Decimal Arithmetic Manipulation Technique

Data Compression and Security Algorithm

George Yuva Raj[1]
Dept. of Aeronautical Engineering[1]
MLRIT[1], Dundigal[1]
Hyderabad-43[1], Telangana[1], India[1]
georgeyuvaraj2@gmail.com[1]

Ch. Dhanunjay Rao[2]
Dept. of IT[2]
MLRIT[2], Dundigal[2]
Hyderabad-43[2], Telangana[2], India[2]
chigurukota@gmail.com[2]

Goyal Himani Sharma[3]
Dept. of Electronics and Communication Engineering[3]
MLRIT[3], Dundigal[3]
Hyderabad-43[3], Telangana[3], India[3]
goyalhimani@gmail.com[3]

***Abstract*—A method to compress data in digital communication system by means of converting or representing data into decimal form and performing certain arithmetic operations to compress data as well as to incorporate security and to retain transmitted data at receiver by performing counter arithmetic operations to the one by which it is manipulated.**

***Index Terms*— Data, digital communication system, decimal conversion, arithmetic operation, data compression and security**

I. INTRODUCTION

Main objective of this technique is to provide maximum security and maximum compression ratio for a message of an information source with maximum symbols of 2^n while fully utilizing modern processing capabilities. In a digital communication system two types of signals are used to represent two different states that are high and low. So, in this manner, n bits are used to represent 2^n symbols. Same symbols can be represented within the limits of a communication channel by using DAM technique with fewer bits as well as providing security for data transmission. It is a data compression technique. It requires both electronic circuit as well as computer program that run in a computer or an embedded system. Special circuits are necessary to accomplish DAM technique.

II. REQURIMENTS

A computer or an embedded system with suitable latches and registers, circuit to perform PSK (Phase Shift Keying) and a special circuit to generate and identify remainder signals.

III. DESCRIPTION

Consider a source for information with 2^n symbols. These symbols range from 0 to 2^n-1 [2]. Here each symbol is again represented by an n bit binary equivalent. Whenever, Phase Shift Keying is used prior to source coding, all the symbols can be represented from -2^{n-1} to $(2^{(n-1)}-1)$. Therefore, n-1 bits are used to represent each symbol for a source of 2^n symbols. To achieve this, each symbol is converted into its decimal equivalent and a subtraction and division operation is performed [1] in such a way that obtained reminder may vary from 0 to divider-1. All the remainder values must be distinguishable by the receiver and should be able to represent within channels limit. To achieve this, various digital modulation techniques are used. Even M-ary techniques can also be used [2]. Now, this data is further processed for channel encoding and modulation and then transmitted [2]. Prior to transmission other data redundancy can also be used for better efficiency.

When data is received at receiver it is demodulated and subjected to noise reduction techniques. Later it is converted into decimal equivalent and counter arithmetic operations are performed on it, to obtain original message.

IV. METHODOLOGY

COMPRESSING METHOD:-

A.STEP 1

Identify the number of symbols in the source and convert or represent every symbol in their decimal form. If the number of symbols are 2^n. Subtract 2^{n-1} decimal equivalent value from every symbol's respective decimal equivalent from. This obtains data ranging from -2^{n-1} to 2^{n-1}-1 [5]. Where, the positive integers are represented with 1`s and 0`s, negative integers are represented with -1`s and 0`s [4]. Here 1 and –1 are represented with same amplitude and frequency only difference is the change in their phase. Representation of positive and negative values are shown in Fig.1 and Fig.2. 180° phase is maintained to distinguish them apart. Subtracting 2^{n-1} also provides security as the receiver must be aware of total number of symbols in the source. 2^{n-1} has to be added to receive data at the receiver end to retrieve transmitted message [4].A compression ratio of 2:1.75 is achieved at this stage for each symbol.

A low state representation may vary in amplitude, frequency or in pulse with in such way that it should not exceed channel capacity in any manner and should not interfere with 1`s or -1`s.

$P_i=N-2^{n-1}$,Here P_i is any integer ranging from 2^{n-1} to 2^{n-1}-1 and N is any symbol in source in its decimal equivalent form.

