ESL-P-684

ON A SPECIAL CLASS OF WIRETAP CHANNELS

by

S.K. Leung-Yan-Cheong*

ABSTRACT

In this note we examine a special class of wiretap channels and give some useful characterizations of its members.

*Massachusetts Institute of Technology. This work was supported in part by NSF Grants GK-33250 and ENG-10173, in part by the U.S. Air Force Office of Scientific Research, under contract F44620-73-C-0065 at Stanford University, in part by ARPA under Grant ONR/N00014-75-C-1183 and in part by NSF/ENC-76-24447, at M.I.T., Cambridge, Mass. 02139.

1. Introduction

The concept of the wiretap channel was first introduced by Wyner [1]. The model which he proposed is shown in figure 1.

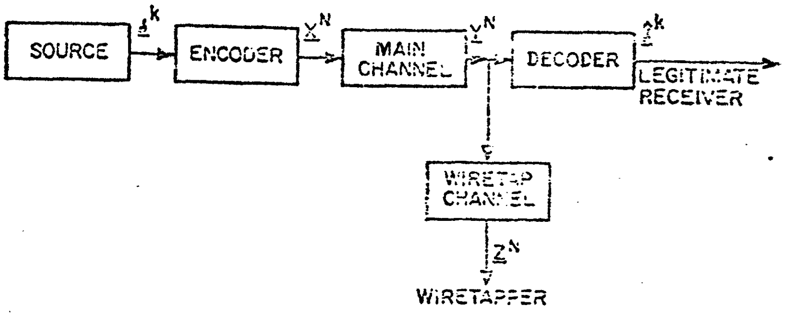

Figure 1. General Wiretap Channel.

It is a form of degraded broadcast channel, with the new idea that one information rate is to be maximized and the other minimized. The object is to maximize the rate of reliable communication from the source to the legitimate receiver, subject to the constraint that the uncertainty of the wiretapper, after he receives his data $\underline{Z}^N$, is no less than some given quantity. The wiretapper knows the encoding scheme used at the transmitter and the decoding scheme used at the legitimate receiver, and is handicapped only by the greater noise present in his received signal. Thus, while the objective is the same as in cryptography, the technique used to achieve privacy is very different.

In this note, we make use of Wyner's basic result for discrete memoryless wiretap channels to analyze a special class of channels.

II. Preliminaries

Referring to figure 1, the source is ergodic and has a finite alphabet. The first k source outputs $\underline{s}^k$ are encoded into an N-vector $\underline{x}^N$ which is the input to the main channel. The legitimate receiver makes an estimate $\hat{\underline{s}}^k$ of $\underline{s}^k$ based on $\underline{y}^N$, the output of the main channel, incurring a block error rate

$$P_e = \Pr\{\underline{s}^k \neq \hat{\underline{s}}^k\} \tag{1}$$

$\underline{y}^N$ is also the input to the wiretap channel and the wiretapper has average residual uncertainty $H(\underline{S}^k|\underline{Z}^N)$ after observing the output $\underline{z}^N$ of the wiretap channel.

We define the fractional equivocation of the wiretapper to be

$$\Delta = H(\underline{S}^k|\underline{Z}^N)/H(\underline{S}^k) \tag{2}$$

and the rate of transmission to the legitimate receiver to be

$$R = H(\underline{S}^k)/N \tag{3}$$

Note that $\Delta = 1$ implies that the wiretapper's a posteriori uncertainty about the source output is equal to his a priori uncertainty. Thus when

$\Delta = 1$, the wiretapper is no better informed after he receives his data than he was before. The pair (R^*, d^*) is said to be achievable if for all $\varepsilon > 0$ there exists an encoder-decoder pair such that

$$R \geq R^* - \varepsilon \tag{4a}$$

$$\Delta \geq d^* - \varepsilon \tag{4b}$$

$$P_e \leq \varepsilon. \tag{4c}$$

Wyner's basic result on the achievable (R,d) region is the following:

Theorem 1 (Wyner)

Let $p_X(\cdot)$ be a probability distribution on the input of the main channel. Define $\rho(R)$, $R \geq 0$ to be the set of p_X such that $I(X;Y) \geq R$. For $0 \leq R \leq C_M$ where C_M is the capacity of the main channel, let

$$\Gamma(R) = \max_{p_X \varepsilon \rho(R)} I(X;Y|Z) \tag{5}$$

Then the set of all achievable (R,d) pairs is given by

$$R^* = \{(R,d) | 0 \leq R \leq C_M,\ 0 \leq d \leq 1,\ Rd \leq \Gamma(R)\} \tag{6}$$

In Section III, we analyze the class of channels for which $\Gamma(R)$ is constant [2].

