Multi-User Information Theory Dec 05, 2011

Lecture 1

Lecturer:Dr. Haim Permuter *Scribe: Ziv Goldfeld*

I. Relay Channel - Channel Model

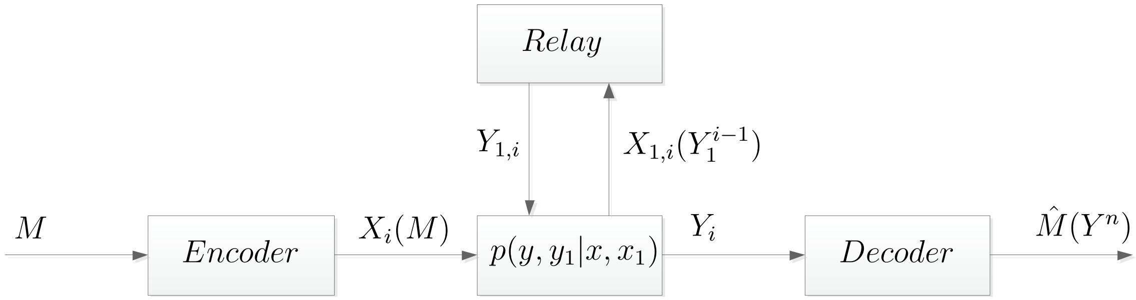

Fig. 1. Relay channel. The user, aided by the relay, sends the message M to the receiver.

The relay channel is a channel in which there is one sender and one receiver with an intermediate node that acts as relay to help the communication. The channel model consists of four finite sets $\mathcal{X}, \mathcal{X}_1, \mathcal{Y}, \mathcal{Y}_1$, and a collection of conditional pmfs $p(y, y_1|x, x_1)$ on $\mathcal{X} \times \mathcal{X}_1 \times \mathcal{Y} \times \mathcal{Y}_1$. The sender X wishes to send a message M to receiver Y with the help of the relay (X_1, Y_1).

The meaning of the channel transition pmf $p(y, y_1|x, x_1)$, is that the channel output at time $i \in \{1, 2, \ldots, n\}$, i.e., $(Y_i, Y_{1,i})$, depends only the channel input at time i, i.e., $(X_i, X_{1,i})$. Namely, this property can be written as

$$p(y_i, y_{1,i}|x_i, x_{1,i}, m, y^{i-1}, y_1^{i-1}, x^{i-1}, x_1^{i-1}) = p(y, y_1|x, x_1) \quad (1)$$

Thus we have the Markov relation $(Y_i, Y_{1,i}) - (X_i, X_{1,i}) - (M, Y^{i-1}, Y_1^{i-1}, X^{i-1}, X_1^{i-1})$.

A $(2^{nR}, n)$ code for the relay channel consists of:

Definition 1 *(The message)*

A Message M, which is an integer of the set $\{1, 2, \ldots, 2^{nR}\}$. The message M is to be sent by the user, with the help of the relay, to the receiver.

Definition 2 *(The encoder)*

An encoder, which is described by a function f:

$$f : \{1, 2, \ldots, 2^{nR}\} \to \mathcal{X}^n \tag{2}$$

i.e., the encoding function assigns a codeword $x^n(m)$, for each message $m \in \{1, 2, \ldots, 2^{nR}\}$.

Definition 3 *(The relay)*

A relay encoder, which is described by a set of n functions $\{f_{1,i}\}_{i=1}^{n}$, which for every time $i \in \{1, 2, \ldots, n\}$:

$$f_{1,i} : \mathcal{Y}_1^{i-1} \to \mathcal{X}_1 \tag{3}$$

i.e., the relay encoder assigns at each time $i \in \{1, 2, \ldots, n\}$ a symbol $x_{1,i}(y_1^{i-1})$ to each past received sequence y_1^{i-1}.

Definition 4 *(The decoder)*

A decoder, which is described by a function g:

$$g : \mathcal{Y} \to \{1, 2, \ldots, 2^{nR}\} \tag{4}$$

i.e., the decoder assigns a message $\hat{m} \in \{1, 2, \ldots, 2^{nR}\}$ (or declares an error) for each received sequence y^n.

