DISTRIBUTED SOURCE CODING WITH LDPC CODES:

ALGORITHMS AND APPLICATIONS

Himani Gandhi

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

December 2022

APPROVED:

Hua Sun, Major Professor
Kamesh Namuduri, Committee Member
Guturu Parthasarathy, Committee Member
Xinrong Li, Interim Chair of the Department of Electrical Engineering
Shengli Fu, Interim Dean of the College of Engineering
Victor Prybutok, Dean of the Toulouse Graduate School

Gandhi, Himani. *Distributed Source Coding with LDPC Codes: Algorithms and Applications*. Master of Science (Electrical Engineering), December 2022, 39 pp., 21 figures, 11 numbered references.

The syndrome source coding for lossless data compression with side information based on fixed-length linear block codes is the main emphasis of this work. We demonstrate that the source entropy rate can be achieved for syndrome source coding with side information when the sources are correlated. Next, we examine employing LDPC codes to apply the channel and syndrome concepts in order to satisfy the Slepian Wolf limit. Our findings indicate that irregular codes perform significantly better when the compression ratio is larger. Additionally, we looked at how well different applications performed when running on two different mobile networks. We have tested those applications which are used in our day-to-day life. Our main focus is to make wireless communication much easier. We know that nowadays data is increasing which led to increase in the transfer of data. There are a lot of errors while doing so like channel error, bit error rate, jitter, etc. To overcome such kind of problems compression and decompression should be done effectively without any complexity to achieve a high performance ratio.

Copyright 2022

by

Himani Gandhi

ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to Dr. Hua Sun, my major advisor and mentor, for his unwavering support and guidance. His insights guided me through my academic career at UNT. I was able to push myself to reach goals with the help of Dr. Hua Sun. Hopefully, I would also be able to make him proud. Dr. Kamesh Namuduri, and Dr. Guturu Parthasarathy have been invaluable in directing me through my studies and providing useful perspectives. I am grateful to the Department of Electrical Engineering for awarding me a Teaching Assistantship. Finally, I would like to express my gratitude to Teleworld Solutions who helped me gain insight of practical applications, and all of my friends who kept my spirits high throughout my time at UNT.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1. Motivation

In this study, it is demonstrated how the Slepian Wolf theorem for correlated binary sources can be achieved by low-density parity check (LDPC) codes. We concentrate on the asymmetric case of side information compression. The strategy is built on using the syndrome notion and interpreting the relationship as a channel. The processes for encoding and decoding, also known as decompression and compression are discussed in depth. The simulated performance results are quite near to the Slepian Wolf limit and outperform the majority of the current turbo-code performances reported in the literature. Furthermore in the paper we discussed the applications of LDPC code. Taking in account different applications in real time communication system, I have analyzed the performance of two networks which was possible because of low complexity encoders and decoders.

1.2. Compression of Correlated Binary Sources

Input alphabet has the structure of a field (for instance, the binary alphabet gives the field F_2). The idea of linear codes with block lengths of n and dimension k that are specified as sub-spaces of a particular sequence of bits. Typically, a messaging is described as a trio consisting of an entry alphabet, an exit alphanumeric characters, and a probability density p for each pair of inlet and outlet items (i, o). The transition probability, in terms of meaning, is the possibility that the symbol o will be received provided that the symbol i was communicated over the channel. Shannon demonstrated the existence of a quantity, referred to as the channel capacity, so that effective communication is achievable for rate arbitrarily near to the capacity and is not feasible for rates over the capacity for a given communication channel. As a result, it cannot assure that there will be transmission plans that can handle the capacity. So Shannon introduced the concept of codes, It is obvious that reliable communication across a channel is impossible if only single elements are transferred through it if even one input element can be processed in at least two conceivable ways (although with different consequences). Even if several unrelated parts are sent, this is the case (in a manner to be made precise). Therefore, it is crucial to communicate input variables that are connected in order to accomplish trustworthy communication. As a result, the idea of a code which is a (limited) sequence of vectors over the source alphabet—is introduced. Assuming that each vector has the same length, we refer to this length as the code block length. Each vector can be specified with k bits if there are $K = 2^k$ vectors. If the vectors are n length, then k bits have been transferred through the channel in n times of use. In that case, we say that the code uses k/n bits per channel. Let's say we submit a codeword and get a vector of the output alphabet in return. How do we determine the sent vector? There is no universal method of knowing which codeword had been sent with 100 percent certainty if the channel permits for errors. The likelihood that this codeword has been sent considering the reported vector is maximized, so we can identify the codeword that was transmitted most likely. List all K codewords and their conditional probabilities one by one to demonstrate that we can indeed find such a codeword. Once you have identified the vector or vectors that produce the greatest probability, return one of them. The maximum likelihood decoder is the name of this decoder. Although it may take a long time and occasionally make mistakes (specifically whenever the code is huge), it is the utmost we can manage. Shannon demonstrated that there are codes with rates arbitrarily near to capacity, for which the greatest likelihood decoder's risk of error decreases as the code's block length increases to infinity. (In reality, Shannon demonstrated that the maximum likelihood decoder's decoding error decreases exponentially with block length, but we won't get into that here.) From the perspective of communication, codes that reach capacity are very good, but Shannon's theories are non-constructive and don't offer any guidance on where to look for such codes. More critically, it is unclear how to efficiently encode and decode them, even if an oracle provided us with code sequences that can operate at a given pace. This note's main focus is on the implementation of codes with effective encoding and decoding algorithms that come close to the channel's capacity. Let us illustrate this with Binary Symmetric Channel. The input and output alphabets for the BSC are both F_2. With probability $1 - p$, a bit is either transferred correctly, or with probability p, it is reversed. At first glance, this channel could appear to be easier than the BEC, but it is actually far more intricate. It becomes complicated since it is unclear which bits are switched. (In the BEC's instance, it is obvious which bits are deleted.) This channel has a capacity of $1 + p\log_2(p) + (1-p)\log_2(1-p)$. Finding the codeword with the minimum Hamming distance from the received word for a given vector of length n over F_2 is identical to maximum probability decoding for this channel. Researchers soon discovered that random codes are capable of obtaining capacity after Shannon's discovery. In fact, Shannon's work itself makes this implicit.

However, reaching capacity is only one aspect of the story. It is necessary to have quick encoding and decoding algorithms if these codes are to be utilized for communication. Just 2 random vectors of length n across the input alphabet make up random codes with a rate of R bpc. To be able to encode information into these vectors, we either need to provide some description of them or compile them all into a "codebook" that specifies which Rn-bit sequence corresponds to which codeword. This requires a codebook of size 2 random vectors, which is too enormous for any code of a manageable size (for example, a code of length 1000 and rate 0.5, which produces 2500 vectors).

One can perform better, at least in terms of encoding, if the the vector space Fn^2 was independently developed by Elias and Golay. Since they are linear and have a rate of k/n, These codes can be defined in terms of a basis made up of k vectors of length n. By mapping a bit vector $(x_1, ..., x_k)$ to the vector created by generating linear combinations of the basis vectors provided by the coefficients $x_1, ..., x_k$, a codebook can now be implicitly stated in a natural way. To demonstrate that there are linear code sequences with rates arbitrarily near to capacity and for which the error probability of the maximum likelihood decoder reaches zero (exponentially rapidly) as the block length grows to infinity, Shannon's arguments can be utilized (nearly verbatim). Additionally, it may be demonstrated that random linear codes reach capacity. Linear codes can be encoded in polynomial time as opposed to exponential time, unlike their non-linear cousins[**8**]. The decoding issue appears to be much more difficult. The maximum likelihood problem on the BSC has indeed been demonstrated to be NP-hard for various kinds of linear codes, as was previously indicated (e.g., general linear codes over F_q for any q). Therefore, polynomial time techniques for maximum likelihood decoding of generalized linear codes are unlikely to exist. Try to replicate the encoding problem's success story and focus on sub-classes of generic linear codes as one strategy to get past this unfavorable consequence. We haven't been able to identify any linear code sub-classes that have polynomial-time maximum likelihood decoding and reach capacity though. As shown in Fig: 1.1 we are trying to compress two sources X and Y which are correlated and they do not communicate their outputs with each other. For example in the below case there are two users trying to play a video game using two different mobile phones there outputs are measured together at one common computer screen. For the compression of such kind of sources we use the concept proved by Slepian and Wolf. They proved that two sources which are correlated can have same joint entropy as two joint sources. In information theory the aim is to have close to zero probability at decoder. By using error-correcting codes we can fulfill this, LDPC is one such kind of error correcting code which we are going to use for compression and decompression. LDPC codes use sparse matrix for determining syndrome. Sparse matrix means number of ones are less compared to number of zeroes. This parity check matrix is used for further syndrome concept and its graphical representation is in form of tanner graph which is shown in Fig: 1.2. By connecting the left variable nodes with right check nodes messages are sent. This reduces the complexity of decoder as well [**6**].

