RFI Channels

R. J. McEliece

Communications Systems Research Section

We present a class of channel models exhibiting varying burst error severity much like channels encountered in practice. We make an information-theoretic analysis of these channel models, and draw some conclusions that may aid in the design of coded communication systems for realistic noisy channels.

I. Introduction

Most of the published research in information theory deals with memoryless channels, whereas most naturally occurring communication channels exhibit at least some degree of bursti-ness, in many cases caused by radio frequency interference (RFI). For example, optical communication with direct detec-tion of photons (Ref. 1), spread-spectrum communication in the presence of hostile jamming (Ref. 2), and communication in the presence of friendly radar transmission (Ref. 3) all lead to channel models in which there are periodic bursts of poor data quality. In this article we shall attempt to model these complicated channels with a class of channels we call "RFI channels." The basic idea behind these models, which we will develop in later sections, is that the channel noise severity is required to remain constant over blocks of n transmitted symbols. However, the channel noise severity may change between one block of n symbols and the next.

Although much further work in this area remains to be done, we are able to draw certain conclusions from this class of models that may prove useful in practical situations. Infor-mally, our main conclusion is that the memory length n should be exploited to determine the noise severity within that block — this is a kind of "soft decision" information; once the noise severity has been estimated, the best strategy is to use n-fold coded interleaving to combat the noise.

II. The Channel Models

Consider the following model for a discrete channel ζ with memory. We start with a finite collection of discrete memory-less channels, $\zeta_1, \zeta_2, \cdots \zeta_k$, each with the same input alpha-bet A, and output alphabet B. When a sequence of letters x_1, $x_2, \cdots$ from A is to be transmitted over ζ, each block of n consecutive letters is sent over one of the auxiliary channels ζ_k, which is selected by an external random variable Z, which takes values in the set $\{1, 2, \cdots, k\}$. If, for example, the ζ_k's are all binary symmetric channels with differing raw bit-error probabilities, the overall channel ζ will be characterized by phased bursts of errors of varying severity.

We consider also another channel $\bar{\zeta}$. This channel is exactly the same as ζ except that it provides to the receiver the index k of the discrete memoryless channel selected by Z.

Our main results are these. First, the capacity of $\bar{\zeta}$ is independent of n, the burst length. We denote this capacity by $\bar{C}$. Second, the capacity of ζ does depend on n, is always less

than $\overline{C}$, and if we denote the capacity of ζ by C_n, we have $\lim_{n\to\infty} C_n = \overline{C}$.

Our results follow fairly easily from calculations with mutual information and entropy. Both channels ζ and $\overline{\zeta}$ can be viewed as discrete memoryless channels with input alphabet A^n. For ζ, the output alphabet is B^n, and for $\overline{\zeta}$, the output alphabet is $B^n \times \{1, 2, \cdots, k\}$. The transition probabilities for ζ are

$$p(\mathbf{y}/\mathbf{x}) = \sum_{k=1}^{n} \alpha_k \prod_{i=1}^{n} p_k(y_i|x_i)$$

where $\mathbf{y} = (y_1, \cdots, y_n)$, $\mathbf{x} = (x_1, \cdots, x_n)$, $p_k(y|x)$ is the transition probability for ζ_k, and α_k is the probability that the channel selected is $\zeta_k : \alpha_k = Pr \{Z = k\}$. For $\overline{\zeta}$, the transition probabilities are

$$p(\mathbf{y}, k|\mathbf{x}) = \prod_{i=1}^{n} p_k(y_i|x_i)$$

From this memoryless viewpoint, the calculation of the channel capacities is simply a matter of minimizing the appropriate mutual informations. For ζ, the capacity is

$$C_n = \frac{1}{n} \max_{\mathbf{X}} I(\mathbf{X}; \mathbf{Y}) \qquad (1)$$

where $\mathbf{X}$ and $\mathbf{Y}$ denote the (n-component) random inputs to and outputs from ζ. For $\overline{\zeta}$, the formula is

$$\overline{C}_n = \frac{1}{n} \max_{\mathbf{X}} I(\mathbf{X}; \mathbf{Y}, Z) \qquad (2)$$

(We have indicated a dependence on n, but as indicated above the capacity $\overline{C}_n$ turns out to be independent of the burst length.)

