NON-SHANNON PERSPECTIVE ON CHANNEL CAPACITY FOR BINARY SOURCE CODES USING WEIGHTED CONDITIONAL ENTROPY

Abstract

Discrete memory-free channels with a very low capacity are known as noisy channels. Our recent study has yielded some new insights into the channel capacity of noisy channels, which could prove useful in the development of mathematical models for these channels and other contexts.

Keywords and phrases: Non Shannon entropy, Channel capacity, Weighted entropy, Conditional entropy, Mutual information, Information theory etc.

1. INTRODUCTION:

By definition, noisy channels are discrete memory-less channels with low capacity. These channels were initially created by Reiffen [10] to explore physical channels operating at low ratios. Researchers [1, 11, and 13] have demonstrated that even in cases where channels are merely functioning at low capacity, these channels nonetheless play a significant role in communication. Majani [6] provided a broad mathematical model of noisy channels and demonstrated that Reiffen's description does not apply to all channels with low capacity. In his classification of noisy channels, he distinguished between two types: type I and type II. All of the Type I Noisy Channels fall within Reiffen's description.Certain Type II Noisy Channels, such as the noisy Z channel, are not covered by Reiffen's definition. Our effort will mostly concentrate on Reiffen's Type I Noisy Channels.

Let as assume that there is a discrete memoryless channel with input and output alphabets of X and Y respectively, input probabilities $P(x)$,output probabilities $P(y)$, and transition probabilities$P(y/x)$, andboth X and Y are taken to be finite. Reiffen states that the channel is extremely loud if, for all $x \in X$ and $y \in Y$, then we have

$$\frac{P(y)-P(y/x)}{P(y)} = \varepsilon_x(y) \ll 1 \quad (1.1)$$

Equation (1.1) also written as

$$P(y/x) = P(y)(1-\varepsilon_x(y)) \quad (1.2)$$

We observe that

$$\sum_{x\in X} P(x)\,\varepsilon_x(y) = 0 \text{ and } \sum_{y\in Y} P(y)\varepsilon_x(y) = 0 \quad (1.3)$$

with $\varepsilon_x(y) \ll 1$.Later, Gallager [3] proposed that $P(y)$ need only be an approximation of the output probabilities rather than the output probabilities itself.

Information theory has been very helpful for researching information transfer through noisy communication channels [24] because it offers a flexible mathematical framework ofSuch paths for communication. The concepts of entropy [3], mutual information [3], parametric measures [17, 21, and 23], multivariate normal distribution [18], divergence [5], directed divergence [19, 20, and 22],quantitative-qualitative measure [2, 4],exponential entropy functional [9, and 16]and many others are derived from information theory. Their generalisations [8, 14, and 15] have been used in the areas of pattern recognition and medical diagnostics, among others.According to Shannon [3], communication over a discrete memoryless channel is always possible as long as the channel capacity is non-zero.

The entropy function described by Shannon [12] is given by

$$H_P(X) = -\sum_{x\in X} P(x) lnP(x)$$

The generalisedRènyi entropy can be expressed as follows:

$$H_P(X) = \frac{1}{1-\alpha} \, ln(\textstyle\sum_{x\epsilon X} P^{\alpha}(x)), \;\; \alpha > 0. \; (1.4)$$

It represents the degree of uncertainty about the input alphabet X. Similar to the output alphabet Y level of uncertainty can be expressed as

$$H_P(Y) = \frac{1}{1-\alpha} ln(\textstyle\sum_{x\epsilon X} P^{\alpha}(y)), \;\; \alpha > 0. \qquad (1.5)$$

X and Y are discrete random variables so joint entropy function is given as

$$H_P(X,Y) = \frac{1}{1-\alpha} \, ln(\textstyle\sum_{x\in X} \sum_{y\in Y} P^{\alpha}(x,y)) \; (1.6)$$

