From Sensors to Video: Information Theoretic Foundations of Source Communication over Wireless Channels

Elza Erkip

New York University Tandon School of Engineering

2016 Australian Information Theory School
Mobile data traffic is expected to grow nearly tenfold in 2014-2019.
57% compound annual growth rate.
Application requirements: high rates, low latency, ultra-high reliability etc.
Video Traffic

- Video: TV, video on demand, Internet, P2P.
- Every second, nearly a million minutes of video content will cross the network by 2019.
- The sum of all forms of video will be in the range of 80-90% of global consumer traffic by 2019.
  - Share of video was 64% in 2014.

Cisco estimates that billions of devices and sensors will be connected by 2020.

- Wireless is pervasive.
- Miniaturization.
- Smart devices and sensors everywhere.
Transmitter-receiver pairs $(T_i, D_i)$, relays $R_i$.

Source $S$, relay side information $U$, destination side information $W$. 
Signals \( (S, U, W) \) correlated in time and space.
- Sensor data.
- Internet of Things.

Signals can be represented at various quality levels.
- Scalable video coding.
  - Successive refinement.

How do we harness these properties for communication over wireless channels?

How does source compression and channel coding interact?
This Lecture

- Information theoretic foundations of source communication over wireless channels.
- Two important source properties:
  - Correlation.
  - Scalability.
- Interaction of source and channel coding.
- New notions for separation of source and channel coding.
Knowledge at the level of an introductory course in information theory is assumed.

- Quick review of source-channel coding fundamentals.

The coverage reflects my own tastes and experience. I apologize in advance for any omissions.

Mathematical expressions is kept to a minimum to maximize the transfer of intuition.
Lecture Outline

- Exploiting source correlation.
  - Why correlation matters.
  - Shannon’s source-channel separation theorem in networks.
  - Informational versus operational separation.
  - Broadcast channel, multiple access channel, compound multiple access channel, interference channel, relay channel.

- Scalable compression.
  - Communication over fading channels.
  - Multicasting.
  - Multicasting with user caches.
Exploiting Source Correlation
Motivation: Relay Channel

- Sending a source with the help of a relay.
  - Reproduce source $S_0$ at the destination as $\hat{S}_0$ (either lossless or lossy reconstruction)

- First compress the source $S_0$, then send over the relay channel.

- Is this always optimal? Does source-channel separation always hold?
Motivation: Correlated Side Information

Nodes have correlated side information: \( p(s_0, s_1, s_2) \).
- Sensor network.
- In classical relay network, relay helps by forwarding its received signal.
  - Now the relay can use its side information to better “hear” the source.
- In the correlated model, relay can forward its side information as well.
Motivation: Relay Doesn’t Hear the Source

If $S_0$ and $S_1$ are independent, relay cannot help.

If $S_0$ and $S_1$ are correlated, relay can still help by forwarding $S_1$.
- Will call this *source cooperation*.
- Also known as the *helper problem* in source coding literature.

In general, relay can forward both received signal and side information.
Correlation can in general be useful.

What do we know about separate source and channel coding in networks?
Point-to-Point Communication

- Lossless/lossy reconstruction.
Shannon’s Source-Channel Separation Theorem

Separate source and channel coding optimal.

\[ S \xrightarrow{\text{Transmitter}} \text{Channel} \xrightarrow{\text{Receiver}} \hat{S} \]

- Source encoder
- Channel encoder
- Channel
- Channel decoder
- Source decoder

\[ S \xrightarrow{\text{Source encoder}} \text{Channel encoder} \xrightarrow{\text{Channel}} \text{Channel decoder} \xrightarrow{\text{Source decoder}} \hat{S} \]
Separation is good, because ...

- Brings modularity.
- We can benefit from existing source and channel coding techniques.
Shannon’s Source-Channel Separation Theorem

- Separation is good, because ...
  - Brings modularity.
  - We can benefit from existing source and channel coding techniques.