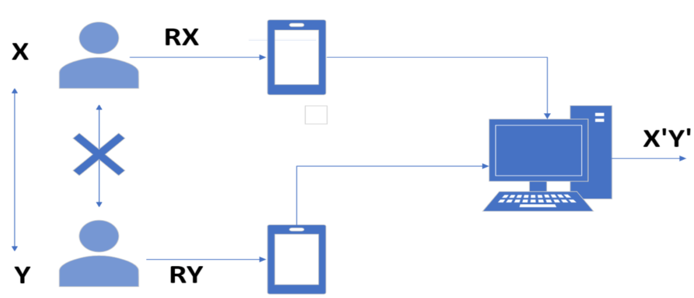


FIGURE 1.1. Compression of binary sources with common decoder

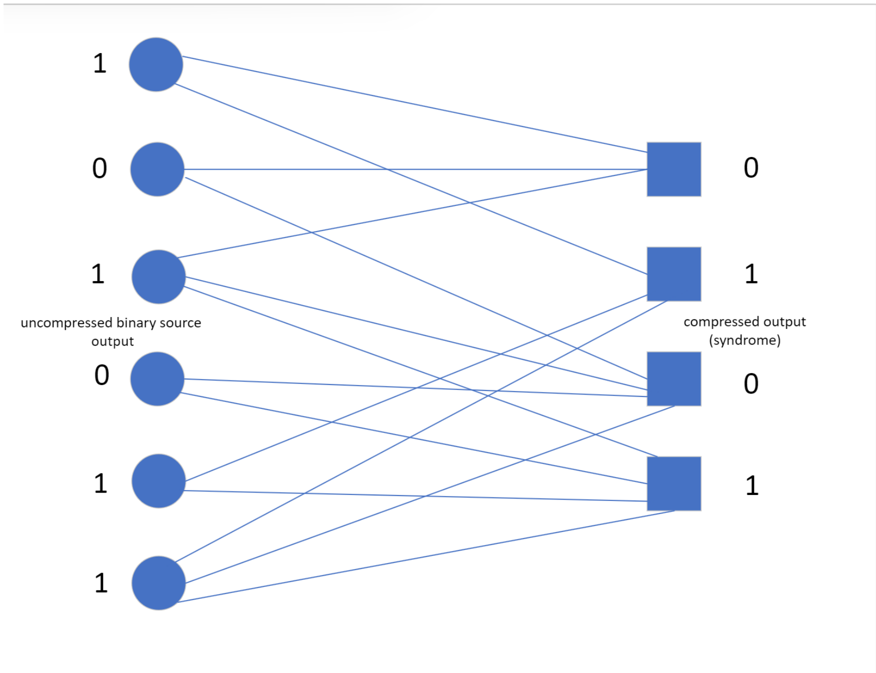


FIGURE 1.2. Tanner graph

1.2.1. Introduction to LDPC Codes

In his PhD thesis, Robert Gallager devised the LDPC codes. They were constructed, essentially disregarded for the next 30 years, and then continuously recreated. One of the most fascinating parts of their history is that they have returned after two independent cultures reinvented codes that are similar to Gallager's LDPC codes almost at the same time but for very different reasons. Bipartite sparse graphs are the source of the linear codes known as LDPCs. Assume that G is a graph with n message nodes on the left and r right nodes (called check nodes). The graph yields the following linear code, which has blocks of length n and at least nr dimensions: The n coordinates of the codewords are connected to the n message nodes. The codewords are all those vectors $(c_1, ..., c_n)$ where the sum of the surrounding locations among the message nodes is zero for all check nodes. By examining the adjacency matrix of the graph, it can be seen that the graph representation is comparable to a matrix representation as shown in Fig: 1.3, let H be a binary r n-matrix in which entry (i, j) is 1 if and only if the ith check node is linked to the jth message node in the graph. The collection of vectors $c = (c_1, ..., c_n)$ such that $Hc = 0$ is the LDPC code provided by the graph. For the code, the matrix H is referred to as a parity check matrix as shown in Fig: 1.3. In contrast, any binary r n-matrix generates a bipartite graph with n message and r check nodes. This graph is related with the code that is specified as the null space of H. Therefore, a bipartite network can be represented by any linear code as a code (note that this graph is not uniquely defined by the code). Not all binary linear codes, however, may be represented by a sparse bipartite network. In that case, the code i referred to as an LDPC (low-density parity-check) code. The sparsity of the graph representation is a crucial element in LDPC codes' ability to have efficient algorithms. [**8**].

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}_{4X6}$$

FIGURE 1.3. Matrix for tanner graph

A branch of probability theory called information theory discusses mistake probability. When the chance of an error at the receiver approaches 0, an error-free transmission is feasible. Ideas, limits, and regulations governing the transit of communications across communications networks are computed numerically. Entropy, which is connected to the idea of information in the information theory and is represented by the letter H, is the average amount of information in any given source of information. The main challenge in a communication system is error-free message transmission [**10**]. The majority of faults are caused by noise or channels. Techniques for error identification and repair are employed to address this problem. Encoding is the process of changing a message into binary, or 1 or 0, form. Extraction of the original message is known as decoding. One such error control coding method used for error detection and correlation is linear block codes. For each information word, information data is divided into a block of length K pieces. Then, each information word is encoded into a codeword, which is a block of bits with a length of n. Every coset of data that is sent to the decoder has the syndrome ascribed to it. The decoder receives the syndrome former and any side information that is readily available to it without any loss [**6**].

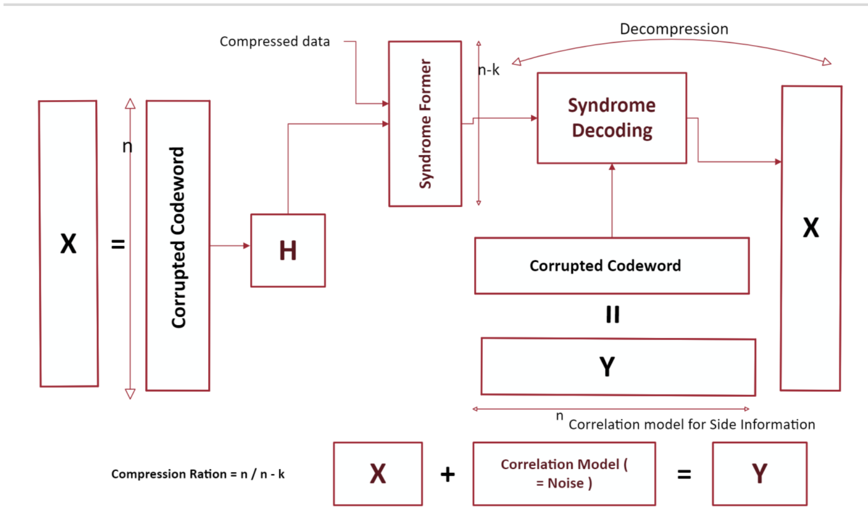


FIGURE 1.4. Overview of the model

1.3. Thesis Organization

This thesis is organized as follows.Chapter 1 explains how the compression an decompression is done using LDPC codes.

Chapter 2 explains in detail the correlation of channel usage along with syndrome concept of two correlated sources out of which one is available losslessly at the decoder.

Chapter 3 describes the algorithm of the LDPC codes and how they can be used in the new generation communications system.

Chapter 4 depicts the results of a mobile phone using two networks like their throughputs, RSRP, Time on Network, Coverage etc.

In Chapter 5, the conclusion and the future scope are addressed.

CHAPTER 2

PROBLEM SETUP AND CORRELATION OF CHANNEL USAGE ALONG WITH SYNDROME CONCEPT

2.1. Compression of Two Sources With Side Information Provided at the Decoder

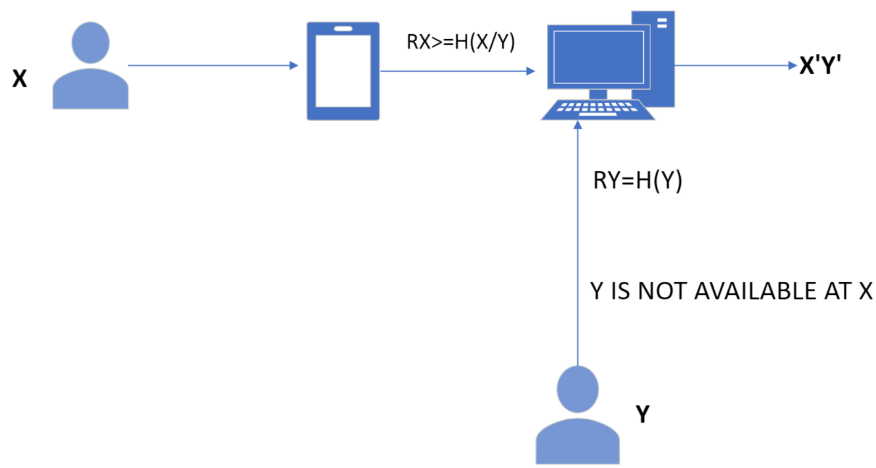


FIGURE 2.1. System for compression with side information.

We take into account the system in Fig: 2.1 under the following presumptions. Since X and Y they are memoryless discrete random variables, knowing one of their past values does not reveal anything about their present value. They are related at the same instant [**3**]. The joint decoder can access Y losslessly, and we make an effort to compress X as well as we can. The theoretical limit for lossless compression of X is derived from the Slepian-Wolf theorem $R_1 \geq H(X|Y)$, since the rate utilized for is its entropy. $R_2 = H(Y)$. Imposing the above three assumptions to the system of Fig: 2.1, we end up with a compression with side information problem. There is an equivalent way one can view the system of Fig: 2.1 in order to allow the use of channel codes. The relation between X and Y can be represented with a channel; X will be the input to the channel and Y its distorted output. Then the compressed version of X, i.e. Z_1 , can be used to make X look like a codeword of a channel.

All of the common X sequences are divided up into disjoint channel codes with the same characteristics, all of which perform equally well across the correlation channel. Z_1 is the codebook's index to which X belongs. Although this method served as the foundation for proving the Slepian Wolf theorem, it has not yet been made explicit in the material that is now available for the construction of effective code. The disjoint channel codes can be created using syndromes if the additional presumption that all values of X and Y are equiprobable is added. With the binary example, when the presumptions are changed as follows, this is relatively simple to accomplish. X and Y are memoryless binary random variables that are equiprobable.

The same instant X and Y are correlated with $\Pr[X \neq Y] = P < 0.5$. This will be called binary symmetric channel correlation from now on. The joint decoder has lossless access to Y, therefore we attempt to compress X as effectively as possible. Since the rate used for Y is its entropy $R_2 = H(Y) = 1$ bit, the theoretical limit for lossless compression of X is from the Slepian Wolf Theorem $R_1 \geq H(X|Y) = H(p) = -p\log_2 p - (1-p)\log_2(1-p)$.

These three presumptions will serve as the foundation for the remainder of the article.

There are $2^n - k$ unique syndromes, each of which indexes a set of $2k$ binary words of length n using a linear (n, k) binary block code. The original code is what we call the linear block code (all-zeros syndrome set). All of the $2^n - k$ sets are disjoint because any two binary words in the same set (corresponding to the same diagnosis) add up to a codeword of the original code, whereas any two binary words in separate sets (corresponding to different syndromes) do not add up to a codeword of the original code.

The Hamming distance characteristics of the original code are also preserved in each set, i.e., all codes operate equally well over the binary symmetric correlation channel. As a result, this scheme's compression ratio is $n/(n-k)$. We will now go over how low-density parity-check (LDPC) codes can be used to compress and decompress with side information because they have been shown to be very effective linear binary block codes.

2.2. Encoding and Decoding With LDPC Codes

The bipartite graph and parity-check matrix H of low-density parity-check (LDPC) codes serve as the finest descriptions of these codes. A binary LDPC code's $(n-k) \times n$ parity-check matrix is similar to that of a regular binary block code, except it is sparse, meaning it contains few ones [**6**]. The degree distribution polynomials tilde $\lambda(x)$ and $\rho(x)$ which represent the proportion of H's columns and rows with various Hamming weights, respectively, describe how these ones are distributed in (number of ones).