The conditional entropy are defined as

$$H_P(X/Y) = \frac{1}{1-\alpha} \, ln(\textstyle\sum_{x\in X} \sum_{y\in Y} P^{\alpha}(x/y)) \; (1.7)$$

$$H_P(Y/X) = \frac{1}{1-\alpha} \, ln(\textstyle\sum_{x\in X} \sum_{y\in Y} P^{\alpha}(y/x))(1.8)$$

The channel capacity C is given by the equation

$$C = \max_{P(x)} I(X,Y) \qquad (1.9)$$

where $I(X,Y)$ is mutual information which was given about X by Y or Y by X and is given by

$$I(X,Y) = H_P(Y) - H_P(Y/X) = H_P(X) - H_P(X/Y) = I(Y,X) \qquad (1.10)$$

The Shannon entropy function additive property is defined by the fact that the information gain function is logarithmic.

The weighted entropy was proposed by Munteanu and Tarniceriu [7] which form is
$H_{Pc}(X) = \frac{\gamma}{1-\alpha}\ ln \sum_{x\epsilon X} P^{\alpha}(x) + \delta \sum_{x\epsilon X} P(x)c(x)$ (1.11)

where $c(x)$ denotes the weights assigned to the input alphabet X, with α and β standing in for arbitrary constants that will be decided by boundary conditions. The weighted variants of (5), (6), (7), and (8) are defined similarly as

$$H_{Pc}(Y) = \frac{\gamma}{1-\alpha}\ ln \sum_{y\in Y} P^{\alpha}(y) + \delta \sum_{y\in Y} P(y)c(y) \quad (1.12)$$

$$H_{Pc}(X,Y) = \frac{\gamma}{1-\alpha}\ ln \sum_{x\in X}\sum_{y\in Y} P^{\alpha}(x,y) + \delta \sum_{x\in X}\sum_{y\in Y} P(x,y)c(x,y) \quad (1.13)$$

$$H_{Pc}(X/Y) = \frac{\gamma}{1-\alpha} ln \sum_{x\epsilon X}\sum_{y\epsilon Y} P^{\alpha}(x/y) + \delta \sum_{x\epsilon X}\sum_{y\epsilon Y} P(x,y)c(x\,/\,y) \quad (1.14)$$

$$H_{Pc}(Y/X) = \frac{\gamma}{1-\alpha} ln \sum_{x\epsilon X}\sum_{y\epsilon Y} P^{\alpha}(y/x) + \delta \sum_{x\epsilon X}\sum_{y\epsilon Y} P(x,y)c(y\,/\,x) \quad (1.15)$$

The channel capacity of weighted entropy is given as

$$C = \max_{P(x)} \bar{I}(X,Y) = \max_{P(x)}\big(H_{Pc}(Y) - H_{Pc}(Y/X)\big) = \max_{P(x)}(H_{Pc}(X) - H_{Pc}(X/Y)) \quad (1.16)$$

2. OUR RESULTS

CHANNEL CAPACITY OF WEIGHTED ENTROPY FUNCTION

The weighted conditional entropy given by (1.15) is

$$H_{Pc}(Y/X) = \frac{\gamma}{1-\alpha} ln \sum_{x\epsilon X}\sum_{y\epsilon Y} P^{\alpha}(y/x) + \delta \sum_{x\epsilon X}\sum_{y\epsilon Y} P(x,y)c(y\,/\,x)$$

If alphabet $x\epsilon X$ is assumed to represent the source codewords and the corresponding weights to be the codeword length, then the constants γ and δ both take values of -1[12] .