- But ...
  - Infinite delay and complexity.
  - No separation theorem for multi-user networks.
  - No separation theorem for non-ergodic source or channel.
Separation in Networks: MAC with Correlated Sources

\[ S_1 \xrightarrow{X_1} \text{MAC} \xrightarrow{Y} \text{Rx.} \rightarrow (\hat{S}_1, \hat{S}_2) \]

\[ p(y | x_1, x_2) \]
Assume the channel is ideal.
- Distributed source coding.
- Lossless reconstruction: \( p(S_1 \neq \hat{S}_1, S_2 \neq \hat{S}_2) \) small.
- For lossless reconstruction we have the Slepian-Wolf region.

\[
\begin{align*}
R_1 &\geq H(S_1 | S_2) \\
R_2 &\geq H(S_2 | S_1) \\
R_1 + R_2 &\geq H(S_1, S_2)
\end{align*}
\]
Assume sources are independent.

Multiple Access Channel (MAC) capacity region

\[ R_1 \leq I(X_1; Y|X_2) \]
\[ R_2 \leq I(X_2; Y|X_1) \]
\[ R_1 + R_2 \leq I(X_1, X_2; Y) \]

Maximum over all \( p(x_1, x_2) = p(x_1)p(x_2) \).
Separation in Networks: MAC with Correlated Sources

With separate source and channel coding, lossless transmission not possible.
Uncoded is optimal: $X_1 = S_1, X_2 = S_2$.

In general, use source correlation to induce correlation of transmitted codewords. (*Cover, El Gamal, Salehi, IT 1980*)
Exploiting source correlations.
First lossless reconstruction.
- Distributed lossy compression is more difficult.
Concentrate on canonical network models.
A generalized notion of source-channel separation.
**Source-Channel Rate**

- **Source-channel rate, \( b \):** Number of channel uses per source sample required for lossless transmission.  
  
  Also known as *bandwidth ratio*.

\[
\begin{array}{c|cc}
S_1 & 0 & 1 \\
\hline
0 & 1/3 & 1/3 \\
1 & 0 & 1/3 \\
\end{array}
\]

\[H(S_1, S_2) = \log 3 = 1.58\]

\[\max_{p(x_1)p(x_2)} I(Y; X_1, X_2) = 1.5\]

- For separate source and channel coding \( b = \frac{1.58}{1.5} = 1.05 \) is achievable.

- Optimal \( b = 1 \) is achieved by uncoded transmission.
Computable expression for source-channel rate?

Optimum source-channel rate by statistically independent source and channel codes?
- Source and channel codes depend only on source and channel distributions respectively.
- Not necessarily optimal for the underlying source and the channel.

Optimal source-channel rate by simply comparing source and channel rate regions?
- Regions not necessarily optimal for the compression and channel coding problems respectively.

Operational separation when above properties are satisfied.

Informational separation when optimal source code + optimal channel code → optimal source-channel rate.
Channel used \( n \times \) times for lossless compression of \( m \) samples of \( S_0 \).

Source-channel rate \( b = n/m \) achievable iff

\[
H(S_0|S_1) \leq bI(X;Y).
\]

Separate source and channel coding is optimal (Verdu, Shamai, *IT* 1995).

Slepian-Wolf compression + optimal channel coding.

We call this informational separation.
Informational Separation

- **Encoding**

\[ 2^{m[H(S_0)+\delta]} \]

source vectors

\[ 2^{m[H(S_0|S_1)+\delta]} \]

bin indices

\[ x^n(i) \]

i.i.d. \( \sim P_X \)

Optimal channel code

Optimal source code

- **Decoding:** First decode the channel code, then the source code.
Operational Separation

- **Encoding**

\[ 2^n [H(S_0) + \delta] \]

source vectors

\[ s_0^n(i) \leftarrow x^n(i) \]

i.i.d. \( \sim P_{S_0} \)

i.i.d. \( \sim P_X \)

- **Decoding**: Find \( i \) such that \((x^n(i), Y^n)\) jointly typical and \((s_0^m(i), S_1^m)\) jointly typical.
- Binning is automatically done in the channel.
Operational Separation

- For point-to-point setting, achieves the same optimal source-channel rate as SW+channel coding.
- Will see it is useful in network settings.
- Optimum source-channel rate by statistically independent source and channel codes.
  - Source and channel codes depend only on source and channel distributions respectively.
  - Not necessarily optimal for the underlying source and the channel.
Exploiting Source Correlation: Road Ahead

- Broadcast channel.
- Multiple access channel.
- Compound MAC.
- Interference channel.
- Relay channel.
  - Lossless.
  - Lossy.
- Source correlations, receiver side information.
- Each case informational or operational separation.
(S₀, S₁, S₂) \sim p(s₀, s₁, s₂).

