A Communication Channel Modeled by the Spread of Disease

Fady Alajaji and Tom Fuja
Electrical Engineering Department, Systems Research Center
University of Maryland, College Park, Maryland 20742

1. Overview

We consider a discrete channel with memory in which errors spread like the spread of a contagious disease through a population. Our motivation is the observation by Stapper *et. al.* that the Polya-Eggenberger (PE) distribution is a better "fit" to the distribution of defects in silicon than the commonly used Poisson distribution. The PE distribution is one of the distributions generated by Polya's urn model for the spread of contagion. We introduce a communication channel with noise modeled by Polya's process. We first present a maximum likelihood (ML) decoding algorithm; we then show that this channel is in fact an "averaged" channel in the sense of Ahlswede and others, and its capacity is zero. Finally, we consider a finite-memory version of the Polya-contagion model; this channel is (unlike the original) ergodic with a non-zero capacity.

2. Polya-Contagion Channel

Consider a discrete binary additive communication channel: $\mathbf{Y}_i = \mathbf{X}_i \oplus \mathbf{Z}_i$, where the random variables $\mathbf{X}_i$, $\mathbf{Z}_i$, and $\mathbf{Y}_i$ are, respectively, the i'th input, noise, and output of the channel. We assume that the input and noise sequences are independent. The noise sequence $\{\mathbf{Z}_i\}$ is generated according to Polya's contagion urn scheme, described as follows. An urn originally contains T balls, of which R are red and S are black. Let $\rho = R/T$ and $\sigma = 1 - \rho = S/T$. We make successive draws from the urn; after each draw, we return to the urn $1 + \Delta$ balls of the same color as was just drawn. In our problem we assume that $\Delta > 0$ (contagion case) and that $\rho < \sigma$. The noise sequence $\{\mathbf{Z}_i\}$ is generated by the draws: $\mathbf{Z}_i = 1$ if the i'th draw yields a red ball and $\mathbf{Z}_i = 0$ if the i'th draw yields a black ball.

For an input block $\underline{\mathbf{X}} = [\mathbf{X}_1, \ldots, \mathbf{X}_n]$ and an output block $\underline{\mathbf{Y}} = [\mathbf{Y}_1, \ldots, \mathbf{Y}_n]$, the block transition probability of the channel is:

$$P(\underline{\mathbf{Y}} = \underline{y} \mid \underline{\mathbf{X}} = \underline{x}) = \frac{\Gamma(\frac{1}{\delta})\,\Gamma(\frac{\rho}{\delta} + d)\,\Gamma(\frac{\sigma}{\delta} + n - d)}{\Gamma(\frac{\rho}{\delta})\,\Gamma(\frac{\sigma}{\delta})\,\Gamma(\frac{1}{\delta} + n)} \tag{1}$$

where $d = d_H(\underline{y}, \underline{x})$, the Hamming distance between $\underline{y}$ and $\underline{x}$.

Channel Properties: Two important properties: (1) *Stationarity:* From equation (1) the noise $\{\mathbf{Z}_i\}$ forms an infinite sequence of exchangeable random variables. Therefore, the noise process is strictly stationary. (2) *Non-Ergodicity:* Let $\mathbf{S}_n \triangleq \mathbf{Z}_1 + \mathbf{Z}_2 + \cdots + \mathbf{Z}_n$. It can be shown that $\mathbf{Z} = \lim_{n\to\infty} \mathbf{S}_n/n$ is (with probability one) a random variable drawn according to the beta distribution with parameters ρ/δ and σ/δ. Thus the noise process $\{\mathbf{Z}_i\}$ is not ergodic since its sample average does not converge to a constant.

Maximum Likelihood (ML) Decoding: Suppose M codewords are possible channel inputs: $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_M$, each of length n. Given an output $\underline{y}$, ML decoding selects as its estimate of the transmitted codeword the $\underline{x}_k$ that maximizes $P(\underline{\mathbf{Y}} = \underline{y} \mid \underline{\mathbf{X}} = \underline{x}_k)$.

Now $g(d) \triangleq P(\underline{\mathbf{Y}} = \underline{y} \mid \underline{\mathbf{X}} = \underline{x})$ is strictly log-convex in $d \in [0, n]$ with a unique minimum at $d_0 = n/2 + (1 - 2\rho)/2\delta$. Thus the ML decoding algorithm for the channel is given as follows:

1. Given the received vector $\underline{y}$, compute $d_i \triangleq d_H(\underline{y}, \underline{x}_i)$ for each i. Compute also $d_{\max} \triangleq \max\{d_i\}$ and $d_{min} \triangleq \min\{d_i\}$.
2. If $|d_{\max} - d_0| \le |d_{\min} - d_0|$, map $\underline{y}$ to the $\underline{x}_j$ for which $d_j = d_{\min}$. In this case ML decoding $\Longleftrightarrow$ minimum distance decoding.
3. If $|d_{\max} - d_0| > |d_{\min} - d_0|$, map $\underline{y}$ to the $\underline{x}_j$ for which $d_j = d_{\max}$. In this case ML decoding $\Longleftrightarrow$ maximum distance decoding.

