EE479: Multiuser Digital Transmission Systems - Lecture 1: Binary Symmetric Channel

John L. Fan

September 22, 1999

1. Introduction

The theme of this course is multiuser detection, in which the basic problem is to consider how to detect when multiple users share a single channel or interfere with each other. In this lecture, the simple case of a binary symmetric channel (BSC) is considered, which serves as a very simple model for understanding the problem in more general situations.

Some results from information theory [1] will be useful for understanding the potential bene…ts of coding and multiuser detection for a situation with interfering users. In particular, the motivation for computing the capacity of the channel is Shannon's Noisy Channel Coding Theorem, which says that there exist error-control codes such that information can be transmitted across this channel at any rate less than or equal to the capacity, with arbitrarily low probability of error. (Conversely, any class of codes which brings the probability of error arbitrarily close to zero must have rate less than the capacity.) In other words, for a …xed amount of noise, the capacity describes the maximum amount of data which can be transmitted along that channel per unit time. Alternately, this relationship can be described by …xing the rate of the code, and looking for the maximum amount of noise that an optimally designed code can tolerate. This is known as the Shannon limit. The proof of Shannon's theorem relies on a random coding argument, which argues that among the codes consisting of randomly chosen codewords, there must exist one which will have good performance, and in fact, it is the case that as block lengths go to in…nity, a randomly chosen code will give near-capacity performance. The problem with these random codes, however,

is the enormous complexity involved in maximum likelihood decoding of these codes.

The theory of error-correcting codes [4] describes the construction of codes that have structure that allows for decoding at reasonable complexities. In particular, there has been extensive use of codes such as convolutional codes and Reed-Solomon codes (or a combination of these codes, known as concatenated coding) that perform reasonably well for many applications. On the other hand, there has traditionally been a gap between the performance and the theoretical limits, due to limits on the complexity of the decoders. In the past few years, since the discovery of turbo codes in 1993, there has been a ‡urry of exciting experimental results and theoretical papers which give error-correcting codes and decoding techniques which achieving performance extremely close to the Shannon capacity. These codes, which include turbo codes and Low Density Parity Check (LDPC) codes, use iterative decoding techniques which involve the passing of probabilistic information. Using these codes, which involve substantial but not infeasible complexity, the full capacity of the channel can be achieved in the single user case.

In the multiuser case, however, it is much more di¢cult to compute the capacity than in the single user case, and the means for achieving these capacities are not always known. The development of decoding techniques for multiuser detection which achieve capacity is an open research question that will have many practical rami…cations.

2. Binary symmetric channel

Suppose that x_i is a binary signal, and that the received signal y_i is also a binary signal which corrupted by some binary noise n_i. The index i indicates that the time of the signal, where the channel allows the transmission of data at discrete time instants.

$$y_i = x_i \copyright n_i$$

It should be noted that x, y, and n belong to $f0; 1g$ and the addition is performed modulo 2. This channel is known as the binary symmetric channel, with the noise n having probability p of being 1 and probability $(1 ¡ p)$ of being 0.

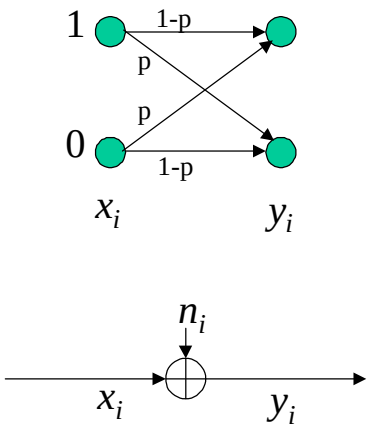

Figure 2.1: Binary Symmetric Channel

2.1. Capacity

The entropy $H(X) = \sum p(x) \log_2 \frac{1}{p(x)}$ denotes the amount of information that is contained in the variable X, and is simply a function of the distribution $p(x)$ on the values that the variable x can take on. In general, the capacity of a discrete memoryless channel (with a discrete input X and discrete output Y) is de…ned as the maximum of the mutual information over all possible input distributions $p(x)$:

$$\begin{aligned} C &= \max_{p(x)} I(X;Y) \\ &= \max_{p(x)} H(Y) - H(Y|X) \end{aligned}$$

In the case of the binary symmetric channel, the capacity is

$$\begin{aligned} C &= \max_{p(x)} H(Y) - H(p) \\ &= 1 - H(p) \end{aligned}$$

where $H(p) = p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}$, and the maximizing distribution on X is the uniform distribution where it has probability 0.5 of being 0 or 1.

