iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 Issue: XII Month of publication: December 2015

DOI:

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

A Comparative Study between Hamming Code and Reed-Solomon Code in Byte Error Detection and Correction

Chukwuma Okeke[1], M.Eng[2]

[1,2]*Department of Electrical/Electronics Engineering, Michael Okpara University of Agriculture Umudike. Abia State, Nigeria*

Abstract:* This work concerns the comparative study between Hamming code and Reed-Solomon (RS) code in byte error detection and correction. Data are either stored in storage applications or transferred through a media. In either case there is problem of data corruption due to noise and interference, which has to be corrected as soon as they occur to prevent erroneous information throughout the system. The basis is to incorporate the original data with some form of redundancy. With this redundancy, even if errors are introduced, the original information can be recovered or at least the presence of error can be detected. Redundancy is added by the encoder and sent over the channel where noise in the form of error distorts the data. The message reaches the receiving end through the decoder where errors are removed, redundancy stripped off and the original message is then produced. This paper compared Hamming code and Reed-Solomon code in respect to their efficiency in adding and removing this redundancy. I find Reed-Solomon code to be highly efficient and more powerful than Hamming code for data communication because of it high coding rate and yet with a minimum coding complexity.

Keyword: Channel, Encoder and Decoder, Error detection and correction, Hamming code, Reed-Solomon code.

I. INTRODUCTION

In this age of information technology, there is an increase demand for efficient and reliable digital data storage and transmission systems. A major concern is to control the error so that a reliable reproduction of data can be obtained. Information media are not 100% reliable in practice, in the sense that interference and noise frequently course data to be distorted. They had always been present but the effect has heightened with increase in technology. To deal with this undesirable but inevitable situation, error correcting codes (ECC) are introduced [1]. The ECC are clever ways of introducing redundancy so that the original information can be recovered even when parts of the data have been corrupted. ECC adds redundancy to the original massage in such a way that it is possible for the receiver to detect the error, correct it and recover the original message. With this redundancy, even if errors are introduced the original information can be recovered or at least the presence of error can be detected. This is crucial for certain applications where resending of the message is not possible, such as in interplanetary communications and storage of data [2], [3]. The issue on how to add this redundancy in order to detect and correct as many errors as possible in the most efficient way makes some of the ECC more efficient than others in some application. This brought about different error detection and error correction scheme [4]. Efficient codes are designed by keeping the ratio of the parity-check bits to the data bits and making processing time involved in encoding and decoding the data stream as minimum as possible [5]. The most efficient code for a particular application is one where the code rate (R) approaches numeric one (R →1) [6]. Error correcting codes are therefore inevitably applied to digital communication for the task of improving reliability.

Fig. 1 below illustrates a communication channel.

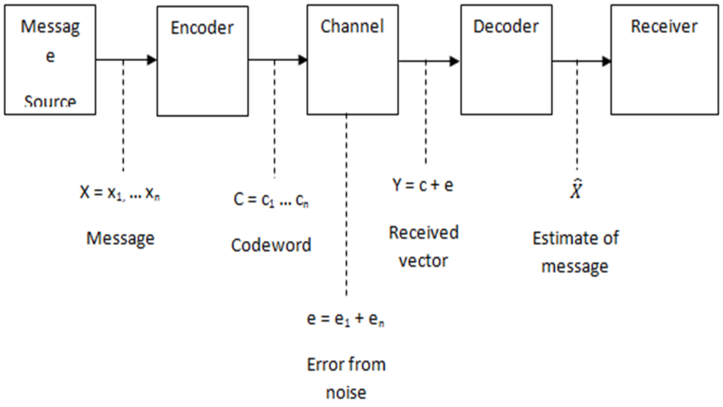

Fig.1 Communication Channel

At the source, a message denoted by X is sent as shown in fig.1. If no modification is made to the massage and it is transmitted directly over the channel, any noise will distort the message so that it is not recoverable. The basic idea is to embellish the message by adding some redundancy to it so that, hopefully the received message is the original message that was sent. The redundancy is added by the encoder and the embellished message called Codeword C in fig.1 is sent over the channel where noise in the form of error vector e distorts the codeword, producing a received vector Y . The received vector is then sent to the decoder where the errors are removed, the redundancy is stripped off and an estimate $\hat{X}$ of the original message is produced. Hopefully $\hat{X} = X$. The job of the decoder is to obtain an estimate $\hat{Y}$ of Y and hoping that $\hat{Y}$ = C.

Error correcting codes guarantees that our hope will be fulfilled to a certain percentage using the right encoding and decoding scheme. Based on the characteristics of the channel this percentage can be made as high as we desire though not 100% [7].