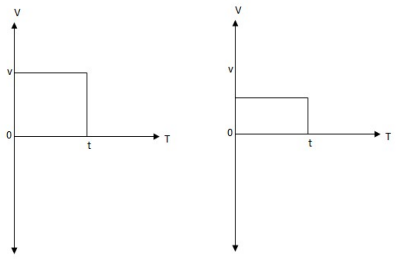


Fig.1. Representation of high state and low states for positive values

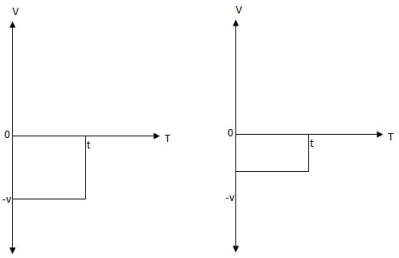


Fig.2. Representation of high state and low states with 180^0 Phase shift for negative values

A low state representation may vary in amplitude, frequency or in pulse with in such way that it should not exceed channel capacity in any manner and should not interfere with 1`s or -1`s.

P_i=N-2^{n-1},Here P_i is any integer ranging from 2^{n-1} to 2^{n-1}-1 and N is any symbol in source in its decimal equivalent form.

B.STEP 2

Now, decimal equivalent values are divided by a divider which is a decimal value [1]. This value is provided by transmitter and provides security to the transmitted data. Since, divider has to be used to retrieve message that is transmitted, receiver must be aware of subtracted value as well as divider value. If P_i is any symbol's decimal equivalent and is divided by x, then quotient and remainder are obtained.

P_i/x where P_i is any decimal equivalent between -2^{n-1} to 2^{n-1}-1 and x is divider. Here, maximum possible value of x depends on the channel capacity and decoding capacity of receiver to identify maximum remainder value that is generated during division. Since, divider is x remainder values may vary from 0 to x-1. So a special circuit to generate x unique remainder signals which are distinguishable and are intended to represent only remainder values. All the signals must have frequency and amplitude in such a way that they do not exceed the limitations of the carrier channel at the same time receiver should be able to identify all remainder signals [7][2]. Based on the number of pulses allocated for remainder signals x value quotient and remainders may vary. If x value increases quotient need fewer bits and remainder signals increases [1].

Greater the x value is higher the security and compression ratio. If the x value is n+1 then compression ratio of 2:1 and security of higher standards is accomplished provided that channel can carry all remainder signals without any loss. Once, division operation is done remainders are represented with separate pulses and quotient is represented with binary digits with phase shift depending upon whether P_i is positive or negative[2].

C.STEP 3

Data redundancy techniques can also be incorporated if necessary that is rejecting data which seem to be unnecessary. Finally, data is channel modulated and transmitted in such way that remainder signals act as marking. First, a remainder signal is transmitted followed by its quotient in binary system. Therefore, every symbol starts with a remainder signal and ends with quotient or vice versa [6]. This technique can be adapted for wireless and wired communication in the form of packets and data stream. It is suitable for both parallel and serial data transmission system. If x is the divider then quotients of symbols ranging from sx to 2sx-1are represented by using s bits. Therefore, sx-1 symbols ranging from sx to 2sx-1 are represented by s bits by ignoring all zeros on the left side before the 1^{st} high value from left side including remainder signals totally s+1 bits or pulses are needed to represent each symbol on the channel.

If data is processed through 1^{st} 2 stages successfully with a x value of 2, then a minimum of 4:3 compression ratio and maximum 2:1 or even higher can be achieved based on the efficiency of the channel and x value.

While representing various factors must be considered to avoid ISI (Inter Signal Interference) and losses due to noise at receiver. Sensitivity of channel and receiver is very important aspect to be considered. A system with higher sensitivity can be used to represent maximum number of remainder signals such that fewer digital pulses are required to transmit data.

Other digital modulation techniques may also be used depending upon the divider and other factors [4]. Huffman coding can yield better results by allocating fewer bits to higher probable symbols and higher symbols to lesser probable symbols[8].Fig.3 shows complete flow chart for compression and encryption.