For instance tilde $\lambda(x) = 0.25 \times 2 + 0.75 \times 3$ means that 25% of all the columns in H have Hamming weight equal to three and the rest 75% have hamming weight four. When both tilde $\lambda(x)$ and $\rho(x)$ have only a single term, the LDPC code is regular. Otherwise it is irregular. In general, a regular LDPC code with the same codeword length and code rate is expected to be less powerful than an optimized irregular one. The coding rate is precisely determined given both tilde $\lambda(x)$ and $\rho(x)$, however multiple H's can be generated. Typically, one is chosen at random and occasionally it satisfies specific restrictions that are easier to relate to the bipartite graph [**3**]. The parity-check matrix H is equivalently represented by the bipartite graph of an LDPC code. Each row is represented by a check or right node, and each column by a variable or left node. Wherever there is a one in H, an edge connects the relevant row and column. All variable nodes (circles in Fig: 1.2 are placed in one column, all check nodes (squares in Fig: 1.2 in a parallel column. The tilde is used in numbers tilde $\lambda(x)$ and $\rho(x)$ because they refer to the node profile, or the percentage of nodes with different degrees, respectively. The edge profiles $\lambda(x)$ and $\rho(x)$, which show how many edges in the bipartite network are related to different degree nodes, are more commonly utilized in the LDPC literature. The message-passing algorithm can be applied since the bipartite graph is used in the decoding process.

2.2.1. Encoding

For our system, the encoding process is rather easy to understand. The degree distribution $\lambda(x)$ and $\rho(x)$, are initially determined when fixing the codeword length n. Next, a random realization of H is generated. An arbitrary binary input sequence can be compressed, or encoded, by forming a vector out of n successive bits, multiplying them by H, and identifying the matching syndrome.

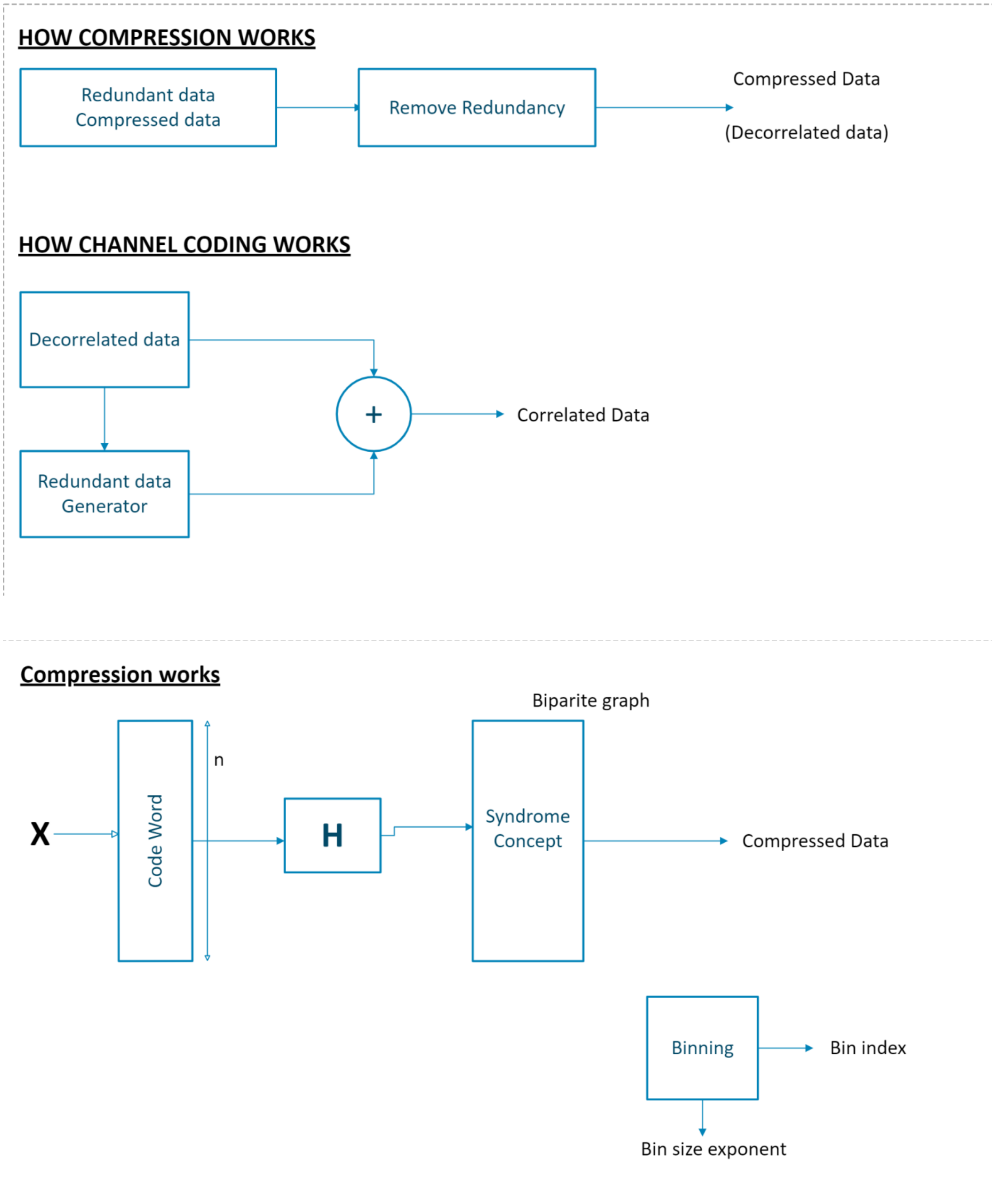


FIGURE 2.2. Working of encoders.

As shown in Fig: 1.2, this can be conceptualized as the binary addition of all the left node values related to the same right node in the bipartite graph. Fig: 1.2's left side displays

6 sequential binary output values of source X from 2.1, whereas Fig: 1.2 right side displays the equivalent 4 output values Z_1 of the first encoder from Fig: 2.1. Since H is a sparse matrix, the difficulty of this encoding process is linear with the length of the codeword. For instance, the code in Fig: 1.2 achieves 3:2 compressions while maintaining the distance attributes of the original LDPC code for all inputs.

2.2.2. Decoding

From its $(n - k)$ long syndrome, the decoder must extract an n-length series of successive X's and the accompanying n-length sequence of successive Y's. It would be equivalent to transmitting X across a binary symmetric channel with crossover probability p given the Y's alone. It appears that these n create successive X's an LDPC-like codeword with the syndrome. For instance, the n X's represent an LDPC codeword if the syndrome is an all-zero vector, and they can be decoded using the message-passing method and the bipartite graph. It is pretty easy to think of the change required in the case of a nonzero syndrome. All of the nonzero points in the syndrome vector are noted in the bipartite graph, i.e., the appropriate check nodes, before the decoding procedure for each binary word of length n begins. The message-passing decoding process is once more applied, but whenever we must operate over a marked check node (using the"tanh rule"), an additional change of sign is made at each output value. The ratio of probabilities must be inverted in the log-domain in order to satisfy the inverse parity, which requires an additional change in sign at the check node's output. LDPC codes can be decoded in two different methods. The decoder performs all parity checks in accordance with the parity equations in the first method. Any bit whose value appears in more unsatisfied parity equations than a predetermined number has its value reversed. Parity equations are recalculated using the new values after the new values have been obtained. When parity-checks are small, this type of decoding technique is feasible. The LDPC graph is subjected to probabilistic algorithms in the second way. The graph is a sparse bipartite graph with two sets of nodes, one set of which stands for the parity equations and the other set for the code symbols. If the code symbol is given in the equation, a line connects the node in the first set to the second. Passing messages along the graph's line allows for decoding. The parity values of the messages are determined when they are passed from message nodes to check nodes and back again. Despite being difficult, this decoding produces better results than the previous.

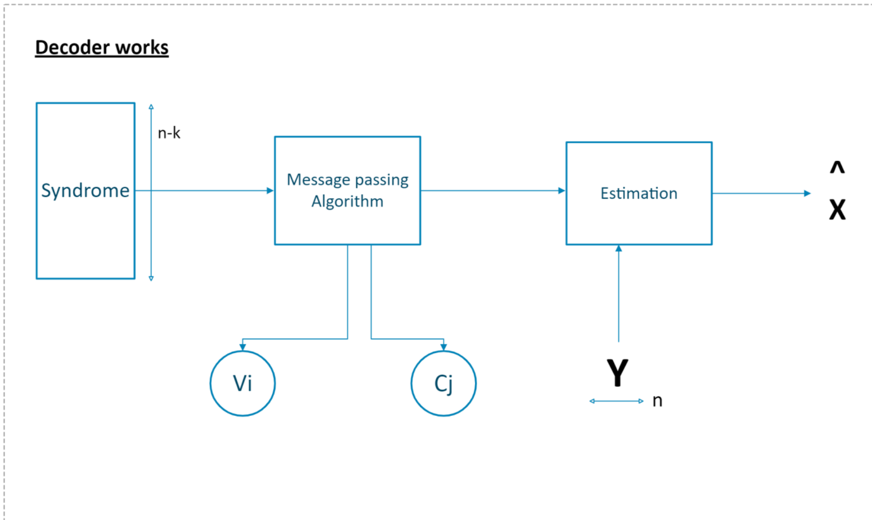


FIGURE 2.3. Working of decoder.

2.2.3. Regular and Irregular Codes

We simulated left standard LDPC codes and compared the outcomes to those of turbo-codes. The length of the codeword is, which is the same as the length of the interleaver. After 40 cycles of the message passing mechanism, more than 2000 blocks were delivered without a single error. The LDPC code distribution includes and, where the variation affects the code rate. The graph shows that even very basic LDPC codes perform slightly better than turbo codes. Simulated with the standard $(3, 6)$ code was this irregular code. The codeword (frame) size was the same for both the regular and irregular codes [**3**]. After the decoding method has completed 100 iterations, the bit error rate (BER) for was calculated. Fig: 2.5 also displays the best turbo code performance known for this code rate as well as the Slepian Wolf theoretical limit of 0.5 bits. The rate 1/2 LDPC code presented in that has been tailored for the binary symmetric channel is the second irregular code, denoted by the letter "bsc" in Fig: 2.5 (BSC). As seen in Fig: 2.5, the threshold for this code is bits. Only good irregular codes could marginally outperform regular codes for compression ratios of 4:1

and 8:1, i.e., only good irregular codes could perform better than regular codes for higher compression. Our findings for length 10 irregular LDPC codes are shown in Fig: 2.4. It is shown that (BSC) irregular codes perform much more better compared to regular codes.

R_1	0.125	0.250	0.500
$H(X\|Y)$ (irreg. LDPC)	0.091	0.204	0.466

FIGURE 2.4. Table 1 [**3**].