Averaged Communication Channels: Consider a family of discrete memoryless channels parameterized by θ:

$$\{W_\theta^{(n)}(\underline{\mathbf{Y}} = \underline{y} \mid \underline{\mathbf{X}} = \underline{x}) = \prod_{i=1}^{n} W_\theta^{(1)}(\mathbf{Y}_i = y_i \mid \mathbf{X}_i = x_i) : \theta \in \Theta\}_{n=1}^{\infty}.$$

A channel is "averaged" if its block transition probability is the expected value of the block transition probability taken with respect to some distribution on θ – i.e., if it's of the form

$$W_A^{(n)}(\underline{\mathbf{Y}} = \underline{y} \mid \underline{\mathbf{X}} = \underline{x}) = \int_\Theta W_\theta^{(n)}(\underline{\mathbf{Y}} = \underline{y} \mid \underline{\mathbf{X}} = \underline{x})\, dG(\theta) \tag{2}$$

where $(\Theta, \sigma(\Theta), G)$ is a probability space for the random variable θ. Note that the averaged channel has memory and is stationary.

Claim: The binary Polya-contagion channel is an averaged channel; specifically, the Polya-contagion channel represents the class of binary symmetric channels with crossover probability θ, where θ is distributed according to the beta distribution with parameters ρ/δ and σ/δ. Furthermore, from the results of Ahlswede we can show that the capacity of this channel is zero.

3. Finite-Memory Contagion Channel

An unrealistic aspect of the Poly-contagion channel is its infinite memory. Consider, for instance, the millionth ball drawn from Polya's urn; the very *first* ball drawn from the urn and the 999,999'th ball drawn from the urn have the identical effect on the outcome of the millionth draw. We now consider a perhaps more realistic model for a contagion channel with finite memory.

As before, consider an urn with T balls, of which R are red and $S = T - R$ are black. At the j'th draw we select a ball from the urn and replace it with $1 + \Delta$ balls of the same color; then, M draws later – after the $(j + M)$'th draw – we retrieve from the urn Δ balls of the color picked at time j. As before, let $\mathbf{Z}_i = 1$ if the i'th draw yields a red ball and $\mathbf{Z}_i = 0$ if the i'th draw yields a black ball. This modification keeps the total number of balls in the urn constant $(T + M\Delta$ balls) after an initialization period of M draws; it also limits the effect of any draw to M draws in the future.

For blocklength $n \le M + 1$, the block transition probability of this new channel is given by (1). For $n \ge M + 2$, we obtain:

$$P(\underline{\mathbf{Y}} = \underline{y} \mid \underline{\mathbf{X}} = \underline{x}) = L \prod_{i=M+2}^{n} \left[\frac{\rho + s_{i-1}\delta}{1 + M\delta}\right]^{e_i} \left[\frac{\sigma + (M - s_{i-1})\delta}{1 + M\delta}\right]^{1-e_i}$$

where $L = [\prod_{i=0}^{k-1}(\rho + i\delta) \prod_{j=0}^{M-k}(\sigma + j\delta)] / \prod_{\ell=1}^{M}(1 + \ell\delta)$. Here, $e_i = x_i \oplus y_i$, $k = e_1 + \cdots + e_{M+1}$, and $s_{i-1} = e_{i-1} + \cdots + e_{i-M}$.

Claim: The new noise process $\{\mathbf{Z}_i\}$ is a stationary ergodic Markov process of order M, and thus the resulting channel is a Markov channel with memory M. The capacity C_M of the channel is given by:

$$C_M = 1 - \sum_{i=0}^{M} \binom{M}{i} L_i\, h_b\left(\frac{\rho + i\delta}{1 + M\delta}\right)$$

where $L_i = [\prod_{j=0}^{i-1}(\rho + j\delta) \prod_{\ell=0}^{M-i-1}(\sigma + \ell\delta)] / \prod_{m=1}^{M-1}(1 + m\delta)$, and $h_b(x)$ is the binary entropy function.

Finally if we let $M \to \infty$, $C_M \to 1 - \int_0^1 h_b(z) f_{\mathbf{Z}}(z) dz$ where $f_{\mathbf{Z}}(z)$ is the beta pdf with parameters ρ/δ and σ/δ. This result is identical to $\lim_{n\to\infty}(1/n) I(\mathbf{X}^n; \mathbf{Y}^n)$ if $\mathbf{X}^n$ and $\mathbf{Y}^n$ are blocks of length n joined by the original Polya-contagion channel (with equally likely inputs). Thus as $M \to \infty$, the stationary ergodic finite-memory contagion channel converges in distribution to the stationary non-ergodic Polya channel, but C_M does not converge to $C_{Polya} = 0$.