III. Constant $\Gamma(R)$ Channels.

A useful characterization of constant $\Gamma(R)$ channels is given by the following theorem:

Theorem 2.

$\Gamma(R)$ is constant if and only if p_X^* maximizes $I(X;Y) - I(X;Z)$ where p_X^* is a capacity achieving distribution on the main channel.

Proof:

We first note that we can rewrite $I(X;Y|Z)$ as

$$I(X;Y|Z) = H(X|Z) - H(X|Y,Z) \tag{7}$$

$$= H(X|Z) - H(X|Y) \tag{8}$$

$$= I(X;Y) - I(X;Z) \tag{9}$$

where in (8) we have used the fact that X,Y and Z form a Markov chain.

If p_X^* maximizes $I(X;Y) - I(X;Z)$, then

$$\Gamma(R) = I_{p_X^*}(X;Y) - I_{p_X^*}(X;Z) \tag{10}$$

since $p_X^* \varepsilon \rho(R)$ for $0 \leq R \leq C_M$. Therefore $\Gamma(R)$ is a constant.

Conversely, suppose p_X^* does not maximize $I(X;Y) - I(X;Z)$. Let the maximizing distribution be denoted by p_X', and let $I_{p_X'}(X;Y) = R_1$. Then, it is clear that $\Gamma(R_1) > \Gamma(C_M)$. This shows that $\Gamma(R)$ cannot be a constant.Q.E.D. We now give a necessary condition for $\Gamma(R)$ to be constant for a special class of discrete memoryless wiretap channels.

Theorem 3.

Let the main channel be a discrete memoryless channel (DMC) with K inputs, and the wiretap channel be some other DMC. Suppose that $I(X;Y)$ is maximized at a unique input distribution $p_X^* = (p^*(1), p^*(2),\ldots,p^*(K))$ where all the components of p_X^* are strictly positive. Then a necessary condition for $\Gamma(R)$ to be constant is that p_X^* should be a maximizing distribution for $I(X;Z)$.

Proof.

First we note that $I(X;Y)$ and $I(X;Z)$ are both concave functions of the input probability assignment p_X to the main channel [3, theorem 4.4.2].

We handle the equality constraint $\sum_{i=1}^{K} p(i) = 1$ by substituting $1-\sum_{i=1}^{K-1} p(i)$ for $p(K)$ in the expressions for $I(X;Y)$ and $I(X;Z)$. Thus we can consider maximizing $I(X;Y)$ and $I(X;Z)$ which are now functions of K-1 variables subject only to the inequality constraints $p(i) \geq 0$.

By hypothesis, $I(X;Y)$ has a maximum at p_X^* and $p_X^* > 0$. Therefore

$$\left.\frac{\partial I(X;Y)}{\partial p(i)}\right|_{p_X^*} = 0, \quad 1 \leq i \leq K-1 \tag{11}$$

Now assume that p_X^* does not maximize $I(X;Z)$. Then there exists at least one $j\varepsilon[1, K-1]$ such that

$$\left.\frac{\partial I(X;Z)}{\partial p(j)}\right|_{p_X^*} \neq 0 \tag{12}$$

Thus, by moving away from p_X^* along the direction of $p(j)$[†], the difference between

[†] Note that this is always possible since we assumed that all the components of p_X^* are strictly positive.

I(X;Y) and I(X;Z) can be made to increase (at least initially). Therefore, p_X^* does not maximize I(X;Y)-I(X;Z) and by theorem 2, this implies that $\Gamma(R)$ is not constant.