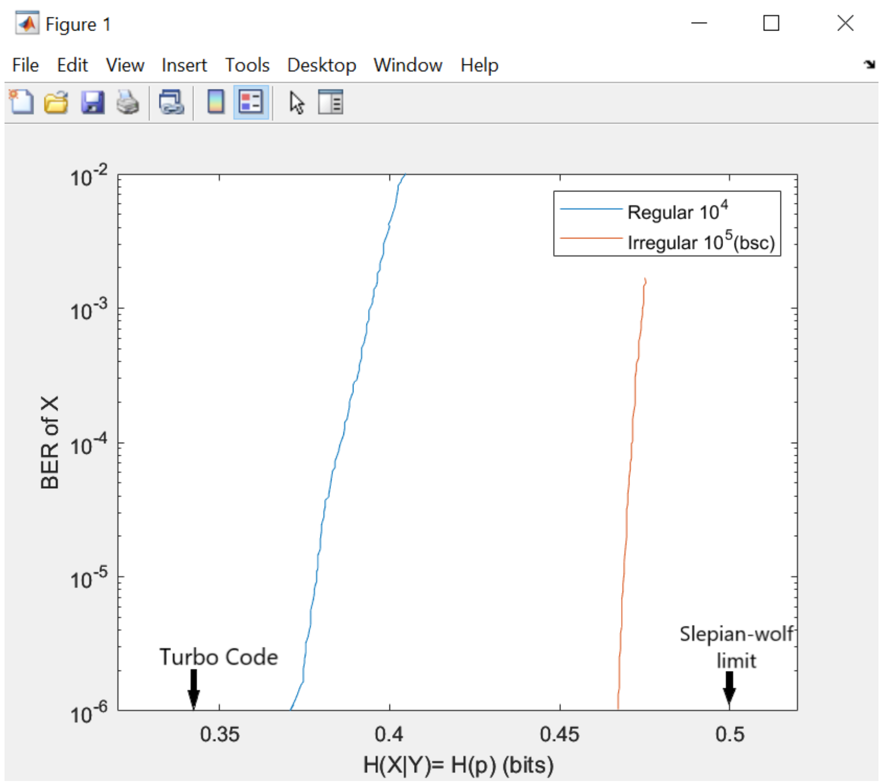


FIGURE 2.5. Simulation Result

2.3. Algorithm of LDPC Codes

Network communications are the foundation of the modern world's expanding information transmission. Few methods are created for network-based data transmission, despite the fact that several cutting-edge technologies have been established to actualize comfortable communication. Even the most modern information technologies still use point-to-point protocols, which is unusual given that the "Internet," a worldwide computer network, already exists [**4**].

Thus, it is now that we should concentrate on multi terminal communication methods. Source coding, often known as data compression, is a method for reducing the size of messages and information representation. Shannon demonstrated that if the code rate $R = M/N$ satisfies $R = H_2(S)$ in the limit N, M , one can use a different representation for an information source represented by a distribution $P(S)$ of an N-dimensional Boolean (binary) vector S in which the message length N is reduced to $M(N)$ without any distortion. Regrettably, Shannon's theorem is not constructive and does not outline specific guidelines for creating the best codes. The fact that a realistic code developed by Lempel and Ziv (LZ) in 1973 reaches the Shannon's optimal compression limit for point-to-point communication is therefore unexpected. However, it should be noted that although the importance of network communications (NC) is growing quickly due to recent developments in the Internet, generalizing LZ codes to sophisticated data compression suitable for NC is challenging. This is primarily due to all the real-world codes that reach Shannon's limit [**5**]. While compression should often be performed independently on each terminal, date require comprehensive information of all source vectors entering the communication network. Therefore, one of the most crucial issues in information theory is the search for more effective compression codes that are appropriate for NC.

The idea is to use current advances in error-correcting code (ECC) research to achieve this. More particularly, future research on the effectiveness and limitations of a linear compression method inspired by Gallager's ECC, which is actively being investigated in the communities of information theory and physics, is required [**10**].

The random binning argument used in the demonstration of the Slepian Wolf theorem is asymptotic and non-constructive, just like in Shannon's channel coding theorem. We can initially try to develop codes that reach the blue corner point A in the Slepian Wolf rate zone of Fig: 2.6 with $R_1 + R_2 = H(X|Y) + H(Y) = H(X, Y)$., as shown in Fig: 2.6 Rate

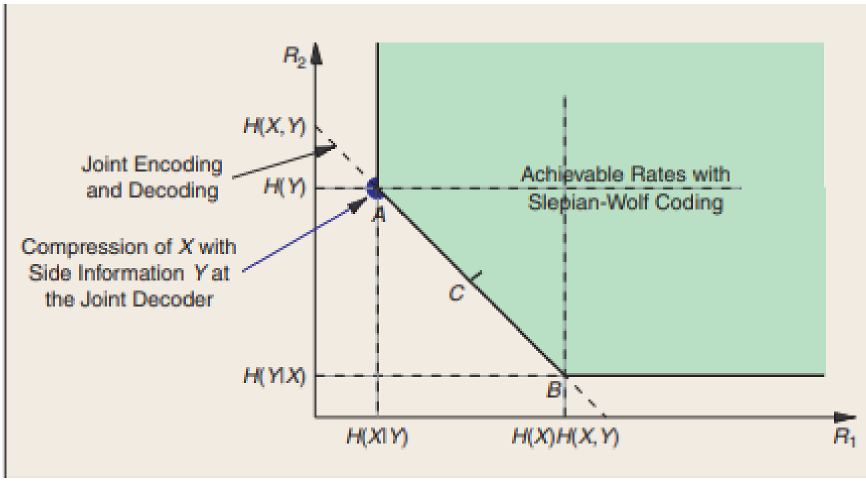


FIGURE 2.6. Active region for distributed source coding [**10**].

region's second corner point B can be approached by switching the roles of X and Y, and all points between these two corner points can be realized. time sharing—for instance, the midpoint C will arise from using each of the two corner point-specific codes 50% of the time. We only take into account code designs for the corner points in this article, even though constructive approaches have been suggested to directly approach the midpoint C in Fig: 2.6 and improvement has recently been made in practical code designs that can approach any point between A and B due to the space limitation. The former approaches are known as symmetric, whereas the latter ones are known as asymmetric. We then compress X into $H(X|Y)$ bits per sample depending on the entire understanding of Y at both the encoder and the decoder. What if, however, a user needs to decode X and Y individually in order to recreate both of them? The rate $R = H(X) + H(Y)$, which is higher than $H(X,Y)$ when X and Y are correlated, is one straightforward method. Slepian and Wolf shown in a study that $R = H(X,Y)$ suffices even for independent encoding of correlated sources. According to the Slepian-Wolf theory, the feasible region of DSC for discrete sources X and Y is provided by $R_1 + R_2 + H(X,Y)$, as illustrated in Fig: 2.6 Random binning is the foundation of the Slepian Wolf theorem's proof. A fundamental idea in DSC is binning, which is the division of the space of all potential outcomes of a random source into disjoint subsets or bins. Using the conditional statistics of (or the correlation model between) X

and Y but not the specific Y at the encoder, we hope to code X at a rate that approaches $H(X|Y)$ in asymmetric coding. In his 1974 article, Wyner proposed the use of linear channel codes as a useful strategy for Slepian Wolf coding after first realizing the tight relationship between DSC and channel coding. The fundamental concept was to divide the space of all potential source outcomes into disjoint bins (sets) that are the cosets of some "good" linear channel code for the particular correlation model. With a linear (n, k) binary block code and binary symmetric sources, there exist two different syndromes that each index a bin (set) of two binary words with a length of n and two Hamming distances. Since each bin is a coset code of the linear binary block code, the original linear binary block code's Hamming distance characteristics apply to each bin[**10**]. Code is kept in each container. A series of n input bits is compressed by mapping it into its corresponding (nk) syndrome bits, yielding a $n : (nk)$ compression ratio. This method, which has been known for some time as "Wyner's scheme," was only recently applied to realistic Slepian-Wolf code designs based on traditional channel codes like block and trellis codes. If X and Y's association can be modeled, By using a binary channel, Wyner's syndrome can be applied to all binary linear codes, and modern near-capacity channel codes like turbo and LDPC can be used to go close to the Slepian Wolf limit. "Turbo and LDPC Codes" gives a brief overview of turbo and LDPC codes. The Slepian Wolf decoder does, however, have a slight chance of overall loss as a result of channel coding. Wyner's technique depends on the correlation model in practice for the linear channel code rate and code design. In order to employ binary (n, k) Hamming codes to precisely accomplish the Slepian Wolf limits, examples like those in "Slepian Wolf Coding Examples" are specifically created with correlation models meeting $H(X) = n$ bits and $H(X|Y) = n-k$ bits. The binary symmetric model, where $(x_1...x_n, y_1...y_n)$ is a sequence of i.i.d. illustrations of a pair of correlated binary Bernoulli(0.5) random variables X and Y, and the correlation between X and Y is modeled by a "virtual" binary symmetric case (BSC) with crossover probability p, is a more useful correlation model than those of "Slepian Wolf Coding Examples. $H(X|Y) = H(p) = p$ in this channel $-p\log_2 p - (1-p)\log_2(1-p)$. Although the Slepian Wolf coding problem appears to be straightforward, it is not. This development could be used to create Slepian Wolf limit approaching codes as the BSC is a well-studied channel model in channel coding and has a variety of capacity-approaching code designs available as a result of the close relationship between Slepian Wolf coding and channel coding. However, unlike in “Slepian Wolf Coding Examples,” the earliest such practical designs did not make the more direct connection with channel coding through the syndromes and the coset codes. Instead, they took concepts from channel coding, particularly turbo codes. The Wyner's system was modified to employ a turbo scheme with structured component codes and parity bits delivered in place of syndrome bits. Liveris et al. used turbo/LDPC codes to develop a code that did adhere to Wyner's scheme for this problem, and they were able to achieve performance that was superior to and very close to the Slepian Wolf limit $H(p)$. Fig: 2.5 displays some simulation results. In Fig: 2.5, the vertical axis represents the likelihood of mistake for the decoded X, and the horizontal axis represents the amount of correlation, where lower $H(p)$ indicates higher correlation. The Slepian Wolf limit is demonstrated to be $0.5b$ at virtually negligible error probability because all of the coding techniques in Fig: 2.5 accomplish 2:1 compression. Wyner-Ziv Coding: The Binary Symmetric Case and the Quadratic Gaussian Case, as well as the section “The Binary Symmetric Case,” describe the performance limit of more general Wyner-Ziv coding. Fig: 2.5 also provides the code-word length for each Slepian Wolf code. It is clear that for a given Slepian Wolf code, the likelihood of mistake decreases with increasing correlation. The lower the correlation required for the code to attain very low error probability, the more robust the Slepian Wolf code is. In the binary symmetric correlation arrangement, stronger channel codes, such as the same family of codes in Fig: 2.5 with greater codeword length, provide better Slepian Wolf codes. We can obtain a good Slepian Wolf code through the syndromes and the accompanying coset codes if the relationship between the source output X and the side information Y can be described with a ”virtual” correlation channel. Any correlation model can be applied to the last sentence. Since Slepian Wolf constraints can be overcome by utilizing near-capacity channel codes like turbo and LDPC codes, the Slepian Wolf coding problem, which at first look seems to be a source coding problem, is actually a channel coding problem.

