MATEMATIČKI VESNIK
МАТЕМАТИЧКИ ВЕСНИК
70, 1 (2018), 89–94
March 2018

research paper
оригинални научни рад

ERROR LOCATING CODES AND EXTENDED HAMMING CODE

Pankaj Kumar Das

Abstract. Error-locating codes, first proposed by J. K. Wolf and B. Elspas, are used in fault diagnosis in computer systems and reduction of the retransmission cost in communication systems. This paper presents locating codes obtained from the famous Extended $(8, 4)$ Hamming code capable of identifying the sub-block that contains solid burst errors of length 2 (or 3) or less. We also make a comparison of information rate between the extended Hamming code and obtained codes. Further, comparisons in solid burst error detection and location probabilities of the codes over binary symmetrical channel are also provided.

1. Introduction and preliminaries

Wolf and Elspas [14] introduced a midway concept (known as error location coding) between error detection and error correction. Error locating (EL) codes have been found to be efficient in feedback communication systems. In such systems, the whole code length is divided into some finite number of sub-blocks which are mutually exclusive. Each sub-block of received digits is investigated for the presence of errors. If error is occurred within a sub-block, then the code has the capacity to locate the corrupted sub-block and the receiver can request the retransmission of the corrupted sub-block instead of the whole block, and this process is repeated for each incoming sub-block. In order to send large amount of data, long code length is desired to increase coding efficiency and which in turn results in a low information rate. The use of EL codes softens this deficiency by dividing long code length into smaller sub-blocks and maintain the system to keep the information rate up. Some of very recent works on error locating codes may be found in [5–7]. A good amount of work dealing with detecting and locating random/burst error can be found in [8] (specially Chapter 6 and Chapter 9).

The type of error occurred on communication channel depends on the behaviour of channel. Solid burst error is one type of error commonly found in many memory

2010 Mathematics Subject Classification: 94B05, 94B20, 94B60
Keywords and phrases: Parity check matrix, solid burst, syndrome, EL-codes, information rate, error probability.

communication channels viz. semiconductor memory data, supercomputer storage system [1–3, 11]. A solid burst may be defined as follows.

Definition 1.1. A solid burst of length b is a vector whose all the b-consecutive components are nonzero and rest are zero.

In what follows an (n, k) linear code is a proper subspace of n-tuples over $GF(q)$. The block of n digits, consisting of k information digits and $n-k$ parity check digits, is divided into s mutually exclusive sub-blocks. Each sub-block contains $t = \frac{n}{s}$ digits. The information rate (data rate) of an (n, k) linear code is $\frac{k}{n}$.

We consider (n, k) linear codes over $GF(q)$ that are capable of detecting and locating all solid bursts of length b or less within a single sub-block. Such an EL-code capable of identifying a single corrupted sub-block containing solid burst of length b or less must satisfy the following conditions:

(a) The syndrome resulting from the occurrence of a solid burst of length b or less within any one sub-block must be distinct from the all zero syndrome.

(b) The syndrome resulting from the occurrence of any solid burst of length b or less within a single sub-block must be distinct from the syndrome resulting likewise from such errors within any *other* sub-block.

The paper [4] studied codes that detect and locate all solid bursts of length b or less. The bounds on parity check digits for the existence of codes are obtained. This paper presents linear codes that are capable of detecting and locating (i) all solid bursts of length 2 or less (ii) all solid bursts of length 3 or less. The codes are obtained from the famous Extended $(8, 4)$ Hamming code (refer [10], also [13, pp. 117–119]). The study of this paper is motivated by the work done by Katti [12] where rearrangement of the columns of the parity matrix of a systematic $(16, 8)$ code (refer Gulliver and Bhargava [9]) gives rise to a code with better error detection and correction. We also give a comparison of information rates between the Extended Hamming code and our second type of codes. Further, we also provide comparisons of solid burst error detection and location probabilities among these codes over a binary symmetrical channel.