II. REVIEW OF RELATED TERMS

A. CodeWord

Codeword refers to a specific element of the code.

B. Code Rate

The code rate (R) is given as $= \frac{k}{n}$, where k = message length and n = code word length. Or $R = \frac{k}{k+q}$, where q = the number of additional check bits. The notion captures the amount of redundancy used in the code. It is an important parameter for comparing error correcting codes. The most efficient code for a particular application is one where the code rate (R) approaches numeric one [8].

C. Encoder

An encoder is a combinational logic function that converts an input device state into a binary representation of ones or zeros. It converts an incoming data or message into codeword and made ready for processing [9].

D. Decoder

This is the reversed process of encoding data, it tries to recover the original data from the received distorted data. A decoder consists of a syndrome generator, syndrome decoder and error correcting logic. It can handle a set of possible error detection and correction depending on the kind of schemed used.

E. Distance

The distance of a code is a parameter that captures how much two different code words differs. More formally the distance is the minimum number of bit that will flip in order to go from one codeword to another.

F. Hamming Distance

Hamming distance of a code is the minimum distance between two distinct code words, over all pairs of code words. Consider C = {c0, c1, c2, c3}, where c0 = 00000, c1 = 10110, c2 = 01011, c3 =11101. C is a {5, 4} code with distance of 3.

G. Hamming Weight

Hamming weight of a code scheme is the maximum number of ones (1's) among valid code words or the number of 1's in the column or row of the parity-check matrix.

H. Channel

Channel link up the transmitter and the receiving end, the Encoder and the Decoder. The reliability of any communication is possible at any rate below the channel capacity. Error is detected and corrected in the receiving end. To enable the receiver/decoder to detect and correct error, parity bits which is the redundant bits are sent along with the actual message through the channel.

III. FUNDAMENTAL CONCEPT OF CODING

The concept of error correction coding is concerned with the protection of digital information against the error that occurs during data storage and transmission. Many indigenous error correcting techniques based on a vigorous mathematical theory have been developed and have many important and frequent applications. The problem with any high-speed data communication system is how to control the error that occurs during processing. In other to achieve a reliable communication, designers should develop good codes and efficient algorithm [10]. Both Hamming Code and RS Code follow this fundamental concept. In fig. 2, let b = b_1, b_2, to b_n represent a sequence of information symbols, which are to be protected from errors in the system.

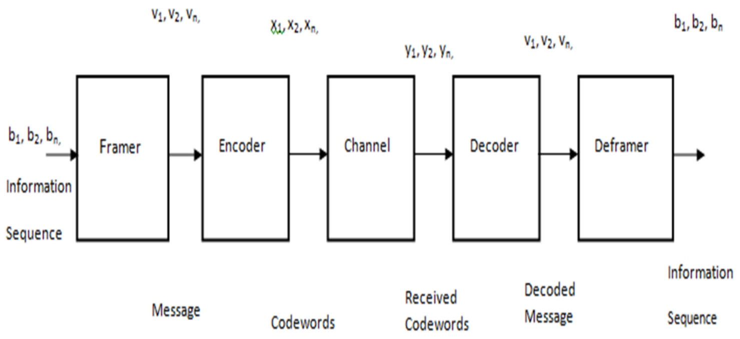

Fig.2a Typical architecture of a system that uses correcting code

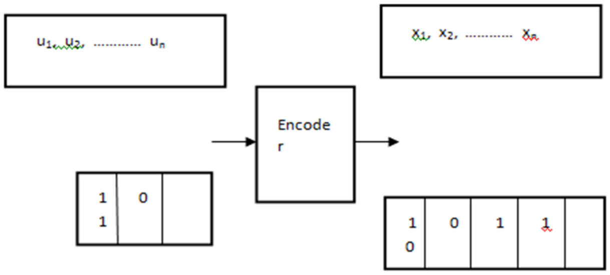

Fig. 2b Encoder overview

At the transmitter the frames convert the information sequence into smaller blocks called message. The messages are then encoded using an error correcting code into codeword which is then transmitted over an unreliable channel. The channel delivers a noisy version of the codeword to the receiver.

The receiver performs the reverse process called decoding, recovering the original message and reconstructing the original information sequences from the decoded messages, this is called deframing. The encoder maps a message of k symbols into a codeword of n symbols in the same alphabet.

In fig. 2b the encoder maps the $k = 3$ bit message into $n = 5$ bit codeword. The decoder performs the exact reverse task of converting the code words back to its original messages.

