Wireless Communication

Anonymous CVPR submission

Paper ID 0001

1. Overview

- At start we want to discuss what exactly mean by information and we will discuss about the systems that will communicate.
- Communication will take a wireless signal and read it as code words we use mainly these two 0,1 and gives output of the taken input from wireless signal.
- Wireless signals mostly collide with external walls and some other matters and hence attains some deviation in signal which we calls it as noise we don't have much control over noise. But generally noise is represented as gaussian noise which will take values as same as gaussian random variables.
- Then to read the signal that came to the decoder will have random noise we define SNR value to it to make sure that noise won't effect much to the driven signal.
- We have some type of Channels which the decoder have some type of conditional pdf we will discuss one case of it BSC(p) which is also called binary symmetric channel
- I would like too introduce terms Huffman codes,shannon codes, entropy,mutual information, channel capacity, rate of a communication system $information \rightarrow channel/communicator \rightarrow decoder \rightarrow received\ information$
- There are probability examples in each of the above cycle.
- Mainly we will deal with simple probability examples as wireless communication has a wide range of examples

2. Information

- Information is the data of an experiment which you are interested in this information will be converted into codes in the communication.
- Lets say we have 4 telephones we want to have the data of each telephone let's say we name it as $00, 01, 10, 11$ then 2 bit information from the telephone will be able to discriminate all telephones so we need a 2 bit information for 4 telephones similarly if u have 8 phones then $000, 001, 010, 011, 100, 101, 110, 111$ so from this we can say that if u have n elements and information required to represent all bits is called $log_2 n$.

We call it as entropy represented as H(x)

$$H(x) = log_2 n;$$

- But here we didn't taken into count about the probability of possibility of using that telephone we assumed all of them has equal probability so by intuitive we get less information from high probability and high information from low probability so given H(x) will include the probabilities of each individuals. Which has n elements in it.

$$H(x) = \sum_{i=1}^{n} P_i log_2 \frac{1}{P_i}$$

Where P_I will be the probability of the I^{th} element to happen

- Information also depends on the probability so it is one example.

3. Noise Addition in passing the signal

- Lets say we made a signal like if the signal value is above 0 then 1 other wise 0 as bitwise information $\{0, 1\}$.
- Now on passing the signal to medium of air which will have many distributions lets take a simple bit wise information sent by cosine signal as below

Recieved signal $x[n] = Acos(2\pi f_o n)$(Transmitted signal)$+V(n)$(gaussian noise)

- Gaussian noise as name suggests it follows the gaussian rv function so $v(n)$ is the gaussian random variable so for this gaussian random variable we have certain variance lets say the variance as σ^2 Now we will take a value SNR(signal to noise ratio) And there are so many ways to define this SNR value we mostly define as

  For zero expectation signals

  $$\text{SNR} = \frac{A^2}{2\sigma^2};$$

  $$\text{SNRdb} = 10{*}log_{10}(SNR);$$

  For non-zero expectation signal

  $$\text{SNR} = \frac{E(signal\, with\, uniform\, distribution)}{E(Noise\, with\, guassian\, rv)};$$

  $$\text{SNRdb} = 10{*}log_{10}(SNR);$$

- For every signal and noise we have some threshold SNRdb value we have to increase the amplitude of the transmitted signal to maintain that signal is least effected by the noise.
- For this we need to do random experiments by passing and we need to find the variance and with that we will find a Threshold SNRdb where the signal is least effected.
- We showed in the MATLAB code results are displayed in the upcoming sections

4. Noise channel/Random channel/ Probability Channel

$$\text{INFORMATION(X)} \longrightarrow \boxed{\text{Noise Channel}} \longrightarrow \text{DECODED(X)}$$

- In above diagram X is an rv we are passing as information Probability channel is a channel that connects $x \in X$ to some $y \in Y$ with $Y \equiv X$; So we can say that for a given $x \in X$ The decoding channel choose a $y \in Y$ but here $y \neq x$ there will be independent probability distribution on y for a given x so we will get an error.
- We choose the channel such that $P(error) \leq e$. We need to take the code words such that Probability of error will be minimised.
- For each $x \in X$ we assign a code word {eg: 0,01,001} which is prefix free code for sure and the set of all code words is called code of a random variable it is denoted by C.