Q.E.D.

Remark:

Theorem 3 can be extended in a straightforward manner to cover the case where I(X;Y) is maximized at non-unique but strictly positive input distributions.

An example of a wiretap channel for which $\Gamma(R)$ is not constant is one in which the main and wiretap channels are as shown in figures 2(a) and 2(b) respectively.

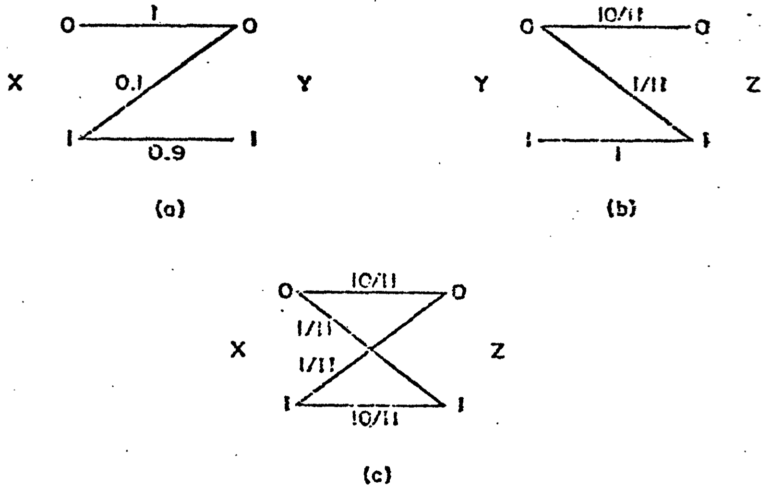

Figure 2.

(a) Main Channel

(b) Wiretap Channel

(c) Cascade of main and wiretap Channels.

The cascade of these two channels is equivalent to a binary symmetric channel (BSC) with a crossover probability of 1/11 as shown in figure 2(c). The maximum of $I(X;Y)$ occurs when $\Pr\{X=0\} = 0.54$. On the other hand, because of the symmetry of the channel from X to Z, $I(X;Z)$ is maximized when $\Pr\{X=0\} = 0.5$. Thus, by theorem 3, $\Gamma(R)$ is not constant.

We now prove a result which gives a sufficient condition for a wiretap channel to have a constant $\Gamma(R)$.

Theorem 4.

Suppose that $I(X;Y)$ and $I(X;Z)$ are simultaneously maximized by an input distribution p_X^*. Then $\Gamma(R)$ is constant (and equals $C_M - C_{MW}$), where C_{MW} is the capacity of the cascade channel.

Proof.

p_X^* achieves the maxima of $I(X;Y)$ and $I(X;Z)$ and hence corresponds to a stationary point of $I(X;Y) - I(X;Z) = I(X;Y|Z)$.

In lemma 1 below, it is shown that $I(X;Y|Z)$ is a concave function of the input probability distribution. Therefore, the stationary point p_X^* also maximizes $I(X;Y) - I(X;Z)$. Finally, use of theorem 2 proves that $\Gamma(R)$ is constant.

$$\text{Also } \Gamma(R) = I_{p_X^*}(X;Y) - I_{p_X^*}(X;Z) \tag{13}$$

$$= C_M - C_{MW}. \qquad \text{Q.E.D.} \tag{14}$$

Corollary: Suppose that the main channel is a symmetric DMC [3, p. 94] and the cascade of the main and wiretap channels is also a symmetric DMC. Then $\Gamma(R)$ is constant $(=C_M - C_{MW})$.

Proof.

It is well known [3, theorem 4.5.2] that for a symmetric DMC, capacity is achieved by using the inputs with equal probability. This fact, in conjuction with theorem 4, proves that $\Gamma(R)$ is constant. Since C_M and C_{MW} can be easily evaluated in this case, $\Gamma(R)$ is readily found.

Remark: The corollary can be used to yield $\Gamma(R)$ for many wiretap channels, e.g. (1) main and wiretap channels are BSC's (2) main channel is a BSC and wiretap channel is a binary erasure channel.

We conclude this section by proving the following lemma.