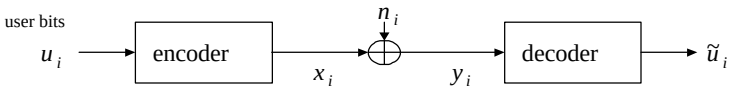

Figure 2.2: Encoding and decoding

2.2. Error-correcting code

To achieve this capacity, it is necessary to use a code. It should be noted that the capacity does not refer to a single use of the channel, but is rather an asymptotic result as the block length of the code increases towards in…nity.

As a very simple example of an error-correcting code, consider the repetition code in which the sequence is divided into blocks of 3 bits, and every three bits are identical. In other words, for all i, we have

$$x_{3i} = x_{3i+1} = x_{3i+2}:$$

It is easy to see that this code encodes a single message bit every three bits, for a coding rate of $\frac{1}{3}$, and this code corrects a single bit error in each block of 3 bits, by using majority voting to determine the correct bit. If the raw bit error rate is p and the errors occur independently, then the probability of an error in a block is given by the part of the binomial expansion

$$P_{error} = 3p^2(1 - p) + p^3;$$

corresponding to having 2 or more errors in a block.

Then suppose that the raw error rate is p = 0:15. Then this code with rate 0:33 lowers the error rate to 0:0607. On the other hand, at this error rate p = 0:15, the capacity is 1 - H (p) = 0:3902. Shannon's theorem then says that it is possible to design a code (using much longer block length) that brings the error rate as close to 0 as desired, and that it is possible to …nd such a code with rate close to the theoretical capacity 0:39.

More generally, error-correcting codes can be constructed using a parity check matrix P with M rows and N columns.

$$P \cdot x = 0$$

$$\sum_{j=1}^{N} P_{ij} x_j = 0 \text{ for all } i \in \{1, \ldots, M\}$$

For example, the repetition code described above can be described using the following parity check matrix

$$P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

so that the parity check equation,

$$P \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

gives $x_1 \oplus x_2 = 0$ and $x_2 \oplus x_3 = 0$, so that x_1, x_2 and x_3 must be equal, giving the repetition code.

From Shannon's theorem, it is known that choosing P randomly with rate $\frac{N-M}{N}$ less than capacity will give a code that gives a probability of error that approach zero as the block length N increases. This assumes that a maximum likelihood decoding of the code is possible—a brute force decoding would involve comparing the received vector y with every possible codeword (there are 2^{N-M} of them).

Low-density parity check (LDPC) codes [2] are error-correcting codes de…ned by a binary parity check matrix that is sparse, meaning that it has a relatively small number of 1's, on the order of 3 or 4 per column. It turns out that there exist iterative decoding methods that can decode LDPC codes with relatively low complexity.

3. Multiuser detection

In this section, three situations involving multiple users are considered and analyzed. These scenarios turn out to be rather trivial in the case of a binary symmetric channel, but give some basic insight into the types of multiuser detection problems. Considering these scenarios for other channels, such as channels with Gaussian noise and channels with memory, will be the topic of further lectures.