Y/X	Y = 0	Y=1
X=0	1-p	p
X=1	p	1-p

Table 1. Probability Table

- So we represent a code word with $\underline{c}$ as we run the source many times the under line below indicates the vector.
- Code length is the the no. of bits for codeword we represent it as $l(x)$ For $x \in X$
- We find the expectation of code length by finding the expectation of code length for the probability distribution of X as an rv.
- We may pass X as an continuous RV or discrete RV mostly we passes discrete RV because we can't continuously control the signal.

  $$\bar{L} = \sum_{x \in X} P(X = x) l(x);$$

- Our objective is to minimise the expectation of code lengths so we assign different length of code words to $x \in X$ based on their probability.
- If we know that $P(x_1) \leq P(x_2) \leq$ ....$P(x_n)$; then by intuition to minimise the length of code words we take the code lengths as $l(x_1) \geq l(x_2) \geq$ ....$l(x_n)$.
- These all will depend on the probability distribution of RV X

5. Example of probability channel is binary symmetric channel

- ***BINARY SYMMETRIC CHANNEL***:
  - ∗ The binary symmetric channel.This type of channel transmits only two distinct characters, generally interpreted as 0 and 1, hence the designation binary. The probability of correctly receiving either character is the same, namely, p, which accounts for the designation symmetric.
  - ∗ We showed in the MATLAB code results are displayed in the upcoming sections
- The channels $P(Y/X)$ will look like the above probability table as we said for each $x \in X$ we have a probability distribution in $y \in Y$
- ***SHANNONS THEOREM***:

- $*$ The rate of a code for a given code C And the channel with $P(error) \leq e$ will always be less than the channel capacity of the channel

- ***RATE OF A CODE***:
  - $*$ Rate of a code is simply defined as R = $log_2\frac{|C|}{n}$ Where source is runned n times and $|C|$ Represents the cardinality of the code.

- ***CHANNEL CAPACITY***:
  - $*$ It is simply defined as the maximum mutual information between X,Y for a channel.

    C = max[I(X;Y)];

- ***MUTUAL INFORMATION***:
  - $*$ mutual information I(X;Y) is defined as the information acquired after knowing Y we already stated as the information is H(X) for an rv X; So,

    I(X;Y) = H(X)-H(Y/X);

- By using the entropy formula using above channel probability distribution we can see that Channel code for this is $1 - H_2(p)$ where

$$H_2(p) = plog(1/p) + (1-p)log(1/1-p);$$

- So from this we can say rate $R < 1 - H_2(p)$; $1 - H_2(p)$ is specially named as $C_{BSC(p)}$

- MAIN PROBABILITY things will appear in achievability of the SHANNONS THEOREM

- ***ACHIEVABILITY***:
  - $*$ We need to prove that for $R = C - e$ We will achive an P(error) = 0; as $n \to \infty$
  - $*$ We will be doing this for the n times runned source of this BSC channel

- $\underline{x}$ be the x and n length vector passed through the channel where $\underline{x} \in X^n$ then $\underline{x}$ will have a codeword $\underline{c} \in C$. $\underline{y}$ be the y and n length vector passed through the channel where $\underline{y} \in Y^n$

- We know that $|C| = 2^{nR}$.$\hat{c}$ is the output codeword for a $\underline{c}$.

- P($\underline{x}$) $= \prod_{i=1}^{n} P(x_i)$ As each of them is independent variables. Similarly P(y/$\underline{x}$) $=\prod_{i=1}^{n} P(y_i/x_i)$

- Decoder function be D(y) = arg[max p(y/$\underline{c}$)]

- P(error) $= P(U_{c\in C}\hat{c} \neq c)$.

- We passed that X as a uniform RV so

  P(error) = $\frac{P(\hat{c}\neq c)}{2^{nR}}$.

  As both of them are independent and as X is an uniform RV

- Now we need to find $P(\hat{c} \neq c)$.

- We need to minimise P(D($\underline{y}$) not equal to $\underline{c}$)

- D($\underline{y}$) choose a y where maximum probability of P($\underline{y}/\underline{c}$) so expanding this individually as they are independent random variables. $d_H(y, c)$ Is the no. Of codewords interchanged.