Source-channel rate \( b \).

*Multicasting* to receivers with correlated information.
Separate Source-Channel Coding for the BC

- Rate region for the BC: All \((R_1, R_2)\) such that

\[
R_1 \leq I(U; Y_1) \\
R_2 \leq I(X; Y_2|U)
\]

for some joint distribution \(p(u)p(x|u)p(y_1, y_2|x)\).

- For separate source-channel coding
  - Two SW compressed streams: \(S_0\) with receiver side information \(S_1\) and \(S_2\) respectively.
  - Send these two compressed streams over the BC.

- Reliable reception possible if

\[
H(S_0|S_1) \leq bl(U; Y_1) \\
H(S_0|S_2) \leq bl(X; Y_2|U)
\]

for some joint distribution \(p(u)p(x|u)p(y_1, y_2|x)\).
For joint source-channel coding reliable reception possible if and only if

\[ H(S_0 | S_1) \leq bl(X; Y_1) \]
\[ H(S_0 | S_2) \leq bl(X; Y_2) \]

for some joint distribution \( p(x)p(y_1, y_2|x) \). \((Tuncel, IT 2006)\)

Performance as if each receiver is alone!
- Need same \( p(x) \) for both channels.
Achievability

- Same strategy as in point-to-point case.
- Encoding

\[
2^n \left[ H(S_0) + \delta \right]
\]

source vectors

\[ s_0^n(i), \quad x^n(i) \]

Decoder \( i \) jointly searches for the codeword typical with \( Y_i^n \) and \( S_i^m \).
Correlated Sources over BC

- Computable expression for source-channel rate.
- Optimum source-channel rate by statistically independent source and channel codes.
  - Source and channel codes depend only on source and channel distributions respectively.
  - Not necessarily optimal for the underlying source and the channel.
- Optimal source-channel rate by simply comparing source and channel rate regions.
  - Regions not necessarily optimal for the compression and channel coding problems respectively.
- Operational separation is optimal.
MAC with Correlated Sources and Receiver Side Information

Separate source and channel coding

\[
\begin{align*}
H(S_1 | S_2, W) & \leq bl(X_1; Y | X_2) \\
H(S_2 | S_1, W) & \leq bl(X_2; Y | X_1) \\
H(S_1, S_2 | W) & \leq bl(X_1, X_2; Y)
\end{align*}
\]

for some \( p(x_1)p(x_2) \).
Separation in MAC

- Separate source and channel coding in MAC optimal when
  - $S_1 \rightarrow W \rightarrow S_2$
    - Receiver knows $W$, additional channel correlation useless.
  - $S_1, S_2$ independent, $W$ correlated with $(S_1, S_2)$.
    - Transmitters cannot exploit correlation given $W$.
    - Providing $W$ to transmitters would result in lower $b$.

- In both cases **informational** separation is optimal.
Exploiting Source Correlation: Road Ahead

- Broadcast channel.
- Multiple access channel.
- **Compound MAC.**
- **Interference channel.**
- Relay channel.
  - Lossless.
  - Lossy.
- **Source correlations, receiver side information.**
- **Each case informational or operational separation.**
Two receivers, each with side information $W_i, i = 1, 2$.
Both receivers want lossless reproduction of both sources.
Two simultaneous multicasts.
Assume \((S_1, W_2)\) independent of \((S_2, W_1)\).

Transmitters far away, each transmitter close to one receiver.
Optimal source-channel rate $b$ given by

$$H(S_1 | W_i) \leq bl(X_1; Y_i | X_2)$$
$$H(S_2 | W_i) \leq bl(X_2; Y_i | X_1)$$
$$H(S_1 | W_i) + H(S_2 | W_i) \leq bl(X_1, X_2; Y_i)$$

for some $p(x_1)p(x_2), i = 1, 2$. 
Resembles intersection of two MACs but informational separation not optimal.