The general multiuser situation can be described as follows

$$y_i = Hx_i + n_i$$

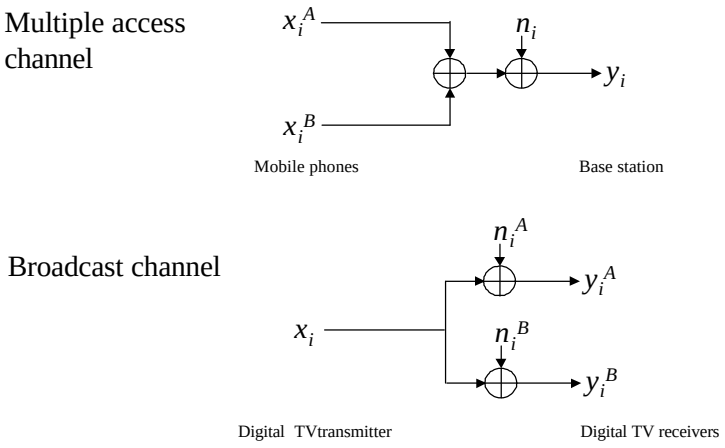

Figure 3.1: Multiple access and broadcast channels

where x_i, y_i and n_i are vectors corresponding to the transmitted signal, the received signal and the noise at time i, respectively. The size of the vectors determines the number of transmitters and the number of receivers in this multiuser system.

3.1. Multiple access channel

In this scenario, multiple users are trying to talk to a single receiver. An example would be where multiple mobile wireless handsets are sending information to a single base station.

$$y_i = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_i^A \\ x_i^B \end{bmatrix} + n_i$$
$$= x_i^A + x_i^B + n_i$$

A simple solution is to have the two users take turns talking. This is known as time-division multiple access (TDMA). If each user uses half of the time, then each user can transmit at a rate which is equal to half the capacity of the channel,

in other words $\frac{1}{2}C = \frac{1}{2}(1 - H(p))$. In general, the channel can be shared at any proportion t:

$$\begin{aligned} r^A &= tC \\ r^B &= (1-t)C \end{aligned}$$

On the other hand, another way to divide the channel is as follows: Let user A transmit in the following pattern, where x_i^A is the encoded bits for user A.

$$x_1^A, x_1^A, x_2^A, x_2^A, x_3^A, x_3^A$$

Meanwhile encode user B's signal using the following pattern

$$x_1^B, 0, x_2^B, 0, x_3^B, 0, \ldots$$

Then it is possible to detect user B by adding together adjacent samples $y_{2i} + y_{2i+1}$. Note that it is assumed that the receiver also knows the timing of the two signal, so that the receiver can then obtain some values for x_i^B (which may possibly have some noise). The receiver then applies the decoder for user B to clean up the signal. Similarly, by looking at the instants when user B does not transmit, the receiver is able to decode user A as well. This separation scheme can be thought of as a code, so that this very simple example falls into the category of code-division multiple access (CDMA).

The advantage of this code-division over time-division is in the case where user B does not signal at all. Then user A can be better decoded, since each bit is transmitted twice, so that there is extra protection against errors. In the time division situations, half of the available slots of signalling would be left empty. Generally speaking, an advantage of code division over time division is that if users do not signal, then the other users improve their performance by automatically utilizing the extra bandwidth.

3.2. Broadcast channel

The opposite of the multiple access channel is the broadcast channel, where a single user transmits to two receivers. This is the situation with digital satellite television, where the transmitter broadcasts a single message to all receivers. Since different receivers may want to view different channels, the receivers must then decode the signal so as to obtain the desired part of the broadcast.

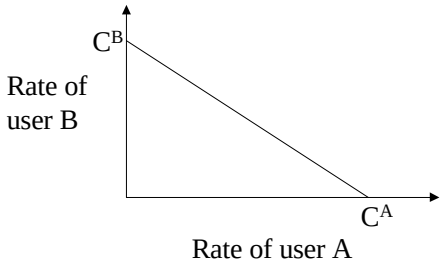

Figure 3.2: Achievable rate region

$$\begin{bmatrix} y_i^A \\ y_i^B \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} x_i + \begin{bmatrix} n_i^A \\ n_i^B \end{bmatrix}$$

Using time division multiplexing (TDM), it is possible for the transmitter to take turns transmitting to the two receivers, or in the case of a television broadcast, the transmitter alternates between broadcasting two programs. Since the two receivers have di¤erent noise levels, the capacities of the channels are di¤erent. The trade-o¤s that can be made by using time division are shown in Figure 3.2. In particular, time division multiplexing yields rates which are a linear combination of the respective capacities

$$\begin{aligned} r^A &= t\,C^A \\ r^B &= (1-t)\,C^B \end{aligned}$$

Notice that to make sure that the same rate is achieved by both users, the transmitter should spend more time transmitting to the weaker user.

