ANALYSIS ON (15,7) BINARY BCH ENCODER AND DECODER FOR 7-BIT ASCII CHARACTERS

P.Mozhiarasi*

C.Gayathri**

Abstract-*In digital communication, the data is transmitted through the channel to the receiver side. If noise is included in the channel, data that is received at the receiver side will not be the actual data. This may create a corruption of data which is intolerable. In such situations, an error control must be employed so that errors may be detected and afterwards corrected. In this paper, (15,7,2) BCH encoder and decoder are designed mathematically. The block length(n), data length(k) and maximum number of correctable errors(t) are 15, 7 and 2 respectively. The encoding and decoding operates over the Galois Field GF (2^4) with an irreducible polynomial x^4+x+1. Decoder includes 3 major steps: 1) Syndrome calculation (SC) 2) Berlekamp Massey Algorithm (BMA) 3) Chein Search (CS). The error is then corrected effectively in decoder part with the help of NAND Flash memory.*

Keywords: *Bose Chaudhuri Hocquenghem (BCH), BCH Encoder, BCH Decoder, Berlekamp Massey Algorithm (BMA), Chein Search(CS), Galois Field (GF), Syndrome Calculation(SC).*

***ME Student VLSI Design, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India

**Asst. Prof Department of ECE, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India

1. INTRODUCTION:

Millions of tiny devices are combined to form the computer memory, each of which can store a 0 or a 1 by magnetically on a disk or as state in a transistor, a core or a vacuum tube. The word "bit" is used to refer to one of these individual atoms of memory. Bits are combined to form some meaningful data. When eight bits are grouped together they are represented as a byte. Using blocks of bits the computer memory can store numbers, text and even pictures. Bits in a byte may be normally numbered from zero to seven, where bit 0 indicates lower order bit or Least Significant Bit (LSB) and bit 7 indicates higher order bit or Most Significant Bit (MSB). All the other bits are referred by their respective number. When the data is transmitted in a noisy channel, it is difficult to retrieve actual data at receiver side. A single error may shutdown the whole processing unit and result in an unbearable loss of data. There are many different coding schemes [1], each having some particular advantages and characteristics. The more sophisticated error correcting code [5] is Bose, Chaudhuri and Hocquenghem (BCH) codes which are overview of the Hamming codes.

BCH code is a general algorithm to correct a small bit error with wide range of applications in digital communication s and storage [2]. They include BCH encoder and decoder with three processing stages 1) Syndrome calculation (SC) 2) Berlekamp Massey Algorithm (BMA) 3) Chein Search (CS). At last, certain errors are corrected automatically and enable reconstruction of the original data.

2. BCH CODES

BCH codes [3] are multiple random error-correcting codes which were discovered by A. Hocquenghem in 1959. It was continued by R. C. Bose and D. K. Ray Chaudhuri in 1960. BCH codes are known as cyclic codes which means that for any cyclic shift, the example of codeword is a valid codeword. BCH codes were also called as primitive BCH codes. As the BCH code operate in Galois Field, it can be defined by two parameters that are length of code words (n)and the number of error to be corrected t. BCH encoder and decoder block diagram is shown in Fig. 1. Fundamentally, BCH codes are defined by

Length of codeword, $n = 2^m-1$

Number of parity check bits, $n-k \leq mt$

Minimum distance, $d_{min} \geq 2t+1$

where, m – the degree or integer of the generating polynomial *k* – length of information

