such a code. The proof is based on the technique used in Theorem 4.13, P e t e r s o n and W e l d o n [8].

Theorem 1. Any (n, k) linear code over GF(q) that detects any key error of length b or less must satisfy

$$q^{n-k} \geq \begin{cases} 1+b & \text{for} \quad q=2, \\ 1+b(q-1)(q-2) & \text{for} \quad q \neq 2. \end{cases}$$

P r o o f: The result will be proved on the basis that no detectable error vector can be a code word.

Let V be an (n, k) linear code over GF(q). Let X be the set of all vectors, such that the non zero components are confined to the first $2b$ positions in the following ways:

(i) for q=2, the first $2b$ positions are

$$\overbrace{x000\ldots00}^{b}\overbrace{y000\ldots00}^{b}$$
$$\overbrace{0x00\ldots000}^{b}\overbrace{y00\ldots00}^{b}$$
$$\overbrace{00x0\ldots0000}^{b}\overbrace{y0\ldots00}^{b}$$
$$\vdots$$
$$\overbrace{0000\ldots0x000}^{b}\overbrace{0\ldots0y}^{b}$$

where x=y=1;

(ii) for $q\neq2$, the first $2b$ positions are

$$\overbrace{x000\ldots00}^{b}\overbrace{y000\ldots00}^{b}$$
$$\overbrace{0x00\ldots000}^{b}\overbrace{y00\ldots00}^{b}$$
$$\overbrace{00x0\ldots0000}^{b}\overbrace{y0\ldots00}^{b}$$
$$\vdots$$
$$\overbrace{0000\ldots0x000}^{b}\overbrace{0\ldots0y}^{b}$$

where x, y belong to GF(q)–{0} and $x\neq y$.

We claim that two vectors of the set X must belong to a different coset of the standard array.

Assume on the contrary, that there are two vectors, say x_1, x_2 in X belonging to the same coset of the standard array. Then their difference viz. $x_1 - x_2$ must be a code vector. But $x_1 - x_2$ is a vector whose all non zero components are confined to $2b$ or less consecutive components, in which the gap of components between any

two consecutive non zero components is less than b, i.e., $x_1 - x_2$ is a key error of length b or less, which is a contradiction. Thus all the vectors in X must belong to distinct cosets of the standard array. The number of such vectors over GF(q), including the vector of all zero, is clearly:

(i) $1+b$ for $q = 2$,
(ii) $1+b(q-1)(q-2)$ for $q \neq 2$.

Since the number of available cosets is q^{n-k}, so

$$q^{n-k} \geq \begin{cases} 1+b & \text{for} \quad q=2, \\ 1+b(q-1)(q-2) & \text{for} \quad q \neq 2. \end{cases}$$

Hence the theorem is proved ■

In the next theorem, an upper bound on the number of check digits required for the construction of a linear code mentioned in Theorem 1 is provided. This bound assures the existence of a linear code that can detect 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 to S a c k s [10], also Theorem 4.7 of P e t e r s o n and W e l d o n [8]).

Theorem 2. There exists an (n, k) linear code over GF(q) that has no key error of length b or less as a code word provided that

$$N-k > \log_q[1+ q^{2b-1}(q-1)].$$

P r o o f: The existence of such a code will be shown by constructing an appropriate $(n-k)\times n$ parity-check matrix H. The requisite parity-check matrix H will be constructed as follows:

Select any non-zero ($n-k$)-tuples as the first $n-1$ columns h_1, h_2, ..., h_{n-1} appropriately, we lay down the condition to add n-th column h_n such that h_n must not be a linear combination of immediately preceding consecutive b columns, together with the sum of the preceding (b+1)-th column, and along with a linear combination of immediately preceding consecutive b columns after the (b+1)-th column.

In other words,

$$h_n \neq (u_{n-1}\, h_{n-1} + u_{n-2}\, h_{n-2} + \ldots\ldots + u_{n-b+1}\, h_{n-b+1}) + u_{n-b}\, h_{n-b} + \\ + (u_{n-b-1}\, h_{n-b-1} + u_{n-b-2}\, h_{n-b-2} + \ldots\ldots + u_{n-2b}\, h_{n-2b}),$$

where $u_i \in \mathrm{GF}(q)$ and $u_{n-b} \neq 0$.

This condition ensures that no key error of length b or less will be a code word, which thereby means that the code will be able to detect key errors of length b or less.

The number of ways, in which the coefficients u_i can be selected, including the vector of all zeros, is

$$1 + q^{b-1}(q-1)q^b.$$

In the worst case, all these linear combinations might yield a distinct sum.

Therefore a column h_n can be added to H provided that

$$q^{n-k} > 1+ q^{2b-1}(q-1),$$

or,

$$n-k > \log_q[1+ q^{2b-1}(q-1)]. \qquad \blacksquare$$

Remark 1. It is worth mentioning that parity check digits $n–k$ do not depend on n. Thus the above theorem is valid for any value of n ($n > 2b$).

3. An illustration

Consider a (8, 4)-binary code with 4×8 parity check matrix H given by

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

This matrix has been constructed by the synthesis procedure outlined in the proof of Theorem 2, by taking $b = 2$, $n = 8$ over GF (2). It can be seen from Table 1 that the syndromes of any key error of length 2 or less are all nonzero. This shows that the code that is the null space of this matrix can detect all key errors of length 2 or less.