We shall consider $\overline{C}_n$ first, since its calculation is the easier of the two. We have, using standard results about mutual information (Ref. 4),

$$I(\mathbf{X}; \mathbf{Y}, Z) = \sum_{k=1}^{K} \alpha_k I(\mathbf{X}; \mathbf{Y}^{(k)})$$

where $\mathbf{Y}^{(k)}$ denotes the output of the channel ζ_k^n, if $\mathbf{X}$ is the input. Since each ζ_k is memoryless, we have

$$\sum_{k=1}^{K} \alpha_k I(\mathbf{X}; \mathbf{Y}^{(k)}) = \sum_{k=1}^{K} \alpha_k \sum_{i=1}^{n} I(X_i; Y_i^{(k)})$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{K} \alpha_k I(X_i; Y_i^{(k)}) \qquad (3)$$

where $\mathbf{X} = (X_1, \cdots, X_n)$ and $\mathbf{Y}^{(k)} = (Y_1^{(k)}, \cdots, Y_n^{(k)})$

To compute $\overline{C}_n$, we are required to maximize this last expression over all random vectors $\mathbf{X} = (X_1, X_2, \cdots, X_n)$. The maximum of the inner sum in Eq. (3), taken over all choices of the random variable X_i, is clearly independent of k, and so from Eq. (2) we have

$$\overline{C}_n = \sup_{\mathbf{X}} \sum_{k=1}^{K} \alpha_k I(\mathbf{X}; \mathbf{Y}^{(k)}) \qquad (4)$$

where the supremum in Eq. (4) is taken over all random variables taking values in the input alphabet A. (If it happens that there is a single input distribution X that simultaneously achieves channel capacity on all K channels ζ_k, then

$$\overline{C}_n = \sum_{k=1}^{n} \alpha_k C_k$$

where C_k is the capacity of ζ_k.) Equation (4) thus shows that $\overline{C}_n$ is independent of n, and that it is in fact the capacity of the DMC with transition probabilities $\{\Sigma \alpha_k p_k (y|x)\}$.

We turn now to the computation of C_n. It is an easy exercise to show that

$$I(\mathbf{X}; \mathbf{Y}, Z) - H(Z) \leqslant I(\mathbf{X}; \mathbf{Y}) \leqslant I(\mathbf{X}; \mathbf{Y}, Z) \qquad (5)$$

It thus follows that for any random vector $\mathbf{X}$,

$$\frac{1}{n} I(\mathbf{X}; \mathbf{Y}, Z) - \frac{H(Z)}{n} \leqslant \frac{1}{n} I(\mathbf{X}; \mathbf{Y}) \leqslant \frac{1}{n} I(\mathbf{X}; \mathbf{Y}, Z) \qquad (6)$$

Since $H(Z)$ is a fixed number $\leqslant \log K$, the left-hand inequality in Eq. (6) shows that

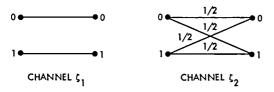$$\liminf_{n\to\infty} C_n \geqslant \overline{C}_n = \overline{C}$$

and the right-hand inequality shows that $C_n \leqslant \overline{C}$, and

$$\limsup_{n \to \infty} C_n \leqslant \overline{C}$$

Together these two inequalities show that $\lim_{n \to \infty} C_n = \overline{C}$, as asserted. (We conjecture, but have not been able to prove, that in fact C_n is a monotonically increasing function of n.)

III. An Example

We illustrate these results with a simple example, with $K = 2$. Channel ζ_1 is a noiseless binary symmetric channel, and channel ζ_2 is a "useless" BSC with raw bit error probability $1/2$.

CHANNEL ζ_1 CHANNEL ζ_2

We assume that the channel selector random variable Z is described by $Pr \{ Z = 1 \} = 1 - \epsilon$, $Pr \{Z = 2\} = \epsilon$. Thus if ϵ is small, the overall channel ζ is characterized by noise-free transmission interrupted by occasional but very severe error bursts.