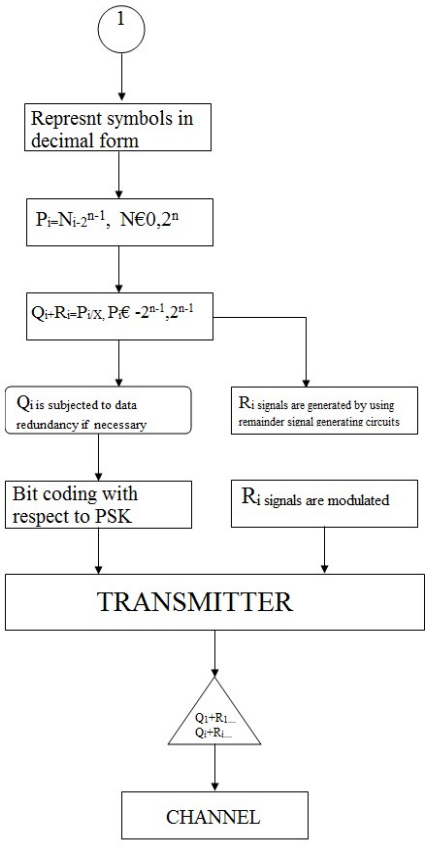


Fig.3. Algorithm to compress and encrypt data

DECOMPRESSING METHOD:-

At receiver decompression is performed. This is exact counter arithmetic operations for compression with same numerical values in reverse order[2].

A.STEP 1

Whenever data is received is subjected to a band pass filter[2]. In other words channel demodulation is performed to suppress high frequency carrier signal.

B.STEP 2

Data is feed to a remainder identifying circuit and separates quotient and remainder[5][7]. Separated quotient is again feed to a micro controller which separates positive and negative values[7]. At the same time remainder identifying circuit identifies corresponding remainder signal and stores it[7].

Now receiver should be aware of x value that is divider. When divider is supplied to system it obtains dividend by.Now, dividend is further processed by multiplying Positive integers with 1 and negative integers by -1[5]. Identification of positive and negative integers is done by identification in phase shift of non-remainder signals that is quotient bits. If quotient bits are negative then their corresponding remainder value is also considered as negative integer. These quotient and remainder yields dividend when divider value is provided to the system.

Quotient(divider)+ Remainder=Dividend. This yields P_i ranging from -2^{n-1} to $2^{n-1}-1$.

C.STEP 3

Once P_i is obtained it is again allowed to addition operation to obtain original symbol's decimal equivalent. 2^{n-1} is added to every integer thus obtained[5].

$N=P_i+2^{n-1}$, where n ranges from 0 to 2^n-1.

All decimal values are converted in binary digits and processing it yields original message.

In Fig.4 working of a remainder identifying circuit is mentioned by using a block diagram.

Fig.5 shows complete flow of decompression and decryption of data using DAM process.

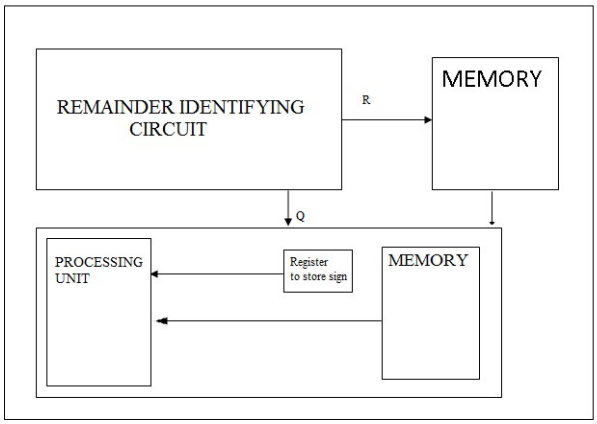


Fig.4. Block diagram of demodulator circuit

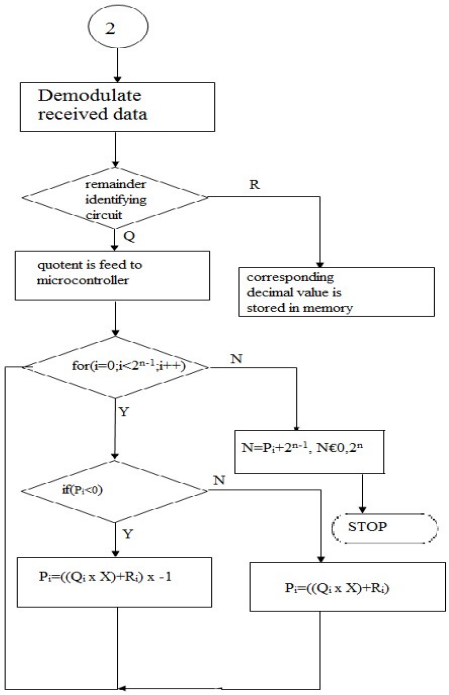


Fig.5. Algorithm to decompress and decrypt data

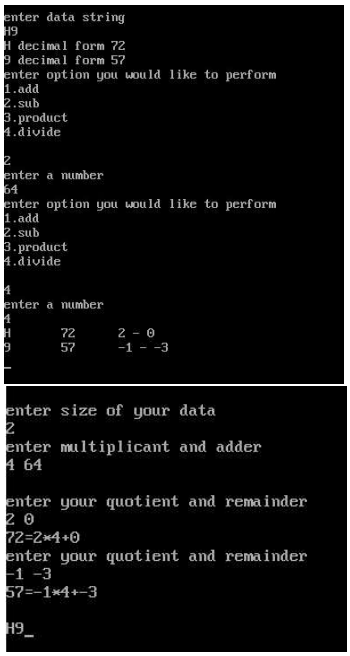
IV. ILLUSTRATIVE EXAMPLE

Consider same algorithm is used to compress data which is coded in ASCII. So, it has symbols with decimal equivalent values varying from range 0 to 127.