t – maximum number of correctable errors

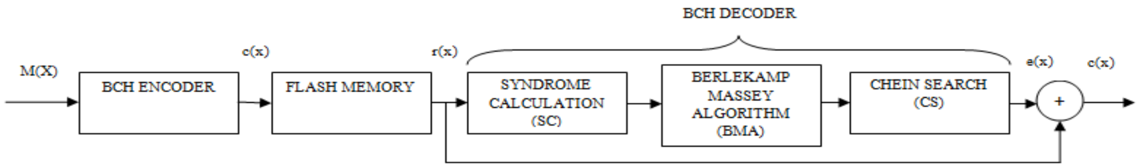

Figure 1 Block Diagram of BCH Encoder and Decoder

Consider a text message "VLSI" should be sent through the channel without any interruption. But it is not possible to receive the actual data when they are passed through noisy channel. Each character is sent one by one with the help of (Linear Feedback Shift Register) LFSR in a network. The system to encode letters and characters into numbers is called the ASCII code (American Standard for Information Interchange).The transmission on computer and radio networks uses binary symbols. By the American Standard Code for Information Interchange the letter "V" is defined by a binary value of 1010110. Similarly, for L,S, I represents 1001100, 1010011 and 1001001 respectively. A sentence is then represented by a string of these values: "VLSI" = 1010110, 1001100, 1010011, 1001001. To detect errors in the transmission of such a string of ones and zeros, then it is better to make a coding system that works with binary symbols.

Assume that 7 bit ASCII character is transmitted through noisy channel there is a chance of error occurrence. In order to detect the error in communication append them with eight zeros before transmitting [6]. This process occurs in encoder part, where Combination of message bit and redundant bit is called codeword. At the receiver side, decoder includes three important steps to detect and correct the bits if error is present.

Table 1: *GF(16)* generated with Primitive polynomial x^4+x+1

Power Form	Polynomial Form	4-Tuple Form $\alpha^3, \alpha^2, \alpha, 1$	Decimal Form	Minimal polynomial
0	0	0000	0	x
$1, \alpha^{15}$	1	0001	1	$x+1$
α	α	0010	2	x^4+x+1 $(f_1(x))$
α^2	α^2	0100	4	x^4+x+1 $(f_2(x))$
α^3	α^3	1000	8	$x^4+x^3+x^2+x+1 (f_3(x))$
α^4	$\alpha+1$	0011	3	x^4+x+1 $(f_4(x))$
α^5	$\alpha^2+\alpha$	0110	6	x^2+x+1 $(f_5(x))$
α^6	$\alpha^3+\alpha^2$	1100	12	$x^4+x^3+x^2+x+1 (f_6(x))$
α^7	$\alpha^3+\alpha+1$	1011	11	x^4+x^3+1
α^8	α^2+1	0101	5	x^4+x+1
α^9	$\alpha^3+\alpha$	1010	10	$x^4+x^3+x^2+x+1$
α^{10}	$\alpha^2+\alpha+1$	0111	7	x^2+x+1
α^{11}	$\alpha^3+\alpha^2+\alpha$	1110	14	x^4+x^3+1
α^{12}	$\alpha^3+\alpha^2+\alpha+1$	1111	15	$x^4+x^3+x^2+x+1$
α^{13}	$\alpha^3+\alpha^2+1$	1101	13	x^4+x^3+1
α^{14}	α^3+1	1001	9	x^4+x^3+1

Let us consider $m=4$, $n=2^m -1= 2^4-1=15$, $k=7$, $t=2$, $d_{min}=2t+1=2(2)+1=5$. Let $f_i(x)$, $i= 1,2,3,\ldots,2t$ be the minimal polynomial, then $g(x) = LCM\{f_1(x),f_2(x),f_3(x),\ldots,f_{2t}(x)\}$. But, $g(x)$ is simplified to $g(x) = LCM\{f_1(x),f_3(x),\ldots,f_{2t-1}(x)\}$ because every even power of primitive element will have same minimal polynomial as some odd power of the elements having the number of factors in the polynomial [7]. Therefore, $g(x) = LCM\{f_1(x),f_3(x)\} = x^8+x^7+x^6+x^4+1$= 111010001 where $f_1(x) = x^4+x+1$= 10011and $f_3(x) = x^4+x^3+x^2+x+1$= 11111. These factors are obtained by primitive polynomial x^4+x+1 which are shown in Table 1 and conjugates of 16 field elements using Galois Field theory [8] are shown in Table 2.