Table 1. Error patterns and syndromes

Error patterns	Syndromes	Error patterns	Syndromes
10000000	1000	00111100	1111
11000000	1100	01110100	0011
10100000	1010	01111100	1011
11100000	1110	00001000	1000
01000000	0100	00001100	1100
01100000	0110	00001010	1010
01010000	0101	00001110	1110
01110000	0111	00101100	1110
11010000	1101	00101010	1000
11110000	1111	00101110	1100
00100000	0010	00011010	1011
00110000	0011	00011110	1111
00101000	1010	00111010	1001
00111000	1011	00111110	1101
10110000	1011	00000100	0100
10101000	0010	00000110	0110
10111000	0011	00000101	0101
01101000	1110	00000111	0111
01111000	1111	00010110	0111
11101000	0110	00010101	0100
11111000	0111	00010111	0110
00010000	0001	00001101	1101
00011000	1001	00001111	1111
00010100	0101	00011101	1100
00011100	1101	00011111	1110
01011000	1101	00000010	0010
01010100	0001	00000011	0011
01011100	1001	00001011	1011
00110100	0111	00000001	0001

4. Simultaneous detection and correction of key errors

In this section we determine the extended Reiger's bound (refer to Reiger [9]; also Theorem 4.15 of Peterson and Weldon [8]) for simultaneous detection and correction of key errors. The following theorem gives a bound on the number of parity-check digits for a linear code that simultaneously detects and corrects such errors.

Theorem 3. An (n, k) linear code over GF(q) that corrects all key errors of length b or less must have at least

$$\begin{cases} \log_q(1+2b) & \text{for } q=2, \\ \log_q[1+2b(q-1)(q-2] & \text{for } q \neq 2 \end{cases}$$

parity-check digits.

Further, if the code corrects all key errors of length b or less and simultaneously detects all key errors of length d ($> b$) or less, then the code must have at least

$$\begin{cases} \log_q(1+b+d) & \text{for } q=2, \\ \log_q[1+(b+d)(q-1)(q-2] & \text{for } q \neq 2 \end{cases}$$

parity-check digits.

Proof: In order to prove the first part, consider a key error of length $2b$ or less. Such a vector is expressible as a sum or difference of two vectors, each of which is a key error of length b or less. These component vectors must belong to different cosets of the standard array because both such errors are correctable errors. Accordingly, such a vector, viz. a key error of length $2b$ or less cannot be a code vector. Applying Theorem 1, such a code must have at least

$$\begin{cases} \log_q(1+2b) & \text{for } q=2, \\ \log_q[1+2b(q-1)(q-2] & \text{for } q \neq 2 \end{cases}$$

parity-check digits.

For the second part, consider a key error of length $b + d$ or less. Such a vector is expressible as a sum or difference of two vectors, one of which is a key error of length b or less and the other is a key error of length d or less. Both component vectors, one being a detectable error and the other being a correctable error, cannot belong to the same coset of the standard array. Therefore, such a vector cannot be a code vector, i.e., a key error of length $b + d$ or less cannot be a code vector. Hence, according to Theorem 1, the code must have at least

$$\begin{cases} \log_q(1+b+d) & \text{for } q=2, \\ \log_q[1+(b+d)(q-1)(q-2] & \text{for } q \neq 2 \end{cases}$$

parity check digits.

5. Conclusion

This paper presents the bounds on parity checks for codes capable of detecting key errors. The bounds will determine the error-detection capability of a linear code. Correcting such errors will remain a further study, which the author will discuss in a future paper.

Acknowledgement: *The author is thankful to Dr. Vinod Tyagi and Prof. B. K. Dass, University of Delhi, for their suggestions.*

References

1. Das, P. K. Codes on *s*-Periodic Random Error of Length *b*. – Palestine Journal of Mathematics, 2014 (accepted for publication).
2. Das, P. K. Codes Detecting and Correcting Solid Burst Errors. – Bulletin of Electrical Engineering and Informatics, Vol. **1**, 2012, No 3, 225-232.
3. Das, P. K., V. Tyagi. Codes on *s*-Periodic Errors. – Ratio Mathematica-Journal of Applied Mathematics, Vol. **22**, 2012, 61-68.
4. Dass, B. K., P. Garg. On 2-Repeated Burst Codes. – Ratio Mathematica-Journal of Applied Mathematics, Vol. **19**, 2009, 11-24
5. Dass, B. K., R. Verma. Repeated Burst Error Detecting Linear Codes. – Ratio Mathematica-Journal of Applied Mathematics, Vol. **19**, 2009, 25-30.
6. Dass, B. K., R. Verma., L. Berardi. On 2-Repeated Burst Error Detecting Codes. – Journal of Statistical Theory and Practice, Vol. **3**, 2009, 381-391.
7. Hamming, R. W. Error-Detecting and Error-Correcting Codes. – Bell System Technical Journal, Vol. **29**, 1950, 147-160.
8. Peterson, W. W., E. J. (Jr.) Weldon. Error-Correcting Codes. 2nd Edition. The MIT Press, Mass., 1972.
9. Reiger, S. H. Codes for the Correction of Clustered Errors. – IRE Trans. Inform. Theory, **IT-6**, 1960, 16-21.
10. Sacks, G. E. Multiple Error Correction by Means of Parity-Checks. – IRE Trans. Inform. Theory, **IT-4**, 1958, 145-147.
11. Sharma, B. D., A. Gaur. Codes Correcting Limited Patterns of Random Errors Using S-K Metric. – Cybernetics and Information Technologies, Vol. **13**, 2013, No 1, 34-45.