2. Code Construction

We start with the binary (8, 4) extended Hamming code that can correct all single errors and detect all double errors. The parity check matrix H of the code is given by

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

This code can not locate double errors. Now we construct a linear codes that is divided into two sub-blocks and is capable of locating solid burst of length 2 or less occurring within a sub-block. We rearrange the columns of H as follows and rename it H_1:

$$H_1 = \begin{bmatrix} 0 & 0 & 0 & 1 & & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & & 1 & 1 & 1 & 1 \end{bmatrix}$$

The null space of the matrix H_1 is a binary $(4+4, 4)$ linear code and is capable of detecting and locating all solid bursts of length 2 or less within a sub-block. This is because the conditions (a) and (b) are satisfied, i.e. the syndromes of all solid bursts of length 2 or less are nonzero and distinct within one sub-block, further the syndromes of such errors within one sub-block are distinct from the syndrome resulting likewise from any such errors within the other sub-block. It can be easily verified from the error pattern-syndrome table (as done in [4]). For location of errors, we proceed as follows. If the syndrome of any solid burst of length 3 or less is any one of the following: 0001, 0010, 0110, 1100, 0011, 0101, 1001, then the error can be located in the first sub-block. Again if the syndrome of such error is any one of the tuples 1011, 1111, 0111, 1101, 0100, 1000, 1010, then the location of the error is the second sub-block.

We now again give another construction from H and which gives rise to a class of linear codes that are capable of locating solid bursts of length 3 or less.

Rearrange the columns of the matrix H as $[h_1h_2h_3h_5h_4h_7h_6h_8]$ and then repeat the first two columns h_1, h_2 alternatively t times as $h_1h_2h_1h_2h_1h_2\ldots$ and consider as the first sub-block, again repeat the next two columns h_3, h_5 alternatively t times for the second sub-block and so on for other two pair of columns $(h_4, h_7), (h_6, h_8)$ for the third and fourth sub-blocks. Then the resulting matrix will give rise to a class of binary $(4t, 4t-4)$ $(t \geq 3)$ linear codes. The new matrix H_2 is given as follows:

$$H_2 = [\overbrace{h_1h_2h_1h_2}^{t} \quad \overbrace{h_3h_5h_3h_5\ldots}^{t} \quad \overbrace{h_4h_7h_4h_7\ldots}^{t} \quad \overbrace{h_6h_8h_6h_8\ldots}^{t}]$$

or,
$$H_2 = \begin{bmatrix} \overbrace{0000\ldots}^{t} & \overbrace{0101\ldots}^{t} & \overbrace{0101\ldots}^{t} & \overbrace{1111\ldots}^{t} \\ 0000\ldots & 1010\ldots & 1111\ldots & 0101\ldots \\ 0101\ldots & 0000\ldots & 1010\ldots. & 1111\ldots \\ 1111\ldots & 1111\ldots & 1111\ldots & 1111\ldots \end{bmatrix}$$

The null space of the matrix H_2 will detect and locate all solid bursts of length 3 or less. This claim is also true as we can verify that the syndromes of all solid bursts of length 3 or less are being nonzero and distinct within one sub-block, further the syndromes of such errors within one sub-block are distinct from the syndrome resulting likewise from any such errors within any other sub-block. For location of errors, we proceed as follows. If the syndrome of any solid burst of length 3 or less is any one of the following: 0001, 0011, 0010, then the code will locate the error in the first sub-block. Again if the syndrome of such error is any one of the tuples

0101, 1001, 1100, then the location of the error is the second sub-block. In the same way, if the syndrome is 0111 or 1101 or 1010, the faulty sub-block is the third sub-block. If syndrome is any of 1011, 1111, 0100, the sub-block with errors is the fourth one.

3. Comparisons of codes with respect to information rate and error probability

An error detecting code can only detect the presence of errors in the received vector, whereas an error locating code can also indicate the position of error and furthermore an error correcting code can correct the errors present in the received vector. As the purpose of the three types of codes to handle errors is varying, so different types of codes are to be constructed accordingly, but in the construction of codes, one has also to keep in mind the information rate. The more is the information rate, the more is the speed of the system which transmits the data. Further, as not all errors can be detected/located, so it is always important to know the probability of errors going undetected/unlocated despite the use of error detection/location scheme.