For $S_2 = 0$, BC with correlated sources.

Achievability

- Map typical source sequences to channel codewords.
- Receiver $i$ decodes each source index jointly using $(Y_i, W_i)$.

Operational separation optimal.
Interference Channel with Side Information

Two receivers, each with side information $W_i$, $i = 1, 2$.

Receiver $i$ want lossless reproduction of $S_i$ only.
For independent \((S_1, S_2)\) and no \((W_1, W_2)\) we have the regular IC.

Capacity region not known in general.

Under strong interference

\[
I(X_1; Y_1|X_2) \leq I(X_1; Y_2|X_2) \\
I(X_2; Y_2|X_1) \leq I(X_2; Y_1|X_1)
\]

capacity region given by that of the compound MAC.

- Decoding interference is optimal.
Interference Channel with Independent Sources, Receiver Side Information

- Assume \((S_1, W_2)\) independent of \((S_2, W_1)\).
- **Strong source-channel interference** condition
  
  \[
  \begin{align*}
  bI(X_1; Y_1|X_2) & \leq bI(X_1; Y_2|X_2) + I(S_1; W_2) \\
  bI(X_2; Y_2|X_1) & \leq bI(X_2; Y_1|X_1) + I(S_2; W_1)
  \end{align*}
  \]

- Considers additional flow of information due to receiver side information.
- Weaker than usual strong interference condition.
Under strong source-channel interference, optimal source-channel rate $b$ given by

$$H(S_1|W_i) \leq bI(X_1; Y_i|X_2)$$
$$H(S_2|W_i) \leq bI(X_2; Y_i|X_1)$$
$$H(S_1|W_i) + H(S_2|W_i) \leq bI(X_1, X_2; Y_i)$$

for some $p(x_1)p(x_2)$, $i = 1, 2$.

Operational separation is optimal.
For more details, see Gunduz, Erkip, Goldsmith, Poor, IT 2009.
Exploiting Source Correlation: Road Ahead

- Broadcast channel.
- Multiple access channel.
- Compound MAC.
- Interference channel.
- **Relay channel.**
  - Lossless.
  - Lossy.
- Source correlations, receiver side information.
- Each case informational or operational separation.
Relay Channel with Correlated Sources

\[ p(y_1, y_2 | x_0, x_1) \]
Assume relay also obtains a lossless version of $S_0$.

Reliable reception possible if for some $p(x_0, x_1)$

- $H(S_0 | S_1) \leq bI(X_0; Y_1 | X_1)$
Reliable reception possible if for some \( p(x_0, x_1) \)

- \( H(S_0|S_2) \leq bl(X_0, X_1; Y_2) \)
Reliable reception possible if for some $p(x_0, x_1)$

- $H(S_0|S_1) \leq bI(X_0; Y_1|X_1)$
- $H(S_0|S_2) \leq bI(X_0, X_1; Y_2)$

Need joint source-channel coding involving a modified form of block Markov encoding and backward decoding.

In each block, relay uses $S_1$ and $Y_1$ to losslessly recover $S_0$, then works with the transmitter to send it to the destination.
### Block Markov Structure

<table>
<thead>
<tr>
<th>Source</th>
<th>Block 1</th>
<th>Block 2</th>
<th>Block B</th>
<th>Block B+1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>$x_0^N(1, w_1)$</td>
<td>$x_0^N(w_1, w_2)$</td>
<td>$\cdots$</td>
<td>$x_0^N(w_{B-1}, w_B)$</td>
</tr>
<tr>
<td>Relay</td>
<td>$x_1^N(1)$</td>
<td>$x_1^N(w_1)$</td>
<td>$\cdots$</td>
<td>$x_1^N(w_{B-1})$</td>
</tr>
</tbody>
</table>
Separate Source-Channel Coding

- Encoder/decoder at each block
Joint Source-Channel Coding

- Encoder/decoder at each block
This strategy is optimal when

- Relay also wants to recover $S_0$ (cooperative relay broadcast channel).
- We have a physically degraded relay channel and side information.

