BANDWIDTH SCALING FOR FADING MULTIPATH CHANNELS

Medard & Gallager

Bottom line: If a CDMA system is spread over too much bandwidth on a fading channel, it ceases to work.

More generally, any system that, in a sense, spreads its power over too much bandwidth ceases to work.

Principle: If the receiver can't measure the channel adequately, it can't detect the data.

CHANNEL MODELING

WLOG we can view a wideband channel of bandwidth $\mathcal{W} = bW$ as a collection of b adjacent narrow band channels each of some fixed bandwidth W.

We look at what happens when b increases with a fixed overall power constraint.

We use the sampling theorem to view each narrow band channel as a discrete time channel with W complex input and output symbols per second.

Thus we view the channel as a time/frequency grid, with 2 degrees of freedom per W Hertz, $1/W$ seconds.

Wireless channels often have multiple electromagnetic paths from source to destination.

The path lengths often change due to motion of source, destination, or reflecting objects.

Each path in motion inserts Doppler shift into corresponding portion of received waveform.

Doppler spread between paths causes impulse reponse of channel to change. The coherence time $\mathcal{T}_c$ is the time for this response to change significantly.

$$\mathcal{T}_c \approx \frac{c}{f_c v}$$

There is also a time spread s from early to late returns.

The channel system function (at a given time) changes significantly over a frequency coherence interval $\mathcal{F}_c$ given by

$$\mathcal{F}_c \approx 1/s$$

Choosing frequency slices of width $W = \mathcal{F}_c$, the output at slice a, discrete time i can be approximated by

$$Y_{i,a} = F_{i,a} X_{i,a} + Z_{i,a}$$

The simplest model is to assume that $F_{i,a}$ is an array of circularly symmetric complex Gaussian rv. They are independent between slices, the same over a segment of $\mathcal{F}_c\mathcal{T}_c$ discrete time intervals and independent between segments.

This model corresponds to Rayleigh fading - an essentially infinite number of small randomly changing paths.

Results are quite different for a finite set of paths (Telatar & Tse, Trans. IT, July 2000).

We want to look at CDMA like systems that spread the available signal energy essentially uniformly over the available degrees of freedom.

The point of the talk is that as $\mathcal{W} \to \infty$, the available energy per degree of freedom $\to 0$.

This makes it impossible to measure the channel gain in that degree of freedom.

This makes that degree of freedom 'useless.'

There is a modeling subtlety here: suppose (in a given degree of freedom) we send 0 with probability $1-\epsilon$ and $\pm\sqrt{E}$ with probability $\epsilon/2$ each. The energy is ϵE.

Interpreted differently, the energy is E if the degree of freedom is used, and 0 otherwise.

If we 'use' an entire time/frequency coherence block or not in this way, then the channel can be measured over the blocks that are used.

This strategy is akin to frequency hopping rather than uniform spreading.

The simplest way to ensure uniform spreading is by a fourth moment input constraint.

Suppose that S is the overall power constraint. Then $S/\mathcal{W}$ is the average energy per complex variable.

Our input constraint (along with the overall power constraint) is that for some constant α,

$$\mathsf{E}[|X|^4] \leq \frac{\alpha S^2}{\mathcal{W}^2}$$

If we constrain $\mathsf{E}[|X|^2] \leq S/\mathcal{W}$, then this would constrain the kurtosis to be at most α.

For example above, $\mathsf{E}[|X|^2] = \epsilon E$, $\mathsf{E}[|X|^4] = \epsilon E^2$, kurtosis$= 1/\epsilon$.

The coding theorem for this channel works the same as for any channel.

The question is how much mutual information can be crammed through the channel.