In this section, we establish a comparison of information rates among the extended $(8, 4)$ Hamming code, the (4+4, 4) code and the new class of $(4t, 4t-4)$ codes. Then, a comparative study of solid burst error undetection and unlocation probability of these codes is followed.

In Table 1 below, we put the information rates of the $(4t,\ 4t-4)$ codes for different values of t. As the information rate of the extended (8, 4) Hamming code as well as the $(4+4,\ 4)$ code is 0.5, we can conclude that the new class of $(4t,\ 4t-4)$ codes has better information rate than the extended (8, 4) Hamming code or $(4+4,\ 4)$ code.

For comparison of solid burst error undetecting/unlocating probability of the codes, let us consider a binary symmetrical channel (BSC) with error probability p.

For the (8, 4) extended Hamming code, we see that the code can detect any solid burst of length 3 or less and probability that solid burst goes undetected is $5p^4(1-p)^4+4p^5(1-p)^3+3p^6(1-p)^2+2p^7(1-p)^1+p^8 \cong 5p^4(1-p)^4$, we can ignore other terms for small value of p. Further, this code can locate only single errors within a sub-block of length 4, so the probability that solid burst of length $2, 3, 4$ can not be located within a sub-block is $6p^2(1-p)^8+4p^3(1-p)^5+2p^4(1-p)^4 \cong 6p^2(1-p)^8$.

For the $(4+4,\ 4)$ code, the code can detect solid bursts of length 3 or less, so the probability that solid burst goes undetected is same as that of the (8, 4) extended Hamming code i.e. $\cong 5p^4(1-p)^4$. But the code can locate solid bursts of length 2 or less occurring within a sub-block of length 4, the probability that solid burst error of length 3 and 4 goes unlocated is $4p^3(1-p)^5+2p^4(1-p)^4 \cong 4p^3(1-p)^5$.

For the $(4t, 4t-4)$ code, it can also detect solid bursts of length 3 or less, so the probability that solid burst goes undetected is given by $(4t-3)p^4(1-p)^{4t-4}+(4t-4)p^5(1-p)^{4t-5}+(4t-5)p^6(1-p)^{4t-6}+\cdots+(p)^{4t} \cong (4t-3)p^4(1-p)^{4t-4}$. As solid bursts of length 3 or less occurring within a sub-block of length t can be located, so the probability of not able to locate solid burst errors of length 4 upto t by the code is

$4\{(t-3)p^4(1-p)^{4t-4}+(t-4)p^5(1-p)^{4t-5}+(t-5)p^6(1-p)^{4t-6}+\cdots+(p)^t(1-p)^{3t}\} \cong 4(t-3)p^4(1-p)^{4t-4}$.

Let us assume the value of p is 0.01. In Table 1 the probabilities of solid burst error going undetected and unlocated by the $(4t, 4t-4)$ codes for different values of t are listed.

Table 1: The probabilities of solid burst error going undetected and unlocated by the $(4t, 4t-4)$ codes

t	Information rate for the $(4t, 4t-4)$ codes (appr. value)	Solid burst error undetecting probability for the $(4t, 4t-4)$ codes (appr. value) for $p = 0.01$	Solid burst error unlocating probability for the $(4t, 4t-4)$ codes (appr. value) for $p = 0.01$
3	0.667	0.0000000830470224985	0.0000000000000000000
4	0.750	0.0000001152300333231	0.0000000354553948686
5	0.800	0.0000001447478210861	0.0000000681166216876
6	0.833	0.0000001717604568954	0.0000000981488325117
7	0.857	0.0000001964195352018	0.0000001257085025292
8	0.875	0.0000002188685932891	0.0000001509438574407
9	0.889	0.0000002392435108661	0.0000001739952806299

The probability that solid burst goes undetected for the (8, 4) extended Hamming code or for the $(4+4,\ 4)$ code for $p = 0.01$ is same i.e. 0.000000048. But the probability that solid burst goes unlocated for the (8, 4) extended Hamming code is 0.0005536468 and the probability that solid burst goes unlocated for the $(4+4,\ 4)$ code is 0.000003804. Thus, for location point of view of solid burst, the $(4+4,\ 4)$ code is a better code than the (8, 4) extended Hamming code. Further, from the table we can say that (8, 4) extended Hamming code or the $(4+4,\ 4)$ code has better detection rate of solid burst error than $(4t, 4t-4)$ codes, but $(4t, 4t-4)$ codes has better location rate of solid burst error than the $(4+4,\ 4)$ code as well as (8, 4) extended Hamming code. Therefore, this new class of binary $(4t, 4t-4)$ codes will be more useful if the purpose is to detect and locate solid burst error.