Interacting separate source and channel codes optimal in each block.

*Operational separation.*

More: Gunduz, Erkip, Goldsmith, Poor, IT 2013.
Operational Separation for the Relay Channel

- Computable expression for source-channel rate.
- Optimum source-channel rate by statistically independent source and channel codes.
  - Source and channel codes depend only on source and channel distributions respectively.
  - Not necessarily optimal for the underlying source and the channel.
- Optimal source-channel rate by simply comparing source and channel rate regions.
  - Regions not necessarily optimal for the compression and channel coding problems respectively.
We now study the lossy case:
- Source and side information jointly Gaussian, the channel is AWGN.
- In the lossless case relay only used $S_1$ to better decode $S_0$.
- Now the relay will also help by forwarding $S_1$. 
No destination side information $S_2$.

$(S_0, S_1)$ jointly Gaussian with correlation $\rho$.

Average source distortion $E(S_0 - \hat{S}_0)^2 = D$.

Source-channel rate $b = 1$. 
Strategies: Channel Cooperation

- Ignore relay side information.
- Separate source-channel coding.
- Possible strategies:
  - Decode-and-forward: DF
  - Compress-and-forward: CF
  - Direct transmission (do not use relay): DT
Strategies: Source Cooperation

- Ignore received signal at the relay.
- $S_0$ and $S_1$ are sent uncoded
  - Uncoded transmission (UT) over MAC optimal for low SNR’s.
- $S_0$ and $S_1$ compressed and sent over MAC.
  - Helper MAC (hMAC)
Compare all strategies with a lower bound on $D$ obtained from the cutset bound.
Low Quality Relay Side Information

$P_1 = 0 \text{ dB}, P_2 = 10 \text{ dB}, \rho = 0.2$

![Graph showing the relationship between Average SNR for S-R link (dB) and Average Distortion for different transmission methods.](image)

- DT
- DF
- CF
- hMAC
- Uncoded
- Lower bound
High Quality Relay Side Information

\[ S-D = 0 \text{ dB}, \quad S-R = 0 \text{ dB}, \quad \rho = 0.9 \]

![Graph showing the relationship between average distortion and average SNR for R-D link (dB)]
Strategies: Hybrid Source-Channel Cooperation

- Combine source cooperation (helper MAC) with channel cooperation.
- Relay reserves $\gamma$ proportion of power for DF/CF relaying, the rest for helper MAC.
- Advanced cooperation: Relay side information is used to facilitate relay operation.
Effect of Side Information Quality

S-D = -5 dB, S-R = 0 dB, R-D = 15 dB

Average Distortion vs. Correlation Coefficient (P)

- DT
- Channel Coop.
- Source Coop.
- hDF
- hCF
- aCF
- Lower bound
Channel cooperation, source cooperation, hybrid forms.

Relay can use side information $S_1$ in various ways:

- Use $S_1$ to better receive the source.
- Send $S_1$ to the destination directly.

Best strategy depends on the source correlation and channel conditions.

More:

- Gunduz, Ng, Erkip, Goldsmith, ISIT 2007.
- Gunduz, Erkip, Goldsmith, Poor, ISIT 2008.
In multiuser networks, availability of source correlation and receiver side information improves performance.

In general, need joint source-channel coding.

Operational separation is optimal in some special cases.
- Allows for statistically independent source and channel codes.
- Source-channel rate can be found by comparing (not necessarily optimal) source and channel rate regions.
Lecture Outline

- Exploiting source correlation.
  - Why correlation matters.
  - Shannon's source-channel separation theorem in networks.
  - Informational versus operational separation.
  - Broadcast channel, multiple access channel, compound multiple access channel, interference channel, relay channel.

- **Scalable compression.**
  - Communication over fading channels.
  - Multicasting.
  - Multicasting with user caches.
Scalable Compression
Multiple Descriptions Problem

- Two descriptions at rates \((R_1, R_2)\).
- Corresponding reconstructions \((\hat{S}_1, \hat{S}_2)\) at distortions \((D_1, D_2)\).
- Combined reconstruction \(\hat{S}_0\) at distortion \(D_0\).
- Goal: All possible \((R_1, R_2)\) for given \((D_1, D_2, D_0)\).