  P($\underline{y}/\underline{c}$)= $\frac{p}{1-p}^{d_H(y,c)}.(1-p)^n$;

- We took the minimal point BSC which means p¡(1/2) so we need to pic a lower dH and we know that interchanges follows binomial RV so for large n we expect np number of inter changes so the sets of all codes y belongs to dH from 0 to np be B($\underline{y}$,np);

  P $(\hat{c} \neq c)$ = P $(\exists\hat{c} \in B(y, np) : \hat{c} \neq c)$

  $\leq \frac{B(y,np)}{2^n}$(Binomial distribution)

  $\leq \frac{\sum_{i=0}^{np} C_i^n}{2^n}$

  $\leq \frac{C_{np}^n}{2^n}$

  As n is larger other ones can be neglected

  $\leq 2^{-n(1-H_2(p))}$

- We can write C_{np}^n as $2^{H_2(p)}$ so we got P$(\hat{c} \neq c)$

  Implies P(error) $\leq \frac{1}{2^{nR}}.2^{-n(1-H_2(p))}$

  $\leq 2^{-n(1-H_2(p)-R)}$

  $\leq$ $2^{-n(C-R)}$ as we know C-R = e

  $\leq 2^{ne}$ HENCE PROVED

- Hence achievability is proved using many tricks from probability

6. SHANNON CODES

- As I mentioned in the Noice channel we need to reduce expectation of code length so here shannon comes with a way to support that as intuitively most probability one should have the lowest length intuitively he given code lengths as below

$$l(x) = GIF(log(\frac{1}{P_x}))$$

$$GIF = (Greatest\ integer\ function)$$

- It is not optimal code but an accurate way to use. Optimal code is a code where it is minimised to the least and cannot be minimised more

7. HUFFMANS CODES

- Huffman's codes will give an optimal codes by taking the least two and taking their parent like a binary search tree and it will be going from child's to parents side using that recursively(taking two min probabilities and attach it to one parent) now parent will have Probability as sum of their child's. We coded a optimal code by this but not much use of probabilities.

$$H(x) \leq \bar{L}_{Huffman} \leq \bar{L}_{Shannon} \leq H(x) + 1$$

- Above in equation is for passing it for one time if it passed for n times then??

$$H(x) \leq \frac{\bar{L}_{Huffman}}{n} \leq \frac{\bar{L}_{Shannon}}{n} \leq H(x) + \frac{1}{n}$$

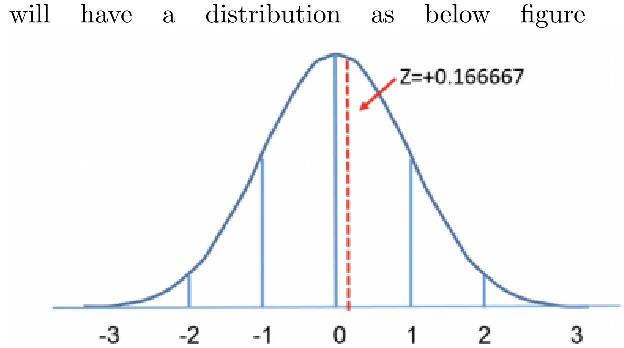
- As $n \to \infty$ Then the code lengths as same as that $H(X)$ per one use of source.
- We decoded information using probabilities

8. Signal Processing on EMW(Electro Magnetic Waves) and collecting Information

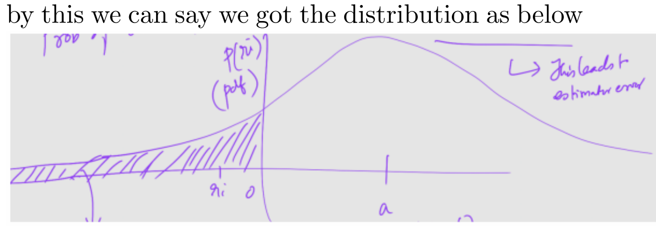
- We can't possibly transfer bitstream manually we have to pass it by some method.here in wire less communication we have to pass the signal as EMW not manually here we have some challenges we are going to discuss.
- If the output coming from the source is 1 we will create a signal will give an output +a from $(k-1)T$ to kT and if the output coming from the source is 0 we will create a signal will give an output -a from $(k-1)T$ to kT

$$Signal(x(t)) = \begin{cases} \text{+a,} & \text{for (k-1)T to kT, if input is 1} \\ \text{-a,} & \text{for (k-1)T to kT, if input is 0} \end{cases}$$

- Now this $x(t)$ is passed through the air medium to transfer and receiver will get the signal as $y(t)$ where $y(t)$ is as follows $y(t) = x(t) + r(t)$,while passing through the medium it will have some errors which are taken as gaussian errors so now we need to find the bit stream from the received signal as 1 or 0.
- From above the $r(t)$ is taken as gaussian error with mean at zero so it will have a distribution as below figure