2.3.1. Wyner-Ziv Coding

As one example of Slepian-Wolf coding, we concentrated on lossless source coding of discrete sources with side information at the decoder in the preceding section. We frequently deal with continuous inputs in network applications, which leads to the issue of rate distortion with side information at the decoder. How many bits are required to encode X under the restriction that the average distortion between X and the coded version X is $Ed(X, X)D$, assuming the side information Y is known at the decoder but not at the encoder? This issue is an example of DSC with Y available uncoded as side information at the decoder, initially raised by Wyner and Ziv. It expands on the design of by coding discrete X according to a fidelity standard rather than a lossless one. Wyner and Ziv provided the rate-distortion function for this issue for both discrete and continuous alphabet situations and general distortion metrics [**11**]. In "Wyner-Ziv Coding: The Binary Symmetric Case," we include. Wyner-Ziv rate distortion functions for the quadratic Gaussian and binary symmetric cases When the side information Y is present at both the encoder and the decoder, Wyner-Ziv coding typically experiences rate loss relative to lossy coding of X (see the binary symmetric case in "Wyner-Ziv Coding: The Binary Symmetric Case and the Quadratic Gaussian Case"). One exception is the quadratic Gaussian situation in "Wyner-Ziv Coding: The Binary Symmetric Case," when X and Y are jointly Gaussian with MSE measure. Since Wyner-Ziv coding recently expanded the no rate loss requirement to $X = Y + Z$, where Z is independently Gaussian but X and Y could follow more broad distributions, there is no rate loss in this situation, which is of particular interest in practice (e.g., IOT networks).

CHAPTER 3

APPLICATION OF DISTRIBUTED SOURCE CODE (DSC)

3.1. 4G and 5G Parameters

We primarily covered lossless (Slepian Wolf) and lossy source coding with side information entirely at the decoder in the talks above because the majority of the effort in DSC to date has focused on these two problems. In order to accomplish this, the nodes transmitting correlated information must cooperate in groups of two or three, especially in communication networks, so that one node provides the side information and another can reduce its information to the Slepian Wolf or Wyner-Ziv limit. This tactic has been applied. Since the decoding loss associated with using lower complexity algorithms may be reduced, this is not a significant issue. In recent years, several low complexity channel decoding methods have been investigated [**1**]. The correlation model is the biggest obstacle to the practical deployment of DSC. In practice, it is typically difficult to come up with a joint probability mass or density function in networks, especially if there is little room for training or little knowledge about the current network topology, despite significant effort being put into DSC designs for various correlation models, and in some cases even application. The correlation statistics in some applications, such video surveillance networks, may primarily depend on the nodes locations. The time-varying correlation could be followed using adaptive or universal DSC if the sensor network includes a training mode option and/or can track the changing network topology. The most effective method for some networks appears to be such universal DSC that can perform well for a variety of correlation statistics, but it is still an open and extremely difficult DSC problem.

DSC can be used to address measurement noise, that is another significant problem in digital communication. If measurement noise is ignored when using the Wyner-Ziv coding method described in the preceding section, there will be a discrepancy between the sources' real correlation and the statistics utilized for noiseless correlation during decoding, which will result in poorer performance. The robust code architecture is one approach to fixing this problem. However, the correlation model can include the noise statistics if they are known, even roughly. The primary market drivers for 5G and beyond are cellular networks and the Internet of Things. Cellular networks and the Internet of Things have a wide range of applications, including e-Health services, autonomous driving, virtual reality, augmented reality, and remote sensing.

With the Internet of Things in mind, 5G is designed to function at faster speeds with almost zero delay, enabling a continuous flow of information. In comparison to the current 4G, mobile IoT-based devices offer a low-cost, low-power option. Comparing it to current wide-area technologies, this 5G enabled mobile network will provide secure connectivity, authentication, and network scalability for capacity expansion. Similar to a 4G system, 5G channel codes for user data should also support hybrid automatic repeat requests and variable code rates and lengths for both control information and user data. Several coding schemes based on the aforementioned specifications are taken into consideration during the 5G standardization process, and LDPC coding has been used for user data focusing on reduced latency in 5G cellular communications [**9**]. In 5G, parallelism in encoding and decoding is crucial to achieving high data throughput. Parallelism in encoding and decoding is naturally maintained in systematic organized LDPC codes, and this parallelism can be used to construct high data rate encoders and decoders. Other crucial 5G channel code characteristics include adaptive rate compatibility, which allows users to choose any number of transmitted codeword bits from the parental code output, and variable code length. These functionalities are demonstrated in the 5G LDPC code design. A recent area of network coding that is also of considerable interest is the coding system proposed for high capacity 5G networks. The trade-off between performance (information processing delay, throughput, bit error rates, etc.) and computational complexity and some necessary overhead is the key consideration in the design of the physical layer for the Internet of Things (e.g., frame overhead, feedback). The fundamental issue with the current cellular system architecture up to LTE 4G is the high data rate downlink for data communication with big packet sizes. The creation of communication physical layer, which can better leverage the wireless channel and can fix the faults that occurred during transmission using error correction techniques like LDPC, is the result of the demand for over growing data rates with limited spectrum resources. The fluctuation of data rate is one of the core components of 5G and IoT. We might only have a little data flow or a very large one. The majority of error correcting codes exhibit performance loss for short code lengths and perform better for larger code lengths. The codeword length and parity check matrix design have a significant impact on the LDPC decoding algorithm's ability to repair errors. Typically, longer codewords and properly constructed parity-check matrices result in greater LDPC decoder performance. In order to fully satisfy the aforementioned needs, we presented medium and small size regular parity check matrices, taking into account the requirements and restrictions of the future communication systems (5G and beyond, IoTs). The suggested small- and medium-size matrices for LDPC codes outperform existing ones in terms of performance, making them more appropriate for these kinds of applications. To implement LDPC codes in telecommunications let us assume two users playing a video game like call of duty in the almost same cluster and the performance of how well the game ran in that specific region in terms of throughput, latency, and user experience is being analyzed by final equipment connected to both the devices in which the game is being played. This is the scenario of testing the network in different regions of United States, by doing the testing we can analyze how to improve the latest generation network(5G) is all the areas. For doing so, we have taken into account certain applications to examine the connectivity of a network like Zoom, WhatsApp, Facebook, YouTube, UDP , Google Maps, and Google Chrome. To do the above testing scenario let us consider two kinds of Networks that is 4G and 5G, as 4G is available in 90% of areas it can give satisfactory output, to establish 5G in all regions we should first check where can 5G technology work properly. In case if we are unable to connect to 5G the device will automatically transfer to 4G, this kind of test scenario will be monitored by observing the transition of networks, in following the above mentioned method there are numerous things to be taken into consideration

UE measures SINR (Signal to Interference plus Noise) based on Resource Blocks (RB). Each RB's SINR is calculated by the UE, converted to CQI, and reported to the eNodeB, which uses it to choose the best MCS for user data transmission in that RB. The MCS to be utilized for an RB, or the number of bits per modulation symbol to be conveyed, determines the throughput that can be accomplished for that specific RB as well as the number of RBs that eNodeB will allot to the user. SINR value is defined as the MCS. The signal-to-interference-to-noise ratio, or SINR, is the ratio of the signal's power to the average of its background noise and other cells' interference. Reference Signal Received Power (RSRP) is a metric for measuring cell-specific signal strength that is used as an input for choices about cell resection and handover. The average power (in Watts) of the Resource Elements (REs) that transmit cell-specific Reference Signals (RSs) inside the analyzed bandwidth is referred to as RSRP for a certain cell. The major use of RSRP measurement, which is often stated in dBm, is to rate various candidate cells according to the strength of their signals. In general, reference signals sent on the first antenna port are used to determine RSRP; however, if the user equipment (UE) can detect their transmission, reference signals sent on the second port may also be used in addition to the RSs on the first port. A signal quality metric that is cell-specific is the Reference Signal Received Quality (RSRQ) measurement. Similar to the RSRP measurement, the major purpose of this metric is to rank various candidate cells according on the quality of their signals. In situations (such as those where the RSRP measurements are insufficient to make valid cell-reselection/handover judgments), this metric might be used as a factor in decision-making. N is the number of Resource Blocks (RBs) in the LTE carrier's Received Signal Strength Indicator (RSSI) measuring bandwidth, and is used to characterize it.

RSRQ = (N.RSRP)/(LTE carrier RSSI).

Received Signal Strength Indicator (RSSI) is the linear average of the total received power observed by the UE from all sources, including co-channel non-serving and serving cells, adjacent channel interference, and thermal noise, within the measurement bandwidth over N RBs for only OFDM symbols carrying reference symbols. To calculate the LTE RSRQ measurement stated above, RSSI is used as an input. As can be seen from the previous Equation, RSRQ takes the combined effect of signal strength and interference into account because RSSI is included. RSRQ and RSRP are mathematically related, as may also be seen.

However, SINR is not defined in 3GPP specifications; instead, UE vendors provide RSRP, RSRQ, and RSSI metrics.

3.1.1. Study of Baseline Devices

In a Case, a mobility scenario is taken into consideration when the handset is travelling at a walking pace from a location with acceptable signal quality to the cell boundary where there is no adjacent cell to make HO to (i.e., the death spot). The measurement results were captured when the phone was playing a game, downloading a file, or making a zoom video call. The user experience for a baseline device will be better on the considered LTE network, with an average throughput of 4.88. The actual bandwidth needed to transfer files of a specified size at a certain time and under specific network conditions is known as throughput. The total speed of data transferred to all terminals in a network is known as the system throughput.

3.1.2. Performance Parameters

The wireless communication link is defined in terms of certain performance measures, such as Quality of Service (QoS). The following are some of the QoS criteria:

Bandwidth: A spectrum of frequencies is referred to. The difference between the highest and lowest frequencies.Bandwidth also increases as the difference between two frequencies grows. The volume of data flowing through a network at any particular time is referred to as throughput. Throughput is the actual link speed as observed in real-time circumstances, whereas bandwidth indicates the total link speed that is available. Uplink Throughput: This indicates that no redundant components have been added to the physical layer. Application data is transferred in the downlink via the PDSCH [**7**].

Because lower frequencies are absorbed by environmental bands and cannot pass through the satellite, up-link performance is lower than downlink throughput. As a result, the frequencies used for uplink from mobile to base station are lower than those used for downlink from base station to mobile. The frequency at which the first earth station interacts with the satellite is known as the uplink frequency.

This signal is changed to a different frequency by the satellite transponder and transmitted to the second earth station using the downlink frequency.