The capacity of ζ_1 is log 2, and the capacity of ζ_2 is 0; both capacities are achieved by a uniform input distribution, and so by Eq. (4) $\overline{C} = (1 - \epsilon)$ log 2. A straightforward calculation shows that the capacities C_n are given by

$$C_n = (1 - \epsilon_n) \log 2 - \frac{1}{n} \{H(\epsilon_n) + \epsilon_n \log (1 - 2^{-n})\}$$

$$\epsilon_n = (1 - 2^{-n}) \epsilon, \quad H(x) = -x \log x - (1 - x) \log (1 - x)$$

Since $\epsilon_n \to \epsilon$ as $n \to \infty$, it follows from this that $C_n \to \overline{C}$, but of course this also follows from the general results of Section II.

How should these results be interpreted? First, we note that the channel $\overline{\zeta}$ is equivalent to a channel exhibiting erasure bursts, since once it is known that channel ζ_2 was used to transmit a block of n bits, the received versions of these bits should be ignored, since they bear no relationship to the transmitted bits. And it is easy to verify that the capacity of such an erasure-burst channel is indeed $(1 - \epsilon)$ log 2, whatever the burst length.

It is evident that C_n ought to be less than $\overline{C}$, since the receiver using ζ will not know when a received block of length n is bad, whereas the receiver using $\overline{\zeta}$ will, and this extra information cannot possibly hurt performance. But if n is very large, the ζ-users could, for example, include in the n bits in each transmitted packet a certain number of parity checks. To be specific, let us assume in fact that each packet includes $\log_2 n$ parity checks. Then if the packet is transmitted over ζ_1, all of these parity checks will be satisfied upon reception. But if the packet is transmitted over ζ_2, these will be parity checks on random data, and the probability that they will all be satisfied is $2^{-\log_2 n} = n^{-1}$. Thus when n is large, the presence of a useless data packet can be detected with high probability and low overhead. In other words, for large n the channel ζ is virtually identical with $\overline{\zeta}$, and this is what our computations with mutual information predicted.

Thus if n is sufficiently large, a good strategy for communication over ζ is to reserve a certain number of the bits in each transmitter package for parity. This number should be large enough so that the presence of bad data can be detected with high probability, but small enough (relative to n) not to substantially reduce the transmission rate. This strategy will, as previously explained, effectively transform the channel into an erasure-burst channel. Then if n' denotes the number of bits in each packet not reserved for parity, one should code for the channel by interleaving n' copies of a code designed for use on the binary erasure channel (BEC). Since the capacity of the BEC is just as large as that of the erasure-burst channel, presumably there will be no performance loss. Furthermore, the decoding complexity of the n' parallel binary code is much less than n' times the complexity of decoding just one such code; see Ref. 5 for details.

IV. Conclusions

On this basis of the mutual information calculation in Section II, and on the basis of the example in Section III, we draw the following conclusions about RFI channels. First, to communicate reliably over $\overline{\zeta}$, nothing is lost by interleaving, and in addition there may be a considerable advantage in doing so. Second, while there will in general be a penalty in performance if interleaving on ζ is used, if n is large enough, it may be possible to accurately estimate the channel index k affecting a given data packet of length n or by using some kind of generalized parity check. If this can be done, then ζ is effectively transformed into $\overline{\zeta}$ and then interleaving can be used without penalty.

References

1. Massey, J., "Capacity, Cut-Off Rate, and Coding for a Direct-Detection Optical Channel," TDA Progress Report article in this volume (42-60). Jet Propulsion Laboratory, Pasadena, Calif.

2. Viterbi, A. J., "Spread Spectrum Communications — Myths and Realities," *IEEE Communications Magazine*, May 1979, pp. 11-18.

3. Berlekamp, E. R., "Error-Correcting Codes," *Proceedings IEEE*, May 1980, pp. 564-593.

4. McEliece, R. J., *The Theory of Information and Coding.* Addison-Wesley, Reading, Mass., 1977.

5. McEliece, R. J., "Interleaved Block Codes for the Photon Channel," TDA Progress Report article in this volume (42-60). Jet Propulsion Laboratory, Pasadena, Calif.