First look at a single frequency slice,

$$Y_i = X_i F_i + Z_i$$

$\{Z_i\}$ complex, circ. symm., iid Gauss, $\mathsf{E}[|Z_i|^2] = N_0$

$\{F_i\}$ complex, circ. symm., stationary Gauss,

$$\mathsf{E}[|F_i|^2] = 1, \qquad \sum_{i=\infty}^{\infty} |\mathsf{E}[F_0 F_i]|^2 = 2\mathcal{T}_c\mathcal{F}_c.$$

Theorem:

$$\frac{1}{n} I(X^n; Y^n) \leq \frac{\alpha S^2 \mathcal{T}_c \mathcal{F}_c}{\mathcal{W}^2 N_0^2}$$

We outline the proof.

$$I(X^n; Y^n) = I(Y^n; X^n) = \sum_{i=1}^{n} I(Y_i; X^n \mid Y^{i-1})$$

$$I(Y_i; X^n \mid Y^{i-1}) = h(Y_i \mid Y^{i-1}) - h(Y_i \mid Y^{i-1}, X^n)$$

As $\mathcal{W}$ gets large, the signal power per slice gets small, so we are looking at a low signal to noise ratio regime. Thus the output entropy density is easily bounded. The conditional entropy is the interesting term.

Consider a sample sequence of inputs x^n. Look at Y_i, F_i and $\vec{Y} = Y^{i-1}$ conditional on x^n. These variables are then jointly Gaussian and jointly circularly symmetric.

Consider estimating (MMSE) F_i from a given sample value $\vec{y}$.

$$\widehat{F}_i(x^n, \vec{y}) = \mathsf{E}[F_i \vec{Y}^{*\mathsf{T}} | x^n] K_{\vec{Y}|x}^{-1} \vec{y}$$

$$\mathsf{B_i}(x^n) = \mathsf{E}[F_i F_i^{*\mathsf{T}}] - \mathsf{E}[F_i \vec{Y}^{*\mathsf{T}} | x^n] \mathsf{K}_{\vec{Y}|x}^{-1} \mathsf{E}[\vec{Y} F_i^* | x^n]$$

where $\widehat{F}_i$ is the estimate and B_i the variance.

Lemma: $h(Y_i \mid x^n, \vec{Y}) = \ln[\pi e(N_0 + |x_i|^2 B_i(x^n))]$.

Proof:

$$Y_i = x_i F_i + Z_i = x_i \widehat{F}_i + x_i \widetilde{F}_i + Z_i$$

Conditional on x^n and $\vec{y}$, $\widehat{F}_i$ is simply a constant.

Thus, conditionally, Y_i is Gaussian with mean $x_i \widehat{F}_i$ and with variance $|x_i|^2 B_i(\vec{x}^n) + N_0$. Thus

$$h(Y_i \mid \vec{y}, x^n) = \ln[\pi e(N_0 + |x_i|^2 B_i(x^n)]$$

Averaging over $\vec{Y}$, the result follows. $\square$

Note the following:

The conditional entropy $h(Y_i|\vec{Y}, x^n)$ is a function only of N_o, x_i, and the variance of the estimate of F_i.

The estimate of F_i is an idealized estimate since the receiver doesn't know x^n.

Any actual estimate at the receiver has at least this much variance.

We still must average over the input x^n

Rewrite $h(Y_i \mid x^n, \vec{Y})$ as

$$h(Y_i \mid x^n, \vec{Y}) = \ln\left[\pi e N_0\right] + \ln\left[1 + \frac{|x_i|^2 \mathsf{B_i}(x^n)}{N_0}\right]$$

For x_i small, the second log term is almost linear in $x_i|^2$, so

$$h(Y_i \mid X^n, \vec{Y}) \approx \ln\left[\pi e (N_0 + \mathsf{E}\left[|X_i|^2 \mathsf{B_i}(X^n)\right]\right]$$

We similarly have

$$h(Y_i) \leq \ln\left[\pi e (N_0 + \mathsf{E}\left[|X_i|^2 \mathsf{B_i}(X^n)\right] + \mathsf{E}[|X_i \widehat{F}_i|^2]\right]$$

Combining,

$$I(Y_i; X^n \mid \vec{Y}) \leq \ln\left[1 + \frac{\mathsf{E}[|X_i \widehat{F}_i|^2]}{|X_i|^2 \mathsf{B_i}(X^n) + N_0}\right]$$