We have,

$$H_{Pc}(Y/X) = \frac{-1}{1-\alpha} ln \sum_{x\epsilon X}\sum_{y\epsilon Y} P^{\alpha}(y/x) - \sum_{x\epsilon X}\sum_{y\epsilon Y} P(x,y)c(y\,/\,x)$$

$$H_{Pc}(Y/X) = \frac{-\alpha}{1-\alpha} \sum_{x\epsilon X}\sum_{y\epsilon Y} ln\, P(y/x) - \sum_{x\epsilon X}\sum_{y\epsilon Y} P(x,y)c(y\,/\,x)$$

$$H_{Pc}(Y/X) = \frac{-\alpha}{1-\alpha} \sum_{x\epsilon X}\sum_{y\epsilon Y} ln\, P(y/x) - \sum_{x\epsilon X}\sum_{y\epsilon Y} P(x)P(y/x)c(y\,/\,x)$$

$$= \frac{-\alpha}{1-\alpha} \sum_{x\epsilon X}\sum_{y\epsilon Y} ln\left(P(y)(1-\varepsilon_x(y))\right) - \sum_{x\epsilon X}\sum_{y\epsilon Y} P(x)P(y)(1-\varepsilon_x(y))c(y\,/\,x)$$

Using the formula

$$ln(1-\alpha) = -\alpha - \frac{\alpha^2}{2} - \frac{\alpha^3}{3} \ldots$$

We get

$$H_{Pc}(Y/X) = \frac{-\alpha}{1-\alpha}\left[\sum_{x\epsilon X}\sum_{y\epsilon Y} lnP(y) - \varepsilon_x(y) - \frac{\varepsilon_x^2(y)}{2} - \cdots\cdots\cdots\right] - \sum_{x\in X} P(x)\sum_{y\in Y} P(y)c(y/x) + \sum_{x\in X}\sum_{y\in Y} P(x)P(y)\varepsilon_x(y)c(y/x) \quad (2.1)$$

We have [8], for noisy channels,

$c(y/x) = c(y)$ and $c(x/y) = c(x)$ (2.2)

After some basic changes and using (1.3) and (2.2) in (2.1), we obtain

$$H_{Pc}(Y/X) = \frac{-\alpha}{1-\alpha}\left[\sum_{x\epsilon X}\sum_{y\epsilon Y} lnP(y) - \varepsilon_x(y) - \frac{\varepsilon_x^2(y)}{2} - \cdots\cdots\cdots\right] - \sum_{x\in X}\sum_{y\in Y} P(x)P(y)c(y)$$

The channel capacity of weighted entropy defined as

$$C = \max_{P(x)} \bar{I}(X,Y) = \max_{P(x)}\left(H_{pc}(Y) - H_{Pc}(Y/X)\right)$$

$$C = \max_{P(x)}\left[\sum_{y\in Y} P(y)c(y)(\sum_{x\in X} P(x) - 1) + \frac{\alpha}{\alpha-1}\left(\varepsilon_x(y) + \frac{\varepsilon_x^2(y)}{2}\right)\right] \quad (2.3)$$

This is the required channel capacity of weighted entropy function.where the terms in equation (2.3) are true up to the second order of $\varepsilon_x(y) \ll 1$.

CONCLUSION

Here, we evaluated the equations for the channel capacity of Type I Noisy Channels using the weighted conditional entropy for Non-Shannon entropy. By applying Majani's Type II Noisy Channels specification, comparable results might be achieved [5].

REFERENCES

1. Abdel, K., Ghaffar, R., and Eliece, Mc. (1986): The ultimate limits of information density, Proceedings of the NATO advanced studyinstitute on performance limits in communication theory andPractice, II Ciocco, Italy.

2. Belis, M., and Guiasu, S. (1968): A quantitative-qualitative measure of information in cybernetic systems, IEEE Transactions on Information Theory, vol. 14, no. 4, pp. 593-594.

3. Gallager, R. (1965): A simple derivation of the coding theorem and some applications, IEEE transactions on Information theory. vol. 11 no. 1, pp. 3-18.

4. Kapur, J. N. (1986): New qualitative-quantitative measures of information, National Academy Science letters, vol. 9, issue 2, pp. 51-54.

5. Kullback, S., and Leibler, S. A. (1951): On information and sufficiency, Annals of Mathematical Statistics, 22, pp. 79-86.

6. Majani, E. (1987): A Model for the study of very noisy channels and applications Ph.D. Thesis, California Institute of Technology, Pasadena.