$$\begin{aligned} t\,C^A &= (1-t)\,C^B \\ t &= \frac{C^B}{C^A + C^B} \end{aligned}$$

3.3. Interference

Finally, suppose that there are two separate users, each processing its own stream of binary data along a binary symmetric channel, and consider what happens if

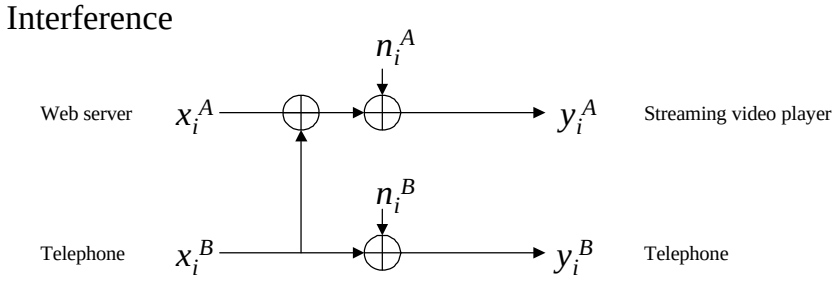

Figure 3.3:

one user's signal interferes with the other user's signal.

$$\begin{bmatrix} y_i^A \\ y_i^B \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_i^A \\ x_i^B \end{bmatrix} + \begin{bmatrix} n_i^A \\ n_i^B \end{bmatrix}$$

For example, imagine a home in the near future where one person is watching television using a set-top box that takes streaming video from a web server, while anohter person is talking on a voice-over-IP telephone. Suppose that all data to the home is sent through the same "pipe" (which could be a cable modem, a digital subscriber loop (DSL) or wireless), and suppose that this future system has been very badly designed so that whenever a phone call comes in the home, the telephone signal interferes with the video signal.

Following the discussion of the binary symmetric channel, we represent this scenario by the highly simpli…ed (and inaccurate) model where all signals are represented by binary data and su¤er from binary noise. First it is assumed that the transmitted signals for user A and user B are independent, so that user A sees user B's signal as additive noise. Then user B has capacity $C^B = 1 - H\ p^B$ as before, but if user A only receives y^A, and knows nothing about user B's signal, then the interference from x^B makes it impossible to gain any knowledge from the signal y^A, so that the capacity C^A is 0.

On the other hand, if user A has access to both y^A and y^B, then user A can …rst decode user B signal from y^B, and then obtain a perfect value for x^B. Then user A can cancel out this signal to remove the interference and thereby decode x^A perfectly, so with multiuser detection the User A's capacity becomes $C^A = 1 - H\ p^A$.

	rate A	rate B $_3$ ´
User A knows y^A	0 $_3$ ´	$C^B = 1 ¡ H_3 p^B$ ´
User A knows y^A and y^B	$C^A = 1 ¡ H\ p^A$	$C^B = 1 ¡ H\ p^B$

It should be noted that when an analogous situation is drawn for other types of channel, such as Gaussian noise channels, it becomes possible (under certain assumptions) to also perform detection even when user A only knows y^A. Computing the capacities in multiuser situations shows a potential for increased transmission rates even in the presence of severe interference. Finding low complexity techniques to achieve capacity on interference channels is an active and promising area of investigation.

References

[1] T.M. Cover and J.A. Thomas, Elements of Information Theory, John Wiley & Sons, 1991.

[2] D.J.C. MacKay, “Good Error-Correcting Codes Based on Very Sparse Matrices,” IEEE Trans. Inform. Theory, Vol. 45, No. 2, March 1999, pp. 399-431.

[3] S. Verdu, Multiuser Detection, Cambridge University Press, 1998.

[4] S.B. Wicker, Error Control Systems for Digital Communications and Storage, Prentice-Hall, 1995.