Remark 3.1. Although the value solid burst error unlocating probability of $(4t, 4t-4)$ codes is increasing which can be seen in Table 1, but its solid burst error unlocating probability is always lesser than that of the $(4+4,\ 4)$ code. This is because of $4(t-3)p^4(1-p)^{4t-4} < 4p^3(1-p)^5$ i.e. $(t-3)p(1-p)^{4t-9} < 1$, for $t \geq 3$ and small value of p. We can verify this by Excel Software.

4. Conclusion

This paper gives the construction of EL codes that can detect and locate solid burst errors of 2 (or 3) or less. The obtained class of $(4t, 4t-4)$ codes is found to have better information rate, locating capability, error location rate for solid burst error point of view. One may work on to obtain EL codes based on other standard codes that can detect and locate solid bursts of length $b(>3)$ or less.

Acknowledgement. The author is very much thankful to anonymous referee(s) for careful reading of the manuscript and for valuable suggestions which improves the paper a lot.

References

[1] C. A. Argyrides, P. Reviriego, D. K. Pradhan, J. A. Maestro, *Matrix-based codes for adjacent error correction*, IEEE Transactions on Nuclear Science, **57** (4) (2010), 2106–2111.

[2] J. Arlat, W. C. Carter, *Implementation and evaluation of a* (b,k)*-adjacent error-correcting/detecting scheme for supercomputer systems*, IBM J. Res. Develop. **28** (2) (1984), 159–169.

[3] D. C. Bossen, *b-adjacent error correction*, IBM Journal of Research and Development, **14** (4) (1970), 402–408.

[4] P. K. Das, *Codes detecting and locating solid burst errors*, Romanian Journal of Mathematics and Computer Science, **2** (2) (2012), 1–10.

[5] P. K. Das, L. K. Vashisht, *Error locating codes by using blockwise-tensor product of blockwise detecting / correcting codes*, Khayyam J. Math., **2** (1) (2016), 6–17.

[6] B. K. Dass, S. Madan, *Repeated low-density burst error locating codes*, Acta Universitatis Apulensis, **33** (2013), 175–191.

[7] Ji-Hao Fan, Han-Wu Chen, *A construction of Quantum Error-Locating codes*, Communications in Theoretical Physics, **67** (1) (2017), 37–40.

[8] F. Fujiwara, Code Design for Dependable Systems, John Willey & Sons, 2006.

[9] T. A. Gulliver, V. K. Bhargava, *A systematic* $(16,8)$ *code for correcting double errors and detecting triple-adjacent errors*, IEEE Trans. Computers, **42** (1) (1993), 109–112.

[10] R. W. Hamming, *Error-detecting and error-correcting codes*, Bell System Technical Journal, **29** (1950), 147–160.

[11] D. W. Jensen, *Block code to efficiently correct adjacent data and/or check bit errors*, Patent number: US 6604222 B1, Date of Patent Aug 5, 2003, (www.google.com/patents/US6604222).

[12] R. S. Katti, *Comments on a systematic* $(16,8)$ *code for correcting double errors and detecting triple-adjacent errors*, IEEE Transactions on Computers, **44** (12) (1995), 1472–1473.

[13] W. W. Peterson, E. J. Weldon(Jr.), *Error-Correcting Codes*, 2nd edition, The MIT Press, Mass, 1972.

[14] J. Wolf, B. Elspas, *Error-locating codes a new concept in error control*, IEEE Transactions on Information Theory, **9** (2) (1963), 113–117.

(received 31.03.2017; in revised form 07.09.2017; available online 31.10.2017)

Department of Mathematical Sciences, Tezpur University, Napaam, Sonitpur, Assam-784028, India

E-mail: pankaj4thapril@yahoo.co.in, pankaj4@tezu.ernet.in