\((El \ Gamal, \ Cover, \ IT \ 1982)\)
Successive Refinement of Information

- In the MD problem, set $D_2 = \infty$.
  - Second description not important by itself.
- Require: $R_1 = R(D_1)$, $R_1 + R_2 = R(D_0)$

Description 1 and combined description operate on the rate-distortion curve.
No loss in 2-stage description.

*(Equitz, Cover, IT 1991)*
Successive Refinability

- Gaussian source is successively refinable.
- Successive refinability extends to multiple layers at rates $R_1, R_2, \ldots$
- Useful when
  - Channel quality varies in time: Fading.
  - Channel quality varies in space: Multicasting.
- Practical implementation: Scalable video coding.
  - Base layer: Acceptable quality.
  - Enhancement layer(s): Further improvement in quality.
  - In practice coding loss with each layer:

$$R_1 + R_2 > R(D_0).$$
Scalable Compression: Road Ahead

- Scalable source-channel coding for the fading channel.
  - Multi-antenna single user channel.
  - Relay channel.
- Scalable video multicasting with relays.
- Distortion-memory tradeoffs in cache-aided multicast.
\[ Y = HX + Z. \]

- Average power constraint \( P = \text{SNR}. \)
- Channel \( H, M_r \times M_t \) matrix.
- Slow fading: \( H \) random, but constant for \( n \) channel uses where \( n \) is large.
- Non-ergodic channel: Capacity \( C(H) \) random.
- Assume no channel state information at the transmitter (CSIT).
- Probability of outage: \( P_{out}(R) = P(C(H) \leq R). \)
At high SNR, for fixed $R$, $P_{out} \approx \text{SNR}^{-M_t M_r}$.

What if $R$ also increases with SNR?

Two inter-related issues at high-SNR:

- Throughput, measured by *multiplexing gain* $r$ when $R \approx r \log \text{SNR}$.
- Reliability, measured by *diversity gain* $d$: Slope of error w.r.t. $\log \text{SNR}$.

More precisely,

$$r = \lim_{\text{SNR} \to \infty} \frac{R(\text{SNR})}{\log \text{SNR}} \quad d(r) = - \lim_{\text{SNR} \to \infty} \frac{\log P_{out}(R)}{\log \text{SNR}}$$

DMT: All possible $(r, d(r))$ pairs.
DMT can be expressed by a piecewise linear function.

For $3 \times 2$ MIMO, DMT is

$\text{(Zheng, Tse, 2002)}$
Send $m$ iid samples of a complex Gaussian source over $n$ uses of a $M_t \times M_r$ slow fading MIMO channel.

- Source-channel rate $b = n/m$.

No CSIT suggests there will be channel outages.

- Not clear which rate $R$ to choose in $P_{out}(R)$.

Separate source and channel coding is not optimal.

Performance measure: End-to-end distortion.

- Optimal strategy not known at finite SNR.
- Concentrate on high-SNR.
Distortion Exponent

\[ \Delta = - \lim_{{\text{SNR} \to \infty}} \frac{\log(\text{Distortion})}{\log \text{SNR}}. \]

- Distortion exponent: \( \Delta \).
- End-to-end distortion decays as \( \text{SNR}^{-\Delta} \) at high-SNR.
- Goal: Fix \( b \), maximize \( \Delta \).
Channel coding rate $R$ bits/channel use.

Source coding rate $nR/m = bR$ bits/source sample.

Expected distortion $ED$:

$$ED = (1 - P_{out}(R))D(bR) + P_{out}(R).$$

For $ED \to 0$, we need $R \to \infty$.

Let $R = r \log \text{SNR}$.

Distortion-rate function: $D(R) = 2^{-bR}$.