Table 2 Minimal Polynomials of the element in *GF(16)*

	Conjugate roots	Minimal polynomial
	0	x
	1	$x+1$
$\varphi_1(x)$	$\alpha,\alpha^2,\alpha^4,\alpha^8$	x^4+x+1
$\varphi_2(x)$	$\alpha^3,\alpha^6,\alpha^9,\alpha^{12}$	$x^4+x^3+x^2+x+1$
$\varphi_3(x)$	α^5,α^{10}	x^2+x+1
$\varphi_4(x)$	$\alpha^7,\alpha^{11},\alpha^{13},\alpha^{14}$	x^4+x^3+1

Each character process is organized as follows:

2.1 Transmission of letter “V”

2.1.1 BCH Encoder

In initial stage, the message polynomial*(M(x))* is padded with *(n-k)* zero bits [13] to form X^{n-k} $M(x)$ i.e., $X^{n-k}\ M(x) = X^8\ M(x) = M + 00000000$=101011000000000. The remainder polynomial*(R(x))* is obtained by dividing the $X^8\ M(x)$ by *g(x)*.Combining remainder polynomial with message polynomial produce the codeword (*c(x)*) [4]. i.e., $c(x) = X^{n-k}\ M(x) + R(x)$ = 101011001000111

2.1.2 BCH Decoder

2.1.2.1 Syndrome Calculation:

Let us consider *c(x)* with 2 errors ($t = 2$) and represent as *r(x)* i.e., when noise is added in the channel while transmitting the data to the receiver [10]. At the receiver side, let us assume *r(x)* =10**0**011001**0**10111. The number of syndrome elements is $2*t = 4$ them, to find $t=2$ errors. Those syndrome elements are represented as S_1, S_2, S_3, S_4 and they can be calculated by

$S_1(\alpha) = r(\alpha) \bmod f_1(\alpha) = \alpha^3 + \alpha^2 = \alpha^6$

$S_2(\alpha^2) = r(\alpha^2) \bmod f_2(\alpha^2) = r(\alpha^2) \bmod f_1(\alpha^2) = (\alpha^2)^3+(\alpha^2)^2 = \alpha^{12}$

$S_3(\alpha^3) = r(\alpha^3) \bmod f_3(\alpha^3) = (\alpha^3)^3+(\alpha^3)+1= \alpha^4$

$S_4(\alpha^4) = r(\alpha^4) \bmod f_4(\alpha^4) = r(\alpha^4) \bmod f_1(\alpha^4) = (\alpha^4)^3+(\alpha^4)^2 = \alpha^9$

By the property of Field element several powers of generating element will have the same minimal polynomial. When f(x) is polynomial over *GF(2)*, α is an element of $GF(2^m)$. If the remainder equals zero, then it is declared that no error in the codeword or else error(*S*) is present [9]. So this error will be preceded for finding error location.

2.1.2.2 Berlekamp Massey Algorithm:

This method has been reduced by a Berlekamp algorithm [11] and error locator polynomial is obtained by iterative method. d_μ represents discrepancy value in Lin and Costello table as shown in Table 3.

Table 3 Finding error locator polynomial

μ	$\sigma^{(\mu)}(x)$	d_μ	l_μ	$2\mu - l_\mu$
-1/2	1	1	0	-1
0	1	α^6	0	0
1	α^6x+1	α^7	1	1
2	$\alpha x^2+\alpha^6x+1$	-	-	-

The “Key Equation” is given by

$\sigma^{(\mu+1)}(x) = \sigma^{\mu}(x) + d_\mu\, d_\rho^{-1} x^{2(\mu-\rho)} \sigma^{(\rho)}(x)$ ---------equation 1

$l_{\mu+1} = L = deg(\sigma^{(\mu+1)}(x))$ --------equation 2

$d_{\mu+1} = S_{2\mu+3} + \sigma_1^{(\mu+1)} S_{2\mu+2} + \sigma_2^{(\mu+1)} S_{2\mu+1} + \cdots + \sigma_L^{(\mu+1)} S_{2\mu+3-L}$--------equation 3