A single radio wave can encode six bits of data using the higher-order modulation method known as 64 QAM. This method changes the radio wave's amplitude and phase into one of 64 distinct and observable states.

The amount of radio resources that can be used to convey data is determined by bandwidth. Throughput depends on Bandwidth, which is also dependent on SINR. Highest SINR for the below graphs is 20dbm. 256 QAM modulation scheme is being used which transfers 8 bits [**2**].

Throughput depends on several factors like modulation scheme, more bits transfer because of higher modulation technique. Eventually will get more bytes to be transmitted for specific second.It is also calculated by the amount of bytes transferred per second. That concludes to if Bandwidth is higher more PRB which increases the Resource Block, more frequency gets modulated which in return leads to more bits to be transmitted.

For stationary testing it depends on RSRP. If RSRP is good the performance is better. SINR will be high, noise gets less in terms of magnitude. Ratio of intelligence signal by noise signal is how we calculate SINR. For example if a room is full of people and if everyone speaks together noise can't be clear(SINR is low). Single signal coming makes it clear. So if intelligence signal can be decoded easily than less Block Error Rate. BLER less than 10% SINR should be more which translates to modulation scheme.

3.2. Working of Two Networks in One Mobile Phone

Consider the above Fig: 3.1 in which there are two cases Non-StandAlone and StandAlone. For NSA the device will have two networks both of which operate from one network. The network provided to the device is 4G and 5G. If the signal is stronger than 5G it will automatically latch to another network. These kinds of devices are less expensive as they will use the frame structure as 4G. When it comes to SA devices they are much stronger and give stronger output compared to NSA. To overcome the above problem let's bring a new mobile phone which can operate two networks on the same device having different operators. It is a technology that allows a capable device to use two sims where one is an active sim and the other one is an inactive sim.

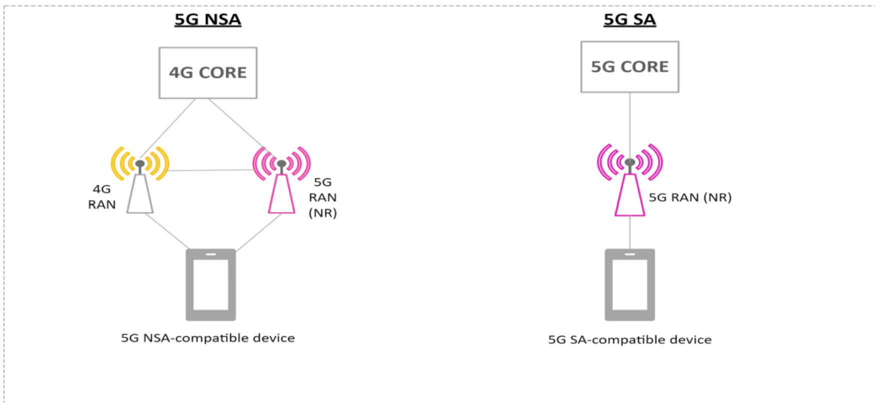


FIGURE 3.1. Non stand alone and stand alone devices

One network can have both 4G and 5G while the other network provides only a 5G network. The basic concept is if you want to latch to another network you will first need to sync with it. Once you latch to LTE it will broadcast the threshold. If the threshold is stronger than LTE you will latch to 5G of the first network p (sim) only. If e sim is stronger than p sim it will shift to another network of p sim.

CHAPTER 4

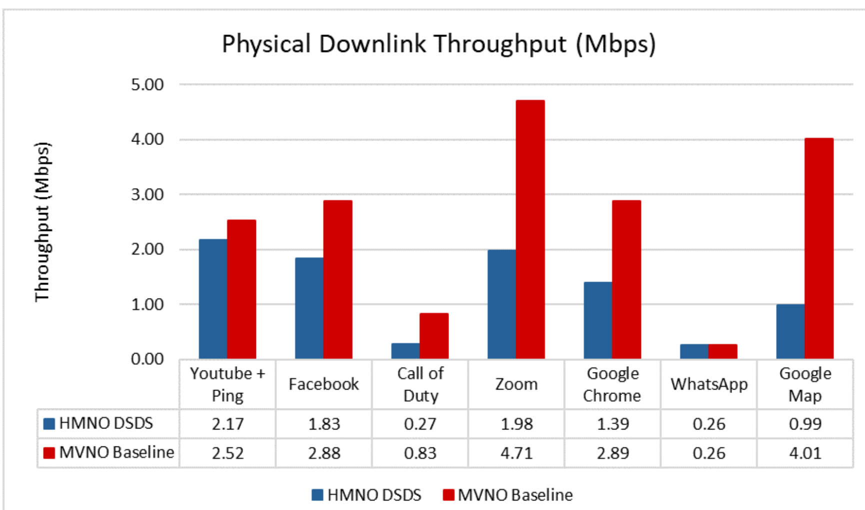
SIMULATION RESULTS OF TWO NETWORKS

4.1. System Working

P Sim(MNO)-Mobile Network Operator E Sim(MSO)-Mobile Service Operator. If it inside the coverage of MSO it will perform data calls, once outside the coverage it will latch to MNO, Uses MNO for voice/sms/mms. For different applications KPI's were measured in order to analyze the performance of the device. As mentioned about throughput it plays an important for signal to be transmitted from transmitter to receiever. Primary sim provides both data and calls whereas E sim only provides data used by MSO for data offload.

4.2. Results

Fig: 4.1 and 4.2 displays the typical throughput numbers. Throughput values have been recorded in various test scenarios while remaining on the same network or switching to a new one.



	Youtube + Ping	Facebook	Call of Duty	Zoom	Google Chrome	WhatsApp	Google Map
HMNO DSDS	2.17	1.83	0.27	1.98	1.39	0.26	0.99
MVNO Baseline	2.52	2.88	0.83	4.71	2.89	0.26	4.01

FIGURE 4.1. Physical downlink throughput

4.2.1. Transition Time and Transition Count

Calculate for how much seconds it is staying on the actual network before transition to another network. This is done by approximately counting if is less than three seconds of network usage it won't be calculated as transition. Let's consider one network as Mobile Service Network and the other one as Mobile Network operator. Wherever it can get access to Mobile Service Network device will start using (it can shift to MNO automatically). Normal calls and messages are still taken place on MNO. MSO is only used for internet related applications. This kind of network transition can be caused when MSO signal changes below or above threshold in cell edge areas. When users move from indoors to outdoors vice versa. If there is any normal text which doesn't uses data it will latch to p sim. When MSO network is at maximum capacity and inactive users' needs to be pushed out. For every application individually transition time and its counts were calculated i.e. for how many times it is latching into another network and for how long it is staying on it. Considering the time for which the device stays on one network by taking into account different test scenarios as shown in Fig: 4.3

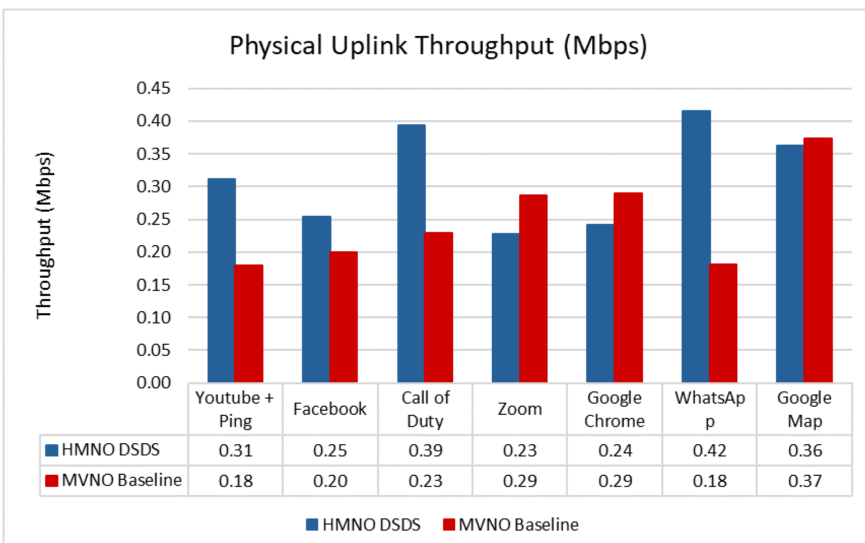


FIGURE 4.2. Physical uplink throughput

UDP DL and Youtube (more data demanding apps) has higher average TP on MSO vs MNO and hence the % data offload to MSO compared with Time on Network as shown in Fig: 4.4

Data Offload: On a WiFi network rather than the 3G or 4G networks, providers redirect traffic. This occurs swiftly and flawlessly, frequently without people even realizing.