IV. HAMMING CODE AND REED SOLOMON CODE

A. Hamming Code

Hamming codes was invented by R.W. Hamming during his employment in Bell Labs and is the most widely used linear block codes. The code send message bits padded with specific party-check bits in the form of $(2n - 1, 2n - n - 1)$ where n is the number of the overhead bits, $2n - 1$ the block size and $2n - n - 1$ the number of data bits in the block. All hamming codes are able to detect up three errors but correct only one. They manage this by having the k parity-check bits set at position $1, 2, \ldots, 2^{k-1}$ to check every element whose binary representation has a 1 in position $k_i - 1$. The parity-check bits are in the positions of powers of 2 (positions 1, 2, 4, 8, 16, 32, 64, etc.) whereas the encode data bits are placed in the other remaining positions (positions 3, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 17, etc.). For an example a message of 100110 bits is encoded as k_1, k_2, 1, k_3, 0, 0, 1, k_4, I, 0. With Hamming arithmetic k_1 =1, k_2, = 0, k_3 = 1, k_4 = 1, So that the final sent string would be 1011001110. If error occurs at position 5 during transmission then the received string becomes 1011101110. With the new parity-check bits due to error, k_1= 0, k_2= 0, k_3 = 0, k_4=1. This shows that k_1 of position 1 and k_3 of position 4 parity-check bits differs. Summing up the positions that are wrong, gives us the location of the bad bit (1+4) = 5.

The maximum number of error a code can fix is less than half of the distance between the closet two code words and any more mistakes will be decoded incorrectly. Hamming code have a distinct problem of inefficiency when sending large amount of data, they get increasingly inaccurate as the number of bits increases. They can only correct one flipped bit for each codeword regardless of length. Common Hamming code sizes are {7, 4}, {15, 11}, {31, 26} and they all have the same Hamming distance between codeword. Hamming distance is a parameters used in Hamming calculation. When Hamming distance is known the capability of the code to detect and correct error can be determined. So that with a Hamming distance (d_{min}), Error Detecting Capability equals d_{min} – 1 and similarly the Error Correction Capability equals $\frac{(\mathrm{dmin} - 1)}{2}$. Hamming codes do very poorly against bursts of errors caused by fading over a channel.

B. Reed Solomon Code

The Reed Solomon (RS) code was invented by Irving S. Reed and Gustave Solomon. The RS is organized on the basis of symbols and not bits. They are non-binary, linear and cyclic codes used for burst error correction with high coding rate and low complexity.

The code consist of m bit symbols and are represented as (n, k) where n is the total number of enclosed symbols and k is the number of data symbols. The minimum distance for RS code is given by

$d_{min} = n - k + 1$, where n is the length of the codeword and k is the number of data symbols being encoded. The code can correct up to t symbol errors, where t is given as $t = \frac{n-k}{2}$. This equation shows that a codeword needs $2t$ parity symbols to correct t errors.

To understand RS encoding and decoding, finite fields or Galois Fields (GF) technique are applied. For any prime number p there exists a finite field, *GF*(*pm*), where m is a positive integer.

RS codes are usually written in the following form $R - s(n, k) = (2^m + 2^m - 1, 2^m - 1 - 2t)$, Where k is the number of data symbols being encoded, n is the length of the codeword, t is the maximum error correction capability and the symbols are made up of m bits each. The generator polynomial for Reed Solomon code takes the following form, $g(x) = g_0 + g_1X^1 + g_2X^2 + \cdots + g_{2t-1}X^{2t-1} + X^{2t}$, where the field elements g_i are from Galois field. The degree of the generator polynomial is the number of parity symbols, which is $2t$. Therefore, there are exactly $\alpha = 2t$ roots of the polynomial. By the algebraic rules of the finite field and binary field, we can express $g(X)$ as $g(X) = \alpha^3 + \alpha^1 x + \alpha^0 x^2 + \alpha^3 x^3 + x^4$. Thus, the resulting codeword polynomial $U(X)$ can be written as $U(X) = p(X) + X^{n-k} m(X)$, where $m(X)$ is the message polynomial, multiplied by X^{n-k} and the parity polynomial is $p(X) = X^{n-k} m(X) \bmod g(X)$ [11].

RS codes can be used for long block lengths with less decoding time than other codes because it work with symbol-based arithmetic. This makes RS coding harder to implement, but it provides better throughput. It seems complicated but the fact is that the larger the codeword the more error this code can correct, though require sending much information to do so.