High rate approximation for non-Gaussian sources.
Distortion Exponent for Single Layer

\[ ED = (1 - P_{out}(R))D(bR) + P_{out}(R). \]

- At high SNR, \( P_{out}(R) \approx \text{SNR}^{-d(r)}. \)
- \( ED \approx \text{SNR}^{-br} + \text{SNR}^{-d(r)}. \)
- To maximize \( \Delta \), choose \( br = d(r). \)
Assume full CSIT.

For each $H$

- The transmitter can transmit at $C(H)$.
- Never in outage.
- Source-channel separation applies.

$ED = E(2^{-bC(H)})$.

Can show $\Delta_{UB} = \sum_{i=1}^{\min(M_t, M_r)} \min(b, 2i - 1 + |M_t - M_r|)$. 
Upper Bound vs Single Layer

2x2 MIMO

Upper bound

Single rate
Think about the fading channel with no CSIT as a broadcast channel with infinitely many receivers, one for each $\mathbf{H}$. (Shamai, Steiner, IT 2003)

Use successive refinability of the source to form many source layers at rates $R_1, R_2, \ldots$.

Aim for the receiver to obtain as many layers at its channel gain permits.

- Not possible to reach $C(\mathbf{H})$ for all receivers.

How to choose layer rates?

- $R_1 < R_2 < \ldots$ since layer $i$ is useless unless all previous layers are received.

What broadcast strategy to use?
Gaussian Broadcast Channel

\[ Y_1 = aX + Z_1, \ Y_2 = X + Z_2 \]

- \( Z_i \sim N(0, 1), \ a > 1, \ E(X^2) \leq P. \)
- Capacity region is obtained by superposition:

\[ R_1 \leq C(a^2 \alpha P), \ R_2 \leq C \left( \frac{(1 - \alpha)P}{1 + \alpha P} \right) \]
Layered Source with Progressive Transmission (LS)

- Transmit source layers progressively in time.
  - Does not achieve BC capacity, but simple to implement.

<table>
<thead>
<tr>
<th>$\alpha n$</th>
<th>$(1-\alpha)n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>Enhancement</td>
</tr>
<tr>
<td>$R_1$ bits/ch. use $\leq$ $R_2$ bits/ch. use</td>
<td></td>
</tr>
</tbody>
</table>

- Number of layers $\rightarrow \infty$ for largest distortion exponent.
Choosing Rates of the Layers

MIMO DMT curve

\[ y = d^*(r_2) + bt_1 r \]

\[ \Delta = d^*(r_1) \]

\[ y = d^*(r_n) + bt_{n-1} r \]

\[ y = bt_n r \]
Distortion Exponent of LS

2x2 MIMO

Distortion exponent vs Bandwidth ratio, b

- LS (infinite layers)
- LS (1 layer)
Analog/Digital Hybrid LS

- Analog (uncoded) is optimal for $b = 1$ AWGN channel.
- Provides graceful performance degradation with channel quality.
- Hybrid digital-analog (HDA) transmission (*Mittal, Phamdo, IT 2002*)
- Combine analog + digital with multiple layers: Hybrid LS (HLS)

<table>
<thead>
<tr>
<th>$\alpha n$</th>
<th>$\beta n$</th>
<th>$(1-\alpha-\beta)n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>Enhancement</td>
<td>Analog</td>
</tr>
<tr>
<td>$R_1$ bits/ch. use</td>
<td>$R_2$ bits/ch. use</td>
<td></td>
</tr>
</tbody>
</table>
Distortion Exponent of HLS

$M_t = 2, M_r = 2$

Distortion exponent, $\Delta$ vs. Bandwidth ratio, $b$
Broadcast Strategy with Layered Source (BS)

- Source layers are transmitted over the channel using superposition.

\[ \text{n channel uses} \]

\[
\begin{align*}
\text{Base layer, R1 bits/ch. use} & \quad \{P_1\} \\
\text{Enhancement layer, R2 bits/ch. use} & \quad \{P_2\}
\end{align*}
\]

\[ P_1 + P_2 = \text{Total power P} \]