7. Munteanu, V., and Tarniceriu, D. (2007): Transmission of Quantitative- Qualitative Information on Noisy Channels, International Symposium on Signals, Circuits and Systems, pp. 1 – 4.

8. Pal, N. R., and Pal, S. K. (1991): Entropy: A new definitions and its applications, IEEE transactions on systems, man and cybernetics. vol. 21 no. 5, pp. 1260-1270.

9. Pal, N. R., and Pal S. K. (1989): Object background segmentation using a new definition of entropy, Computers and Digital Techniques, IEE Proceedings E . vol. 136 no. 4, pp. 284-295.

10. Reiffen, B. (1963): A note on very voisy channels Information & Control, vol. 6, pp. 126- 130.

11. Sanmin, Lee., Kim, A., and Winick (1994): Very noisy channels, reliability functions, and exponentially optimum codes, IEEE transactions on Information theory. vol. 40 no. 3, pp. 647-661.

12. Shannon, C. E. (1948): A Mathematical theory of communication, Bell System Technical Journal, vol. 2, pp. 623-659.

13. Silverman, R. (1955): On Binary Channels and their Cascades, IRE transactions of Information Theory, pp. 19-27.

14. Srivastava, A. (2012): Some New Information Inequalities Involving *f*- Divergences, Cybernetics and Information Technologies, vol. 12 no. 2, pp. 3 -10.

15. Srivastava, A., and Maheshwari, S. (2016): Decision making in medical investigations using new divergence measures for intuitionistic fuzzy sets, Iranian Journal of Fuzzy systems, vol. 13 no. 1, pp. 25-44

16. Srivastava, A., Singh, A. K., and Maheshwari, S. (2013): Dichotomous Exponential Entropy Functional and Its Applications in Medical Diagnosis, IEEE- ICSC-2013, pp. 21 – 26.

17. Verma, R. K., Dewangan, C. L. and Jha, P. (2012): An Unorthodox Parametric Measures of Information and Corresponding Measures of Fuzzy Information, Int. Jour. of Pure and Appl. Mathematics, Bulgaria, Vol. 76, No. 4, pp. 599-614.

18. Verma, R. K., Dewangan, C. L. and Jha, P. (2012): Some Multivariate Normal Distributions Based on Parametric and Non-parametric Measures of Information Int. Jour. of Math. Sci. &Engg. Appls., Vol. 6 No. 1, pp. 343-353.

19. Verma, R. K., Dewangan, C. L. and Jha, P. (2012): Some Generating Functions for Measures of Probabilistic Entropy and Directed Divergence, Int. Jour. of Pure and Appl. Mathematics, Bulgaria, Vol. 74, No. 1, pp. 21-32.

20. Verma, R. K. and Verma, Babita (2013):A New Approach in Mathematical Theory of Communication (A New Entropy with its Application), Lambert Academic Publishing Edn., pp. 84-85.

21.Verma, R. K. (2023): Information Radius Via Verma Information Measure in Intuitionistic Fuzzy Environment, International Journal of Mathematics Research (IJMR), 2023, Vol. 15(1), pp. 1-8.

22.Verma, R. K. (2023): On Optimality of Entropy Like Functional in Terms of Distance Function, as a book Chapter-2 in Research Highlights in Mathematics and Computer Science (RHMCS), Vol. 7, pp. 10-20, DOI: 10.9734/bpi/rhmcs/v7/18679D.

23.Verma, R. K. (2023): On Optimization Policy for Verma Entropy by Dynamic Programming, Asian Journal of Probability and Statistics (AJPAS), Vol. 21(4), pp. 14-21, DOI: 10.9734/AJPAS/2023/v21i4469.

24. Verma, R. K. (2023): On Optimal Channel Capacity Theorems via Verma Information Measure with Two-Sided Input in Noisy State, Asian Journal of Probability and Statistics (AJPAS), Vol. 22(2), pp. 1-7, DOI:10.9734/AJPAS/2023/v22i2478.