Steps:

(i) Initialize μ = 0, $d_\mu \neq 0, \rho =$ -1/2

(ii) Substitute $\sigma^{(\mu)}(x) = \sigma^{(0)}(x) = 1$, $d_\mu = d_0 = S_1 = \alpha^6$, $d_\rho^{-1} = d_{-\frac{1}{2}}^{-1}$ =1,

$\sigma^{(\rho)}(x) = \sigma^{\left(-\frac{1}{2}\right)}(x) = 1$ in equation 1

$\sigma^{(1)}(x) = \sigma^{0}(x) + d_0\, d_{-1/2}^{-1} x^{2\left(\frac{1}{2}\right)} \sigma^{\left(\frac{-1}{2}\right)}(x)$

= $1 + \alpha^6(1)^{-1}x(1)$

= $\boldsymbol{\alpha^6 x + 1}$

(iii) Substitute $\sigma^{(1)}(x) = \alpha^6 x + 1$ in equation 2

$l_1 = \deg(\sigma^{(1)}(x)) = \deg(\alpha^6 x + 1) = \mathbf{1}$

(iv) Substitute $S_3 = \alpha^4$, $S_2 = \alpha^{12}$, $\sigma_1^{(1)} = \alpha^6$ in equation 3

$d_1 = S_3 + \sigma_1^{(1)} S_2$

$= \alpha^4 + (\alpha^6)(\alpha^{12})$

$\mathbf{d_1 = \alpha^7}$

(v) In similar way, consider μ = 1, $d_\mu \neq 0$, $\rho = 0$ and substitute

$\sigma^{(\mu)}(x) = \sigma^{(1)}(x) = \alpha^6 x + 1, d_\mu = d_1 = \alpha^7, d_\rho^{-1} = d_0^{-1} =$ (α6)-1, $\sigma^{(\rho)}(x) = \sigma^{(0)}(x) =$ 1 *in equation 1* $\sigma^{(2)}(x) = \sigma^{1}(x) + d_1\, d_0^{-1} x^{2(1)} \sigma^{(0)}(x)$

= *(α6x+1)+ (α7)(α6)-1 x2(1)*

$\sigma^{(2)}(x) = \boldsymbol{\alpha x^2 + \alpha^6 x + 1}$

The last polynomial represents the final error locator polynomial.

2.1.2.3 Chien Search Algorithm:

The roots of $\sigma^{(\mu)}(x)$ in GF (2^4) should be found out by following order [12]. If $\sigma^{(\mu)}(x) = \sigma^{(2)}(x) = \alpha x^2 + \alpha^6 x + 1 = 0$ is obtained by substituting x = 0,1,α, α^2, α^3, ,..., α^{14} then they are considered as roots.

Example: $\sigma^{(2)}(0) = \alpha(0)^2 + \alpha^6(0) + 1 = 1 \neq 0$

$\sigma^{(2)}(1) = \alpha(1)^2 + \alpha^6(1) + 1 = \alpha + (\alpha^3 + \alpha^2) + 1 \neq 0$

…

$\sigma^{(2)}(\alpha^{3})$ = α $(\alpha^3)^2+\alpha^6(\alpha^3)+1$ = $\alpha^7+\alpha^9+1 = (\alpha^3+\alpha+1)+(\alpha^3+\alpha)+1 = 0$

…

$\sigma^{(2)}(\alpha^{11})$ = α $(\alpha^{11})^2+\alpha^6(\alpha^{11})+1 = \alpha^{23}+\alpha^{17}+1 = (\alpha^2+1)+(\alpha^2)+1 = 0$

…

Therefore, α^3, α^{11} are the roots for $\sigma^{(2)}(x) = \alpha x^2+\alpha^6 x+1$. The bit position of error location will be the inverse of their roots (α^3, α^{11}) i.e., 00**1**0000000**1**0000

Thus, the error pattern polynomial can be written as $e(x) = x^{12}+x^4$.

c(x) = r(x) + e(x)

= 100011001010111 + 001000000010000

$= \underbrace{10101100}_{M(x)}\underbrace{1000111}_{R(x)}$

As a result, the actual data is recovered by performing modulo-2 addition for r(x) and e(x).