- Now lets say we have the transferred signal as +a then received signal will have $a + r(t)$ so by this we can say we got the distribution as below



- Lets say we will point if the receiver got +ve value then it plots to bit 1 other wise 0, Then we will get error if we get $a + rI$ is negative so $p(error) = p(a + rI <= 0)$ implies that $p(error) = p(rI <= -a)$ which is equal to the value

$$F(X = -a) = \frac{1}{2}[1 + erf\frac{-a}{\sqrt{2}\sigma}]$$

its value will be so low when the value in erf is high so higher the input of erf lower the error we named its square as SNR ,

$$SNR = \frac{a^2}{2\sigma^2}$$

So we will adjust a as such the error is low

- In the similar way for an input is -a the probability of getting error is same because it is the gaussian distribution with mean 0 so in the same way $p(error) = F(X = -a)$ where F is gaussian distribution with mean 0.

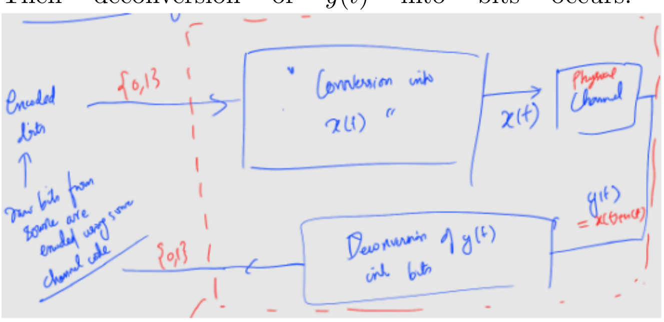
- From the gaussian distribution analysis we found the probability of error and we can use it minimise the error

9. A short discussion on wirelines and wireless

$x(t)$ is a signal which is transmitted through a physical channe and we get $y(t)$

$$y(t) = \alpha(t) + n(t)$$

Noise signal is a random process. Some statistics are known about this suppose $\alpha(t) = \alpha$ is a constant, $\forall$t then we can handle it by boosting the signal power.This comes under wireline communication. So the channel model can be simplified as $y(t) = x(t) + n(t)$. If $\alpha(t)$ is a function of time, then it comes under wireless communication model.This also means that $\alpha(t)$ is another random process. Now how is $x(t$ generated? Raw bits from source are encoded using some channel code. Then the encoded bits are converted to $x(t)$. Then $x(t)$is passed through a physical channel. This gives $y(t) = x(t) + n(t)$. Then deconversion of $y(t)$ into bits occurs.



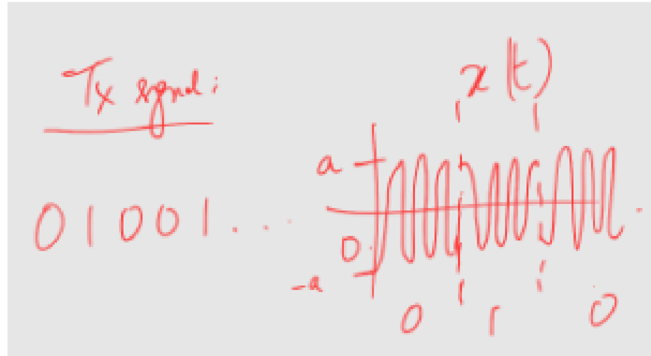
The binary aspect is clear here. It becomes symmetric under some assumptions on- 1)The physical channel 2)How we do conversion and deconversions. Conversion of bits to signals: First there is a bits to message signal mapper. Thus we get a digital signal. The digital signal is modulated by using carrier signal.

$$c(t) = cos(2\pi f_c t)$$

$c(t)$ is a sinusoid with carrier frequency f_c If we do the mapping in this way: 0 is mapped to $(-a)rect((k-1)T, kT)$ 1 is mapped to $(a)rect((k-1)T, kT)$ Therefore the transmitted signal is:

$$x(t) = m(t)c(t)$$

0 is mapped to $(-a)cos(2\pi f_c t)$
1 is mapped to $(a)cos(2\pi f_c t)$



Therefore the received signal is $y(t) = x(t) + n(t)$, where $x(t)$ is as shown above. Then $y(t)$ is passed through a demodulator in which we multiply the received signal with $c(t)$ and pass it through a low pass filter. We get $m(t) + n'(t)$. We pass $m(t) + n'(t)$ though a sampler which samples at intervals of T. T_x samples are:

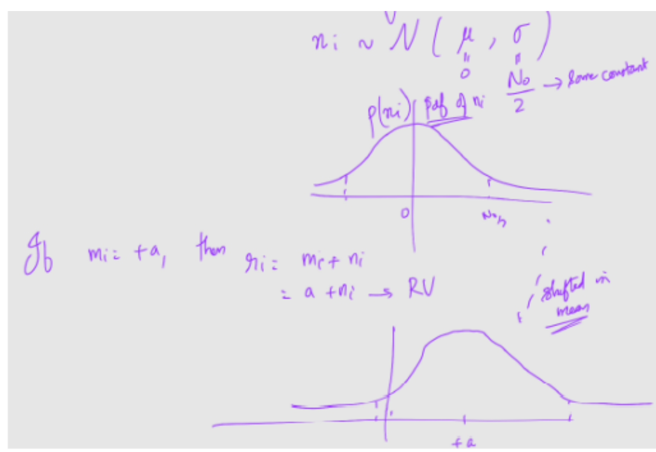
$$m_i \in (+a, -a)$$

R_x samples are:

$$r_i = m_i + n_i \in (+a + n_i, -a + n_i)$$

n_i is the sample of a random process. Therefore it will be a random variable. We will assume that n_i is Gaussian distributed.

$$n_i \sim N(\mu, \sigma)$$



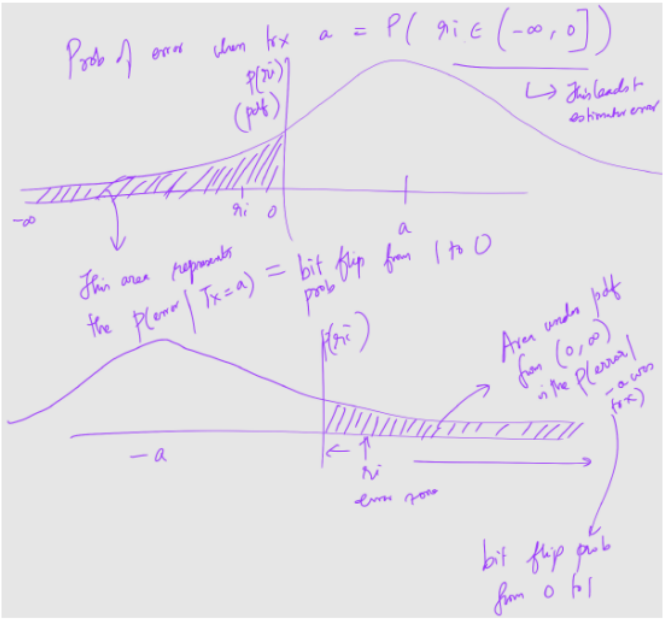
If $m_i = +a$, then $r_i = m_i + n_i = a + n_i$ If $r_i = m_i + n_i$ is passed through an estimator for m_i , it is estimated by Maximum likelihood estimate.

$$m_i' = argmaxp(r_i|trxwas = m)form \in (+a, -a)$$

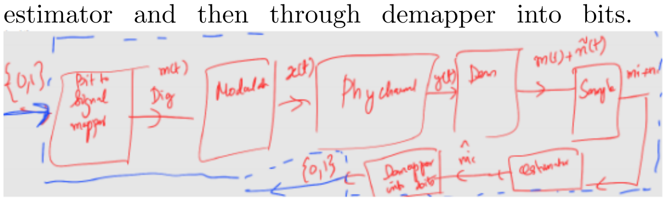
$P(R_x = r_i|trx = a) = P(noise = r_i a)$ $P(R_x = r_i|trx = a) = P(noise = r_i + a)$ Now plugin the noise pdf(Gaussian).

$$argmaxp(r_i|trxwas = m) = argmin|rim|$$

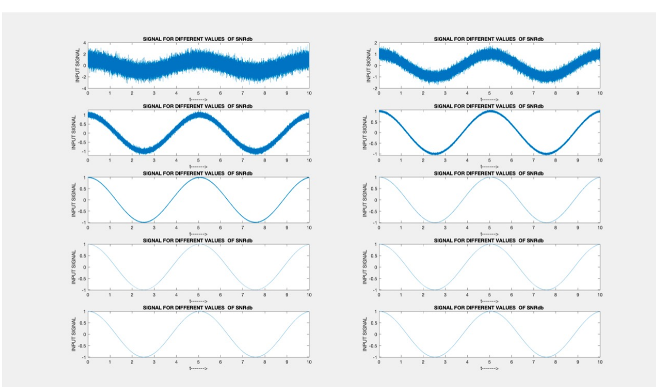
for $m \in (+a, a)$ ML decoder decodes outputs either +a or -a whichever is closer. Probability of error when trx is $a = P(r_i \in (\infty, 0])$