When each symbol is subtracted with 64 or 63 the resulting decimal equivalent ranges from -64 to 63 or -63 to 64 respectively[5]. This has to be done prior to source encoding

Now, a message has been generated using symbols from this source and that message is represented in modified decimal equivalent and then allowed to perform an arithmetic operation such a division[1]. If transmitter uses any remainder identifying circuit then remainder value may depend upon the efficiency of circuit or else maximum possible divider without using remainder identifying circuit is 2. When divider is the 2, then number of bits needed to represent 2^n symbols are n-2 bits.

After arithmetic coding that is suppressing data which seems to insignificant[3]. Here, zeros on left side of a byte are suppressed as they don't carry any significant data. Total number of bits needed to represent all 128 symbols by general representation is 1024 bits[3]. Total number of bits used to represent 128 bits by using DAM technique is fewer bits along with their remainder values that is as follows. Since, we are using 2 as divider possible remainder values are 1 and 0.So there is no need of complex remainder identifying circuit. A circuit which can identify two different signals and change in their phase can resolve the task.



0 to 3 needs 2 bits
4 to 7 needs 3 bits
8 to 15 needs 4 bits
16 to 31 needs 5 bits
32 to 63 needs 6bits

By using this techniques all symbols together that is ranging from -32 to 32 need only 258 bits where as by general method 1024 bytes must be used to transmit the data.

Further more, Huffman coding can yield better results by allocating fewer bits to higher probable symbols and higher symbols to lesser probable symbols.

V. CONCLUSION

This algorithm provides maximum security, compression and efficiency by using all the available resources and techniques available in modern digital communication system. Hence, it is reliable and cheaper technique to adopt for a digital communication system. A compression ratio of 2:1 or even higher can also be achieved. It is a patch that is either a software or hardware which will make it suitable for any communication system. So, it is adaptable for any communication system. By implementing this technique huge data is transmitted with higher security and better quality. A better grade of service can be achieved in as telecommunication switching system. Availability of network is increased to greater extent. Probability to lose data is greatly reduced. Furthermore exploration can be done on fields of digital signal processing and free space optical

communications by using 3 signals that is using 1,0 and -1 by means of RGB colours.

References

[1] Ronald Fink, US 6373986 B1, 2002. [Compression of data transmission by use of prime exponents.

[2] Simon Haykin, Digital Communications. 4Th edition, John Wiley & Sons Inc., 2009.

[3] Rissanen, J.J.; Langdon, G.G., Jr. "Arithmetic coding". IBM Journal of Research and Development 23 (2): 149–162. doi:10.1147/rd.232.0149. Archived from the original on 28 September 2007. Retrieved 2007-09-22,March 1979.

[4] "Modulation," http://en.wikipedia.org/wiki/Modulation, 25 June, 2014.

[5] Osseily, H.A. and Haidar, A.M., "Arithmetic Algorithms and Circuits to Convert MVL to MVL Coded Decimal and Vice Versa,"Information and Communication Technologies: From Theory to Applications, 2008. ICTTA 2008. 3rd International Conference on Damascus, IEEE, INSPEC Accession Number:10053411,April, 2008.

[6] Ramesh R. Sarukkai, "Prime numbers and output codes",Neural Networks, 1995. Proceedings., IEEE International Conference on Perth, WA, Page(s):564 – 568 vol.1,INSPEC Accession Number:5255858, 1995.

[7] Tamir, "Discrete complex fuzzy logic", Fuzzy Information Processing Society (NAFIPS), 2012 Annual Meeting of the North American on Berkeley, CA, INSPEC Accession Number:12965226,6-8 Aug, 2012

[8] Huffman. D."A Method for the Construction of Minimum-Redundancy Codes". Proceedings of the IRE 40 (9): 1098–1101. doi:10.1109/JRPROC.1952.273898, Sept,1952.