2.2 Transmission of letter "L", "S", "I"

<table>
<tr><th></th><th>L</th><th>S</th><th>I</th></tr>
<tr><td>Binary value</td><td>1001100</td><td>1010011</td><td>1001001</td></tr>
<tr><td>Remainder</td><td>10011</td><td>1110000</td><td>100100</td></tr>
<tr><td>Codeword c(x)</td><td>100110000010011</td><td>101001101110000</td><td>100100100100100</td></tr>
<tr><td>Received polynomial r(x)</td><td>100111001010011</td><td>101001111111000</td><td>110101100100100</td></tr>
<tr><td>Syndrome Calculation(SC): Syndrome (S_1, S_2, S_3, S_4)</td><td>$\alpha^5, \alpha^{10}, \alpha^{10}, \alpha^5$</td><td>$\alpha^4, \alpha^8, \alpha^5, \alpha^2$</td><td>$\alpha^{10}, \alpha^5, \alpha^5, \alpha^{10}$</td></tr>
<tr><td>Berlekamp Massey Algorithm(BMA):

Final error locator polynomial ($\sigma^{(\mu)}(x)$)</td>
<td><table>
<tr><th>μ</th><th>$\sigma^{(\mu)}(x)$</th><th>d_μ</th><th>l_μ</th><th>$2\mu-l_\mu$</th></tr>
<tr><td>-1/2</td><td>1</td><td>1</td><td>0</td><td>-1</td></tr>
<tr><td>0</td><td>1</td><td>α^5</td><td>0</td><td>0</td></tr>
<tr><td>1</td><td>α^5x+1</td><td>α^5</td><td>1</td><td>1</td></tr>
<tr><td>2</td><td>$x^2+\alpha^5x+1$</td><td>-</td><td>-</td><td>-</td></tr>
</table>
$\mathbf{x^2+\alpha^5x+1}$</td>
<td><table>
<tr><th>μ</th><th>$\sigma^{(\mu)}(x)$</th><th>d_μ</th><th>l_μ</th><th>$2\mu-l_\mu$</th></tr>
<tr><td>-1/2</td><td>1</td><td>1</td><td>0</td><td>-1</td></tr>
<tr><td>0</td><td>1</td><td>α^4</td><td>0</td><td>0</td></tr>
<tr><td>1</td><td>α^4x+1</td><td>α^{14}</td><td>1</td><td>1</td></tr>
<tr><td>2</td><td>$\alpha^{10}x^2+\alpha^4x+1$</td><td>-</td><td>-</td><td>-</td></tr>
</table>
$\mathbf{\alpha^{10}x^2+\alpha^4x+1}$</td>
<td><table>
<tr><th>μ</th><th>$\sigma^{(\mu)}(x)$</th><th>d_μ</th><th>l_μ</th><th>$2\mu-l_\mu$</th></tr>
<tr><td>-1/2</td><td>1</td><td>1</td><td>0</td><td>-1</td></tr>
<tr><td>0</td><td>1</td><td>α^{10}</td><td>0</td><td>0</td></tr>
<tr><td>1</td><td>$\alpha^{10}x+1$</td><td>α^2</td><td>1</td><td>1</td></tr>
<tr><td>2</td><td>$\alpha^7x^2+\alpha^{10}x+1$</td><td>-</td><td>-</td><td>-</td></tr>
</table>
$\mathbf{\alpha^7x^2+\alpha^{10}x+1}$</td></tr>
<tr><td>Chein Search(CS): Roots
Error pattern polynomial e(x)</td><td>α^6, α^9
x^9+x^6</td><td>α^8, α^{12}
x^7+x^3</td><td>α^2, α^6
$x^{13}+x^9$</td></tr>
<tr><td>Recovered data (c(x) = r(x) + e(x))</td><td>100110000010011</td><td>101001101110000</td><td>100100100100100</td></tr>
</table>

From above table it is understood that the letters "L", "S", "I" which are received with errors at receiver part is detected by Syndrome Calculation (SC) and corrected with help of Berlekamp Massey Algorithm (BMA) and Chein Search (CS).