Lemma 1.

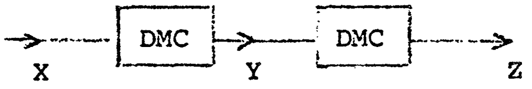

Figure 3. Cascade of two DMC's

Let X, Y and Z denote the inputs and outputs of two cascaded DMC's as illustrated in figure 3. Then $I(X; Y|Z)$ is a concave function of the input probability assignment.

Proof.

Suppose $\underline{Q}_0$ and $\underline{Q}_1$ are two arbitrary probability distributions on X. Let

$$I_{\underline{Q}_0}(X; Y|Z) = I_0 \tag{15}$$

and

$$I_{\underline{Q}_1}(X; Y|Z) = I_1. \tag{16}$$

Let $\theta\varepsilon(0,1)$ be arbitrary and $\underline{Q} = \theta\underline{Q}_0 + (1-\theta)\underline{Q}_1$ and denote the corresponding mutual information by I. Then we want to show that

$$\theta I_0 + (1-\theta) I_1 \leq I. \tag{17}$$

Consider an auxiliary random variable U which takes on values in $\{0,1\}$ with probabilities

$$\Pr\{U=0\} = \theta, \qquad \Pr\{U=1\} = 1-\theta. \tag{18}$$

We now take $\underline{Q}_0$ and $\underline{Q}_1$ to be conditional probabilities conditioned on U: if U=0, the input probability distribution is $\underline{Q}_0$ and if U=1, the input distribution is $\underline{Q}_1$. So U,X,Y and Z form a Markov chain and in particular the following equalities hold:

$$p(z|y,x,u) = p(z|y) \tag{19a}$$

$$p(y|x,u) = p(y|x) \tag{19b}$$

$$p(z|x,u) = p(z|x) \tag{19c}$$

$$p(z|x,y) = p(z|y) \tag{19d}$$

$$\text{Now } p(y|z,u,x) = \frac{p(z|y,x,u)p(y|x,u)}{p(z|x,u)} \tag{20}$$

$$= \frac{p(z|y)p(y|x)}{p(z|x)} \quad \text{using (19a,b,c)} \tag{21}$$

$$\text{Also } p(y|z,x) = \frac{p(z|y,x)p(y|x)}{p(z|x)} \tag{22}$$

$$= \frac{p(z|y)p(y|x)}{p(z|x)} \quad \text{using (19d)} \tag{23}$$

From (21) and (23) we conclude that

$$p(y|z,u,x) = p(y|z,x). \tag{24}$$

Hence

$$H(Y|Z,U,X) = H(Y|Z,X). \tag{25}$$

The left hand side of (17) is $I(X;Y|Z,U)$ and the right hand side of (17) is $I(X;Y|Z)$. So it remains to show that

$$I(X;Y|Z,U) \leq I(X;Y|Z). \tag{26}$$

$$I(X;Y|Z,U) = H(Y|Z,U) - H(Y|Z,U,X) \tag{27}$$

$H(Y|Z,U) \leq H(Y|Z)$ since conditioning can only decrease entropy. (28)

Substituting (25) and (28) in (27) we obtain

$$I(X;Y|Z,U) \leq H(Y|Z) - H(Y|Z,X) \tag{29}$$

$$= I(X;Y|Z). \tag{30}$$

Q.E.D.

IV. Conclusion

We have shown that the complete achievable (R,d) region for "symmetric" wiretap channels can be easily evaluated explicitly. Necessary and sufficient conditions for the function $\Gamma(R)$ to be constant have been derived and should prove useful in evaluating the set of all achievable (R,d) pairs for many channels.

References

[1]. A.D. Wyner, "The Wire-tap Channel",Bell System Technical Journal, Vol. 54, pp. 1355-1387, October 1975.

[2]. S.K. Leung-Yan-Cheong, "Multi-User and Wiretap Channels Including Feedback", Technical Report No. 6603-2, Center for Systems Research, Stanford University, Stanford, California 94305.

[3]. R.G. Gallager, Information Theory and Reliable Communication, Wiley, New York, 1968.