C. Application

Error correcting codes are generally used in communication systems for transmitting data over the telephone lines or via cell phones. They are used in deep space communication and in satellites broadcast. The RS codes are widely used in many applications that involve data transmission, like wireless computer networks, in telephony such as GSM, GPRS, in broadcasting such as DVB-T, DVC-C, in data storage such as hard disk drives (HDD), photo digital storage systems, Memory cards, in optical storage such as Compact Discs (CD), Digital Versatile Discs (DVD), digital audio tape and Blu-ray Discs (BD) are all using Reed-Solomon codes [12]. A Reed Muller Code whose algorithm is gotten from Reed Solomon code where used to detect and correct up to 7 errors in 32-tuple when Mariner space probe transmitted pictures from Mars to earth in 1972. Hamming codes are incorporated in Random Access Memories and in Disk storage to correct error in computers and communication systems. Error correcting codes are particularly suited in noisy transmission lines, and nowadays, all wireless communication uses error correcting scheme.

V. CONCLUSION

The comparative study between Hamming code and RS code shows that RS code is more efficient for data communication because of its low coding complexity and high coding rate. Hamming code is computationally simple and understanding and has the capacity to detect up to three errors but can correct only one error per message. If more than one bit flipped it will incorrectly decode to a different string. But this inefficiency of Hamming code in handling multiple errors can be taken care of by RS code. Hamming codes are relatively efficient when sending small amount of data but they get increasingly inaccurate as the number of bits increases. They can only correctly locate one flipped bits for each codeword regardless of it length. In real situation errors tend to cluster, these cause Hamming code option infeasible in spite it being nice theoretically. Though both Hamming code and RS code are under the category of Bose, Chaudhari and Hocquenghem (BHC) code but Reed Solomon codes sacrifice overall efficiency for broad accuracy.

REFERENCES

[1] Computer-Based Communications Systems and Networks (Spring 2000). Error-Correcting Codes IS250 project 1. Online at: http://people.ischool.berkeley.edu/~rosario/projects/error_correcting_codes.html. Accessed on Nov. 4, 2014.

[2] Robin, Schriebman. Error Correcting Code (April 13, 2006). Online at: http://www.cs.hmc.edu/~mike/public_html/courses/security/s06/projects/robin.pdf. Accessed on Nov. 14, 2014.

[3] Error Correcting Code. Online at: www.cse.buffalo.edu/~atri/papers/coding/thesis-chaps/chap1.pdf. Accessed on Nov. 14, 2014.

[4] Katrepalli, Bhargava. Efficient Implementation of Error Correction Coding in a Communication System by Using VHDL. VSRD-IJEECE, Vol. 2 (6), 2012, 359-365.

[5] S., Baloch, T., Arslan, & A., Stoica (2005). Efficient Error Correcting Codes for On-Chip DRAM Applications for Space Missions. IEEE Conference Publications

[6] Fachpraktikum, Signalverarbeitung (March 9, 2012). SV5: Error Correcting Codes. Online at: http://www.isi.ee.ethz.ch/content/dam/ethz/special-interest/itet/signal-n-information-processing-lab-dam/documents/labs/SV5.pdf. Accessed on Nov. 20, 2014.

[7] W., Cary Huffman, & Vera, Pless (2003). Fundamentals of Error-Correcting Codes. Cambridge University Press

[8] AN1823 APPLICATION NOTE (May 2004). Error Correction Code in NAND Flash Memories. Online at: http://community.qnx.com/sf/docman/do/downloadDocument/projects.filesystems/docman.root/doc1869?logged=1. Accessed on Nov. 4, 2014.

[9] Barry, Paton, (March 1998 Edition). Fundamentals of Digital Electronics (pp. 2 - 3). Dalhousie University.

[10] Muzhi, AL-ANI, & Qeethara, AL-SHAYEA. Unidirectional Error Correcting Codes for Memory Systems: A Comparative Study. International Journal of Computer Science Issues, Vol. 7, Issue 1, No. 3, January 2010.

[11] Gregory, Mitchell. (July 2009). Investigation of Hamming, Reed-Solomon, and Turbo Forward Error Correcting Codes. ARL-TR-4901.

[12] Okeke, C., and Iroegbu, C. (July 2014). Simulink Modelling of Reed-Solomon (RS) Code for Error Detection and Correction. International Journal of Latest Trends in Engineering and Technology (IJLTET). Vol. 4 Issue 2.

Biography

Engr. C. Okeke, B.Eng, M.Eng, MNSE, holds a Bachelors Degree in Electrical/Electronic Engineering, Enugu State University of Science and Technology, Enugu, Nigeria in 1998. A Masters Degree in Electrical/Electronic Engineering, majoring in Electronics and Communication from the same University in 2006. He is a member of Nigerian Society of Engineers (NSE) and Council for the Regulation of Engineering in Nigeria (COREN). He is a member of International Research and Development Institute

(IRDI) and also a member of International Association of Engineers (IAENG). His research area is in Communication Systems and control. Engr. Chukwuma Okeke is currently lecturing in Michael Okpara University of Agriculture, Umuduke, Abia State, Nigeria.