Thus, a text message "VLSI" that was sent through the channel reaches the other side of communication channel without any mismatch in text.

3. CONCLUSION:

Thus, by the use of 7 bit ASCII characters, the numbers, text and even pictures can be sent from one side to other side. While transmission is carried out, the actual data is recovered at the receiver side with the help of Berlekamp Massey Algorithm (BMA) and Chein Search (CS) even when the data is corrupted. Therefore, (15,7)BCH encoder and decoder [14] are derived mathematically which detects and corrects upto two errors in a efficient manner and implemented in future using hardware description language (HDL) known as Verilog, VHDL and synthesized by Xilinx ISE 13.2 simulator.

REFERENCES:

[1] Berlekamp.E.R, "Algebraic Coding Theory", Mc GrawHill, New York, 1968.

[2] Berlekamp, E.R., Peile, R.E. and Pope, S.P. (1987), "The application of error control to communications", IEEE Communication

[3] Hank Wallace, "Error Detection and Correction using BCH Codes", Copyright © 2001 Hank Wallace.

[4] http://en.wikipedia.org/wiki/Polynomial_long_division, an article on polynomial long division in Wikipedia.

[5] Lin. S and Costello Jr. D.J, "Error Control Coding", Prentice-Hall, New Jersey, 1983.

[6] M.Y. Rhee - "Error Correcting Coding Theory", McGraw-Hill, Singapore, 1989.

[7] P.Mozhiarasi , Ms.C.Gayathri, "A Mathematical Derivation for Error Correction and Detection in Communication using BCH Codes", IPASJ International Journal of Electronics & Communication, Vol. 2, Issue 10, p. 7-12, Oct. 2014.

[8] R.E. Blahut, "Theory and practice of error-control codes", Addison-Wesley, Reading, MA, 1983.

[9] Ryoh Fuji-Hara, "Galois Field Manual", University of Tsukuba, Second edition, June 2008.

[10] S. B. Wicker, "Error Control Systems for Digital Communication and Storage". Upper Saddle River, New Jersey 075458: Prentice Hall, Inc1995.

[11] Shu Lin, Daniel J. Castello, "Error control coding: Fundamentals and applications", Prentice-Hall, New Jersey, 1983.

[12] Simon Haykin, "Communication Systems", 4th Edition, John Wiley & Sons, Inc.

[13] W.W.Peterson, “Encoding and error-correction procedures for the Bose-Chaudhuri Codes”, IRE Trans. Inf. Theory, September 1960.

[14] Yunghsiang S. Han, “BCH Codes”, Graduate Institute of Communication Engineering, National Taipei University Taiwan.

Authors

P. Mozhiarasi received her Bachelor’s degree in Electronics and Communication Engineering from Dr. Mahalingam College of Engineering and Technology, Pollachi, Coimbatore District, Tamil Nadu in 2013 and pursuing Master’s degree (VLSI Design) in Sri Eshwar College of Engineering, Kinathukadavu, Coimbatore, Tamil Nadu.

Ms. C.Gayathri received her Bachelor’s degree in Electronics and Communication Engineering from Sengunthar Engineering College, Tiruchengode in 2010 and Master’s degree in Communication Engineering from Karpagam College of Engineering, Coimbatore in 2012. She joined as an Assistant Professor, Electronics and Communication Engineering in Sri Eshwar College of Engineering, Kinathukadavu, Coimbatore, Tamil Nadu in 2012. She is well-versed in Computer Networks.