Therefore a communication system is as follows: First we pass the bits through bit to signal mapper. Then we pass the digital signal through modulator. We get $x(t)$ which we pass through a physical channel and we get $y(t)$. Then we pass $y(t)$ through a demodulator. Then we pass it through sampler and then through estimator and then through demapper into bits.

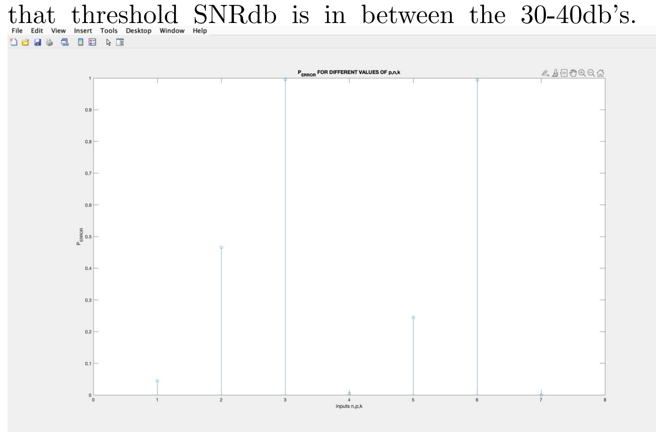


10. Results from MATLAB code



graphs for input signal vs time FOR DIFFERENT VALUES OF SNRDB VALUES STARTING FROM 1 to 91 WITH SPACES 10

In above we found A by changing different values of SNRdb and plotted the signal output from the figures we can say that A increases the signal is least effected as SNRdb is directly proportional to A. Now we do know that threshold SNRdb is in between the 30-40db's.



FOR DIFFERENT VALUES OF N,K,P where p defines the flip probability K denotes the value RATE X N, N is the no. of times we run the source to get vector of x(underlined) as $0, 1^N$

Here we observe the binary coding channel error with various n, k, p value anad observing the results of probability error and providing exmaples for Shannons Thereom

- Number - 1
  - Channel Capacity for n = 15,k = 10,p = 0.015

    0.8876
  - Rate of the process is

    0.6667
  - Rate < Channel capacity in this binary channel we get low probability error

    (p = 0.0535)
- Number - 2
  - Channel Capacity for n = 15,k=10,p=0.1

    0.5310
  - Rate of the process is

    0.6667
  - Rate > Channel capacity in this binary channel we get high probability error

    (p = 0.4565)
- Number - 3
  - Channel Capacity for n = 15,k=10,p=0.45

0.0072

- Rate of the process is

0.6667

- Rate > Channel capacity in this binary channel we get high probability error

(p = 0.9955)

- Number - 4
  - Channel Capacity for n = 20,k=10,p=0.015

    0.8876

  - Rate of the process is

    0.5000

  - Rate < Channel capacity in this binary channel we get low probability error

    (p = 0.0045)

- Number - 5
  - Channel Capacity for n = 20,k=10,p=0.1

    0.5310

  - Rate of the process is

    0.5000

  - Rate < Channel capacity in this binary channel we get low probability error

    (p = 0.2410)

- Number - 6
  - Channel Capacity for n = 20,k=10,p=0.45

    0.0072

  - Rate of the process is

    0.5000

  - Rate > Channel capacity in this binary channel we get high probability error

    (p = 0.9970)

- Number - 7
  - Channel Capacity for n = 50,k=10,p=0.015

    0.8876

  - Rate of the process is

    0.2000

  - Rate << Channel capacity in this binary channel we get low probability error

    (p = 0)

we created a huffman codes and encoded the data and decoded the data in codes and made files we will submit it with the code

11. References

- Applied Digital Information Theory – James L.Massey
- https://ieeexplore.ieee.org/document/7465728

12. Work by

- Banala Vishwa teja Reddy– 2020102058
- Gokul Praneeta – 2020113004
- Vivek Mathur – 2020113002
- Syed Imami – 2020113012