Definition 5 *(Probability of Error)*

We define the probability of error for the relay channel as:

$$P_e^{(n)} = Pr\{m \neq \hat{m}(Y^n) | m \text{ was sent}\} \tag{5}$$

Definition 6 *(Achievable Rate)*

A rate R is called achievable if there exists a sequence of $(2^{nR}, n)$ codes such that $P_e^{(n)} \to 0$ as $n \to \infty$.

Definition 7 *(Capacity)*
The capacity of the relay channel is the supremum of all achievable rates R.

The main difficulty regarding the relay channel is the fact that the relay does not know the message m that is to be sent by the user, but nevertheless is required to help the user to send that message using the information that is available in the relay.
Moreover, until this day no solution for the general relay channel was presented and the capacity have not been found. However, lower and upper bounds on the capacity of the relay channel does exists and shall be presented now.
For some families of channels, such as the degraded relay channel and the semi-deterministic relay channel, the upper and lower bound on the capacity coincide, thus yielding the capacity for those settings.

II. Upper Bound

Theorem 1 *(Upper bound on the capacity of the relay channel)*
If R is an achievable rate, then

$$R \leq \max_{p(x,x_1)} \min \left\{ I(X, X_1; Y), I(X; Y, Y_1 | X_1) \right\} \tag{6}$$

Proof: Given an achievable rate R we need to show that there exists joint distribution of the form $P(x_1, x)$ such that,

$$R \leq \min \left\{ I(X, X_1; Y), I(X; Y, Y_1 | X_1) \right\} \tag{7}$$

Since R is an achievable rate, there exists an $(2^{nR}, n)$ code with a probability of error $P_e^{(n)}$ arbitrarily small. By Fano's inequality,

$$H(M|Y^n) \leq nR \cdot P_e^{(n)} + H(P_e^{(n)}) \triangleq n\varepsilon_n, \tag{8}$$

and it is clear that $\varepsilon_n \to 0$ as $P_e^{(n)} \to \infty$.

We can now bound the rate R_1 as

$$
\begin{aligned}
nR &= H(M) \\
&= H(M) + H(M|Y^n) - H(M|Y^n) \\
&= I(M;Y^n) + H(M|Y^n) \\
&\overset{(a)}{\leq} I(M_1;Y^n) + n\varepsilon_n \\
&\overset{(b)}{=} \sum_{i=1}^{n} I(M_1;Y_i|Y^{i-1}) + n\varepsilon_n \\
&\overset{(c)}{=} \sum_{i=1}^{n} I(M_1, X_i;Y_i|Y^{i-1}) + n\varepsilon_n \\
&\overset{(d)}{\leq} \sum_{i=1}^{n} I(M_1, X_i, X_{1,i};Y_i|Y^{i-1}) + n\varepsilon_n \\
&= \sum_{i=1}^{n} \left[H(Y_i|Y^{i-1}) - H(Y_i|M, X_i, X_{1,i}, Y^{i-1})\right] + n\varepsilon_n \\
&\overset{(e)}{=} \sum_{i=1}^{n} \left[H(Y_i|Y^{i-1}) - H(Y_i|X_i, X_{1,i})\right] + n\varepsilon_n \\
&\overset{(f)}{\leq} \sum_{i=1}^{n} \left[H(Y_i) - H(Y_i|X_i, X_{1,i})\right] + n\varepsilon_n \\
&= \sum_{i=1}^{n} I(X_i, X_{1,i};Y_i) + n\varepsilon_n
\end{aligned}
$$

where

(a) follows from Fano's inequality.

(b) follows from the mutual information chain rule.

(c) follows from the fact that X if a function of the message M.

(d) follows from the fact that conditioning reduces entropy.

(e) follows from the fact that the channel output at time i depends only on the pair $(X_i, X_{1,i})$.

(f) follows from the fact that conditioning reduces entropy.