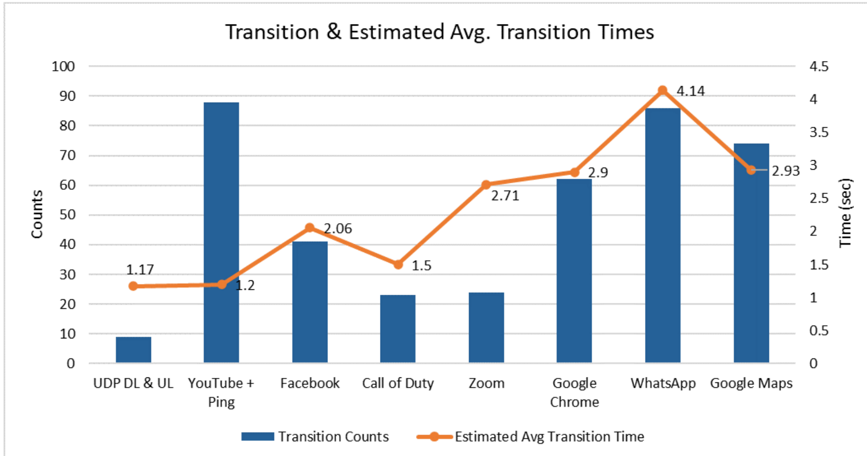


FIGURE 4.3. Transition time

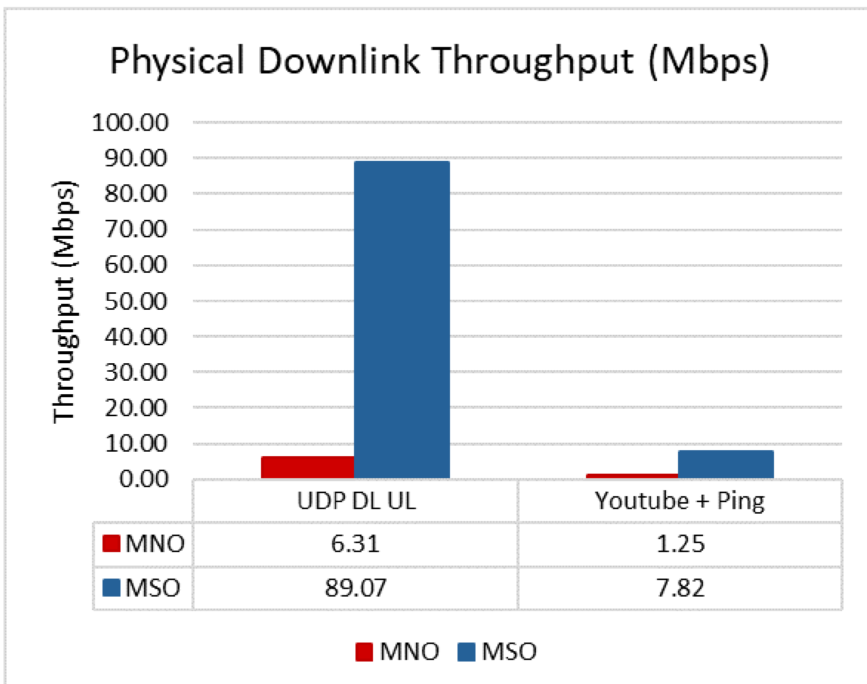


FIGURE 4.4. Physical downlink throughput

It will either switch immediately or, in some situations, always request permission to connect. Specifically for data offloading because using WiFi requires no user action. Since the switch and authentication are embedded into the service itself, they happen automatically. MVNOs are free to keep up their hotspots and WiFi networks so that subscribers can access them without doing anything. If the move is seamless, offloading happens soon. It is adaptable to increase capacity and bandwidth. The time on MSO network varied between 5% – 25%. Data offload in generally range expect a few cases where DL TP is as shown in Fig: 4.5 and 4.6.

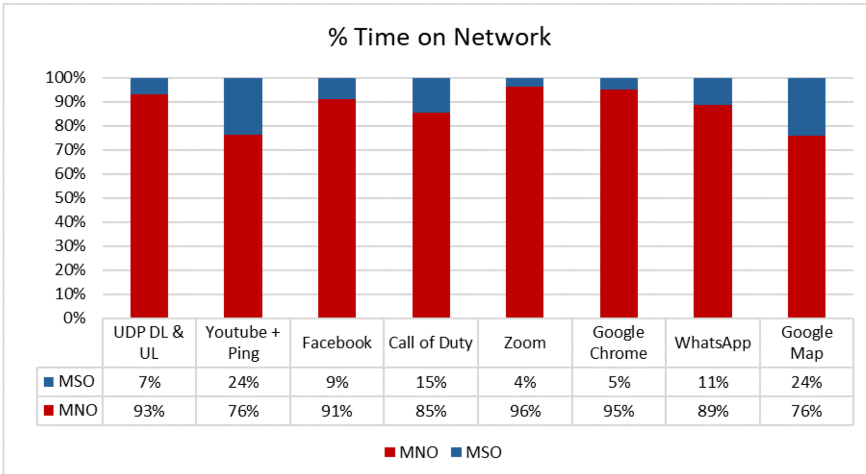


FIGURE 4.5. Time on network

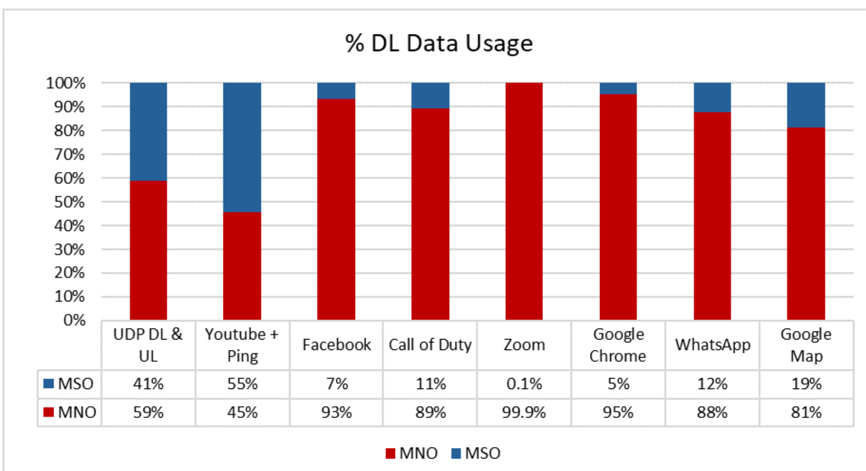


FIGURE 4.6. DL data usage

Its impact on the user experience is if there is a subscription change from one sim to other involves a break before making switch. The user may experience an interruption(jitter) when data switches from one network to another if there is an on-going data session. If user is on MSO and an incoming call takes place, there is a small probability that the incoming calls are missed. If there is an on-going data session and pure SMS/MMS originated or termination, there will be data packet flow interruption. So here we have analyzed several trial and error to implement this new generation technology in a better way.

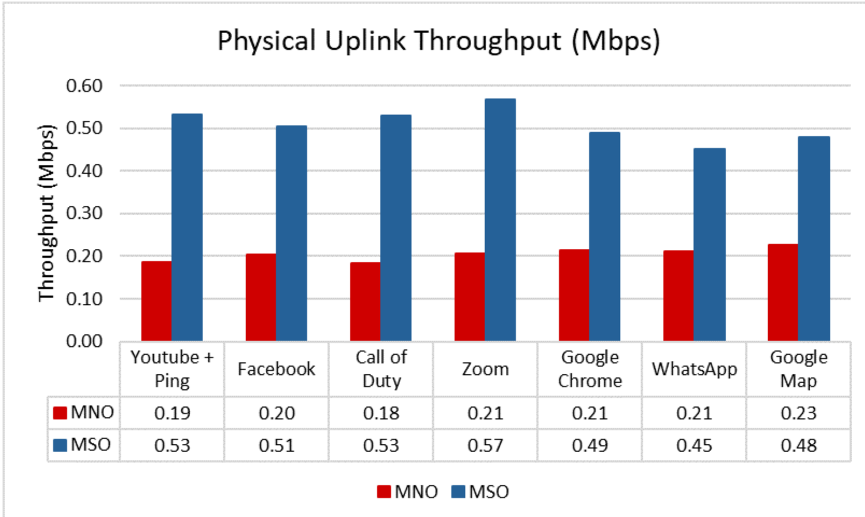


FIGURE 4.7. Physical uplink throughput

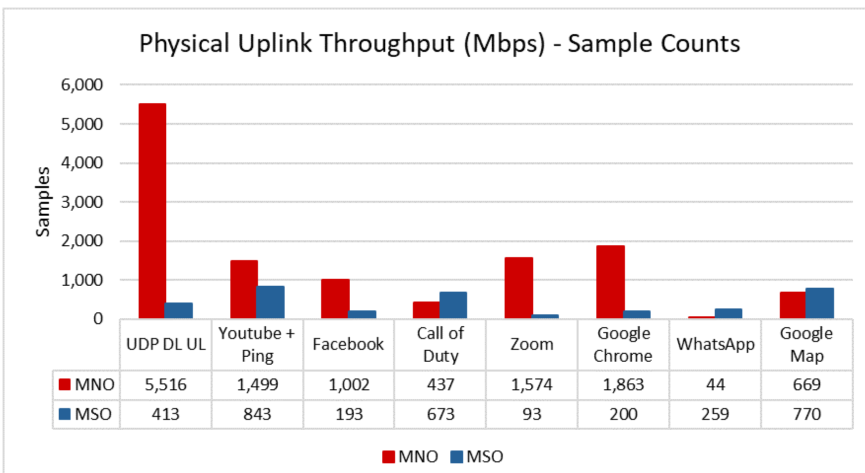


FIGURE 4.8. Sample counts

Samples and Throughput are calculated on the same graph: Samples mean how much time it stays on that network. They want to validate whether the throughput value had many samples on it. Throughput should be analyzed for every sample. 1000's of samples per second were calculated, some of them don't report the throughput values but we want to make sure that we provide average of multiple points and not only one value. Throughput which is above 0.1 is only calculated. More samples less throughput which will result into less average. If we send low throughput and more PRB bits/sec will be less as shown in Fig: 4.7 and 4.8. In the same way it is demonstrated for Downlink Throughput in Fig: 4.9 and 4.10. Let us take an example of a water tank if we find 5 litre of water per 1 second the average speed will be 5litre but for 1 litre of water it takes 3000 seconds speed will be 1l . Speed here is indicated as average throughput and time as total number of samples.

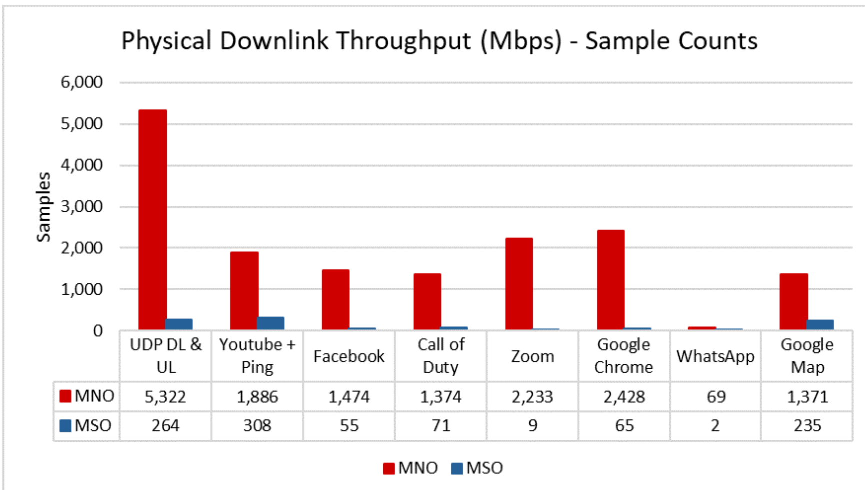


FIGURE 4.9. Sample counts

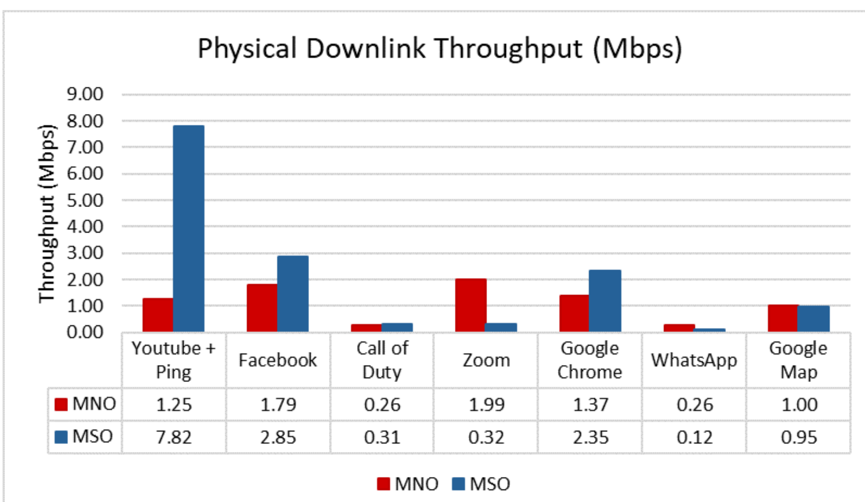


FIGURE 4.10. Physical downlink throughput

No. of samples is calculated by how much time it is staying on that network.

4.3. Benefits and Challenges

4.3.1. Benefits

The advantages of 4G networks contribute to a wider variety of services and user-cases. However, from the user and service provider perspectives, the commercial models and eco-systems that are necessary to drive adoption have not yet been built. Increased uplink and downlink throughput, as well as reduced latency and network capabilities, are provided by the technology. Everyone agrees that the amount of mobile data traffic will increase dramatically over the next few years. Despite the utilization of 4G technology (LTE or WiMAX) in relation to 3G, the bulk of the fundamental throughput and transport bottlenecks will surely be caused by the technology itself. With 4G technologies, spectrum is used at least twice as efficiently and effectively real-time applications are better supported, and maximum speeds are increased. Despite the fact that there are still network and capacity issues to be fully resolved in order to maximize the benefits of the upgrade, these issues include edge or gateway management and signaling management [**7**]. It makes it possible for new mobile applications to improve on already-existing ones (Streaming Music). The increased 4G bandwidth and latency will benefit a number of 4G services, including digital storage and smart home monitoring. Other services like MMS, digital image frames, and numerous applications won't experience any noticeable improvements when using a 4G network. Therefore, it is vital to pay particular attention to the services and applications that 4G developments are likely to improve. We can see that applications like video streaming, MMOG/gaming, and knowledge applications like interactive learning benefit the most from the adoption of 4G technology. Expansion of Addressable Devices: Network potentials and chipset scalability may increase connectivity to a variety of cutting-edge devices. Through the use of smart phones and other more specialized devices, handset technologies continue to advance along a wide spectrum of features and value-added services. The Terminal operating model has traditionally supported a carrier-controlled service experience. Heavy data users have been drawn to commercial operating systems like Windows Mobile which has led to network congestion by losing some control. Additionally, the expanding open eco-systems, made possible by 4G, present operators with a difficult opportunity because third parties create services, applications, and personalization tools to satisfy customer wants. Because of open standards, devices are becoming much more customizable, and specialized devices like netbooks, eReaders, tablets, etc. are entering the market. Vendors, in our opinion, need to consider a micro-segmentation-based device roadmap to cater to the needs of smaller user segments; additionally, new channels of distribution are necessary to enable the adoption of converged mobile devices and 4G apps. Managing user expectations and experiences with new features and services is made possible by a differentiated customer experience. We take into account the user's experience in gaining a thorough understanding of how these services are entirely facilitated and how they integrate into the fabric of our daily lives, the necessity or capability to deploy expertise or configured gadgets to support enhancement, and finally, how to make money and when to share the income from the service delivery. It has been difficult to comprehend what a 4G user goes through, and it is unclear how much this will change as more and different 4G services become available. We employ monetization strategies. Additionally, we think that customers expect that other services and applications will be combined into a "solution" that enhances their quality of life. Therefore, resolving the most likely User-Cases for 4G services will be essential for the successful adoption of 4G services. Being very conscious of the fact that user expectations for price points are shifting, with growing demands for "a little for a little," which challenges the current costs. The bandwidth requirements of various 4G use cases indicate that, if current pricing systems continue, the aforementioned issue would only worsen. One option that operators are currently considering is shifting toward tiered pricing based on traditional factors like time, speed, and service quality. The bandwidth-on-demand service model, together with the associated pricing strategy to charge premium prices for these burst requirements, is an extra capable service type. This could be useful for organizing high bandwidth events like video streaming or LIVE TV. Given what we know at this moment, pricing structures for 4G wireless technologies will need to be expanded in order to favor lower upfront costs (subscriptions, one time purchases, ad-based, fermium and per-use).

4.3.2. Challenges

Security and privacy: To enable the safest method for data transfer, security measures must be implemented in the creation of 4G Wireless Networks. "The 4G core delivers mobility, security, and QoS by leveraging the existing approaches while still working on a few mobility and handover challenges," it states explicitly. Therefore, the company must create a set of efficient and effective technologies that will support the strictest 4G security procedures in order to protect data being carried across the network from hackers and other security breaches. Due to the nature of the 4G wireless network, security incursions are more likely to occur, thus various degrees of security, including greater requirements for validation, will be crucial for protecting data and information transmitted across the network. Therefore, it is crucial for service providers to create a method for the 4G Wireless Networks that is efficient and effective, will improve quality, bestow effective security measures, and will ensure that all users are given access to a wide range of options for quickly downloading music, video, and picture files. Integrating IP-based and non-IP-based devices is the main challenge for 4G wireless networks. We are aware that devices without IP addresses are frequently utilized for VoIP services. Contrarily, devices that rely on IP addresses are typically utilized to deliver data.

CHAPTER 5

FUTURE SCOPE AND CONCLUSION

We have provided a helpful method for source coding in the presence of side information at the decoder in this study. This strategy, which is based on concepts from algebraic channel coding, produces encouraging outcomes. These systems are based on the fundamental idea that the source is quantized first using a source codebook created for the marginal probability density function of the source. Cosets of a channel code created for the fake channel between the side information and the quantized source are divided up into the source codebook. The encoder simply delivers the index of the coset containing the quantized codeword; it does not send the index of the quantized codeword. By decoding the side information in the designated coset of the channel code, the decoder is able to recover the quantized codeword. It quantifies and analyzes the intricate interplay between the estimation component, the channel coding, and the source coding. We began by considering a construction in which both the source code and the channel code lack memory before moving on to the scenario in which both do. The design of such codes is also covered, and simulations are used to assess performance. We have observed that irregular codes perform much better and can reach the Slepian Wolf limit compared to regular and turbo codes. We studied in brief about encoders and decoders in detail and how the complexity of decoders is reduced using LDPC codes. This study used a mobile device application to measure the Quality of Service characteristics based on time to the cellular network. The service card package provided by is the method utilized to gauge the end-to-end parameters of the three providers. Similar to PING, this approach only evaluates overall network speed without taking into account the specifics of a certain application's content. Additionally, measurements have been made by testing in other places around the USA. Test scenarios designed to show how mobile ping and other capabilities can be used while also giving a basic overview of performance for each instance in each region. Our main aim is to establish strong coverage in all areas for better user experience. In order to test any potential correlations between

LTE measures like SINR, RSRP, RSSI, and RSRQ as well as to assess the effects of SNR on throughput, an examination of some practical measurement findings taken from a live LTE network is presented in this study.

The handover events that took place within LTE during the test time are also investigated. The average relationship between RSRP and SINR has been observed, and the smaller the gap between RSRP and RSSI, the better the RSRQ; if the SINR is better for a measurement slot, higher throughput is attained. Additionally, it has been observed that the handover event occurs to the adjacent cell when the RSRP and/or RSRQ of a serving cell fall below that of the neighbour cell.

The acquisition and creative use of spectrum will be more crucial than ever as the demands placed on mobile communication networks increase. Better utilization of the spectrum now accessible to mobile networks, access to extra bandwidth at comparable frequencies, and manipulation of higher frequencies in the centimeter-wave and millimeter-wave bands are all necessary to meet upcoming needs. High tenacity and bi-directional shaping of big bandwidth. All networks can be connected together using technology more dynamic and efficient. Utilizing technology to streamline subscriber administrative tools. Most likely, will offer a sizable amount of transmitting data supporting over 60,000 connections. Simple to handle for earlier generations. A solid technology base to provide a wide range of service areas (including private network).It is possible to pay for global connectivity that is constant, unbroken, and reliable. In the near future, 5G will provide greater service quality, reduced latency, and increased bandwidth, all of which will enhance user experiences in both the consumer and corporate sectors, ranging from telemedicine use cases to cloud gaming. The Internet of Things will change thanks to 5G networks (IoT). However, it will take some time until the majority of the earth is covered by technology. WiFi will manage the local wireless connection for the majority of users, while 5G will handle wide-area wireless connections. However, there may eventually come a point where only one of them will be necessary. It might seem absurd to consider that WiFi might disappear, especially given how commonplace it is now. Greater capacity, more users, and faster speed on the improved spectrum.

REFERENCES

[1] J.E. Barceló-Lladó, Antoni Morell, and Gonzalo Seco-Granados, *Distributed source coding in large wireless sensor networks*, 2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers, 2010, pp. 1535–1539.

[2] Ashish Kumar, Ankit Aswal, and Lalit Singh, *4g wireless technology: A brief review*, International Journal of Engineering and management Research 3 (2013), 35–43.

[3] Angelos Liveris, Zixiang Xiong, and Costas Georghiades, *Compression of binary sources with side information at the decoder using ldpc codes*, Communications Letters, IEEE 6 (2002), 440 – 442.

[4] Sukhleen Bindra Narang, Kunal Pubby, and Hashneet Kaur, *Low-density parity check (ldpc) codes: A new era in coding*, Signal And Image Processing 1 (2016), no. 1, 7–14.

[5] Soon Xin Ng, Kai Zhu, and Lajos Hanzo, *Distributed source-coding, channel-coding and modulation for cooperative communications*, 2010 IEEE 72nd Vehicular Technology Conference - Fall, 2010, pp. 1–5.

[6] S.S. Pradhan and K. Ramchandran, *Distributed source coding using syndromes (discus): design and construction*, IEEE Transactions on Information Theory 49 (2003), no. 3, 626–643.

[7] Vinayak Pujari, Rajendra Patil, Mr Tambe, and Asst Prof, *Research paper on future of 5g wireless system*, (2021), 2021.

[8] Amin Shokrollahi, *Ldpc codes: An introduction*, Coding, cryptography and combinatorics, Springer, 2004, pp. 85–110.

[9] Velotiaray Toto-Zarasoa, Aline Roumy, and Christine Guillemot, *Source modeling for distributed video coding*, IEEE Transactions on Circuits and Systems for Video Technology 22 (2012), no. 2, 174–187.

[10] Zixiang Xiong, A.D. Liveris, and S. Cheng, *Distributed source coding for sensor networks*, IEEE Signal Processing Magazine 21 (2004), no. 5, 80–94.

[11] Yang Yang, S. Cheng, Zixiang Xiong, and Wei Zhao, *Wyner-ziv coding based on tcq and*

ldpc codes, The Thrity-Seventh Asilomar Conference on Signals, Systems and Computers, 2003, vol. 1, 2003, pp. 825–829 Vol.1.