So we have:

$$R \leq \frac{1}{n} \sum_{i=1}^{n} I(X_i, X_{1,i}; Y_i) + \varepsilon_n \tag{9}$$

Another way to bound R is

$$\begin{aligned}
nR &= H(M) \\
&= H(M) + H(M|Y^n) - H(M|Y^n) \\
&= I(M;Y^n) + H(M|Y^n) \\
&\overset{(a)}{\leq} I(M_1;Y^n) + n\varepsilon_n \\
&\overset{(b)}{\leq} I(M_1;Y^n, Y_1^n) + n\varepsilon_n \\
&\overset{(c)}{=} \sum_{i=1}^{n} I(M_1;Y_i, Y_{1,i}|Y^{i-1}, Y_1^{i-1}) + n\varepsilon_n \\
&\overset{(d)}{=} \sum_{i=1}^{n} I(M_1, X_i;Y_i, Y_{1,i}|X_{1,i}, Y^{i-1}, Y_1^{i-1}) + n\varepsilon_n \\
&= \sum_{i=1}^{n} \left[H(Y_i, Y_{1,i}|X_{1,i}, Y^{i-1}, Y_1^{i-1}) - H(Y_i, Y_{1,i}|M, X_i, X_{1,i}, Y^{i-1}, Y_1^{i-1})\right] + n\varepsilon_n \\
&\overset{(e)}{=} \sum_{i=1}^{n} \left[H(Y_i, Y_{1,i}|X_{1,i}, Y^{i-1}, Y_1^{i-1}) - H(Y_i, Y_{1,i}|X_i, X_{1,i})\right] + n\varepsilon_n \\
&\overset{(f)}{\leq} \sum_{i=1}^{n} \left[H(Y_i, Y_{1,i}|X_{1,i}) - H(Y_i, Y_{1,i}|X_i, X_{1,i})\right] + n\varepsilon_n \\
&= \sum_{i=1}^{n} I(X_i;Y_i, Y_{1,i}|X_{1,i}) + n\varepsilon_n
\end{aligned}$$

where

(a) follows from Fano's inequality.

(b) follows the fact that conditioning reduces entropy.

(c) follows from the mutual information chain rule.

(d) follows from the fact that X_i if a function of the message M and $X_{1,i}$ is a function

of Y_1^{i-1}.
(e) follows from the fact that the channel output at time i depends only on the pair $(X_i, X_{1,i})$.
(f) follows from the fact that conditioning reduces entropy.

So we have:

$$R \leq \frac{1}{n}\sum_{i=1}^{n} I(X_i; Y_i, Y_{1,i}|X_{1,i}) + \varepsilon_n \tag{10}$$

Combining (9) and (10) we get

$$\begin{aligned}
R &\leq \min\Big\{\frac{1}{n}\sum_{i=1}^{n} I(X_i, X_{1,i}; Y_i), \frac{1}{n}\sum_{i=1}^{n} I(X_i; Y_i, Y_{1,i}|X_{1,i})\Big\} + \varepsilon_n \\
&\overset{(a)}{\leq} \max_{\{p(x_i,x_{1,i})\}_{i=1}^{n}} \min\Big\{\frac{1}{n}\sum_{i=1}^{n} I(X_i, X_{1,i}; Y_i), \frac{1}{n}\sum_{i=1}^{n} I(X_i; Y_i, Y_{1,i}|X_{1,i})\Big\} + \varepsilon_n \\
&\overset{(b)}{\leq} \frac{1}{n}\sum_{i=1}^{n} \max_{p(x_i,x_{1,i})} \min\Big\{I(X_i, X_{1,i}; Y_i), I(X_i; Y_i, Y_{1,i}|X_{1,i})\Big\} + \varepsilon_n \\
&= n \cdot \frac{1}{n} \max_{p(x_i,x_{1,i})} \min\Big\{I(X_i, X_{1,i}; Y_i), I(X_i; Y_i, Y_{1,i}|X_{1,i})\Big\} + \varepsilon_n
\end{aligned}$$

where
(a) follows by bounding each element in the sums, for each $i \in \{1, 2, \ldots, n\}$ by the maximum it gets over all possible distributions $p(x_i, x_{1,i})$.
(b) follows from the stationarity of the channel (the channel does not depend on the time index i), thus the maximum and the minimum are the same for all $i \in \{1, 2, \ldots, n\}$.

Therefore, finally, we get

$$R \leq \max_{p(x,x_1)} \min\Big\{I(X, X_1; Y), I(X; Y, Y_1|X_1)\Big\} \tag{11}$$

■

In the next lecture we will proceed our analysis of the relay channel and present lower bounds on its capacity.

REFERENCES

[1] T. M. Cover and J. A. Thomas, *‘Elements of Information Theory.‘*. Wiley, New York, 2nd edition 2006

[2] Abbas El Gamal, Young-Han Kim, *’Lecture Notes on Network Information Theory’*.