- Base layer is received when channel quality is low.
- If channel quality is better, both base and enhancement layers are received.
- Infinitely many layers for largest distortion exponent.
  - Rate and power allocation among the layers.
Distortion exponent of BS can be calculated as

$$\Delta_{BS} = \min(b, M_t M_r).$$

Can show $\Delta_{BS} = \Delta_{UB}$ and hence BS is optimal when

- $M_t = 1$ or $M_r = 1$: SIMO or MISO channel
- $b = 1$
Distortion Exponent of BS–MIMO

\[ M_t = 2, \ M_r = 2 \]

<table>
<thead>
<tr>
<th>Distortion exponent, ( \Delta )</th>
<th>Bandwidth ratio, ( b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>3.5</td>
<td>3.5</td>
</tr>
<tr>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>4.5</td>
<td>4.5</td>
</tr>
<tr>
<td>5.0</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Erkip 95/135
You can find details and further results in

- Gunduz, Erkip, IT 2008.
- Bhattad, Narayanan, Caire, IT 2008.
Assume half-duplex relay.
Source and relay each have one antenna.
Inter-user channels slowly fading.
No CSIT.
Relaying strategies:
  - Decode and forward (DF).
  - Amplify and forward (AF).
  - Dynamic decode and forward (DDF).
  - Can also combine with direct transmission (DT).
Relaying strategies used in conjunction with LS, HLS, BS.
Consider LS with one layer.

Depending on $b$, relaying or direct transmission is better.
- Low $b$: Direct
- High $b$: Relaying

More in Gunduz, Erkip, IT 2007
Scalable Compression: Road Ahead

- Scalable source-channel coding for the fading channel.
  - Multi-antenna single user channel.
  - Relay channel.
- Scalable video multicasting with relays.
- Distortion-memory tradeoffs in cache-aided multicast.
Goal: Transmit the same source to multiple receivers each having different channel quality.

Similar to fading channel with no CSIT.
- Now there are multiple receivers instead of multiple channel states.

Two strategies:
- Use half-duplex relays to improve channel quality.
- Use scalable source coding to deliver more to users with better channel qualities.

Consider video and practical constraints.
Conventional Multicast

- AP sends the same video to all subscribers in a coverage region.
  - AP transmits at the lowest transmission rate to reach users at the coverage edge.
  - Receivers with good channel quality unnecessarily suffer.
Cooperative Multicast

- Multicast using relays.
- AP targets receivers with good channel quality.
- These receivers relay video to others.
- Both links have better quality and sustain higher transmission rates.
How to Relay?

- Time-division among the relays: Too much loss in degrees of freedom.
- Relays should transmit simultaneously.
- Synchronization among distributed nodes is difficult.
- Relays can use randomized distributed space-time coding (R-DSTC)
  - DSTC allows relays to act as antennas of a MIMO transmitter. *(Laneman, Wornell, IT 2003)*
  - R-DSTC includes randomization that removes synchronization constraint. *(Sirkeci-Mergen, Scaglione, SP 2007).*
Use coding over the packets to minimize packet losses.
  - Video quality susceptible to packet losses.
  - Application layer forward error correction (FEC).

Use two source layers: Base and Enhancement.
  - Practical video coders incur a coding loss for more source layers.

Transmit layers in time: LS.

Use decode and forward relaying.
Cooperative Multicast with R-DSTC

- AP transmits a packet in the first hop
- Nodes that receive the packet become relays and forward simultaneously using RDSTC in the second hop.
Rates of both layers and time allocation are chosen so that:

- Base layer reaches all nodes with negligible loss probability.
- Base layer provides a target video quality.
- Enhancement layer reaches a target percentage of nodes.
- Base + Enhancement provide maximum video quality.
Simulation Set-Up

- IEEE 802.11g network
- Transmission rates: 6, 9, 12, 18, 24, 36, 48, 54 Mbps.
- BS transmit power set so that all users within 100m can receive at the base rate of 6Mbps.
- Real orthogonal STC of dimensions 2,4,8.
Relaying significantly improves over direct transmission (using base rate) and rate-adaptive direct transmission.

Gains higher when the number of nodes increases (more relays available).
Rate and PSNR when the system is configured so that 30% nodes receive both layers, 70% receive base layer only.
Trade-off between enhancement layer rate and the target percentage of users receiving enhancement layer.
You can find details and further results in

Scalable Compression: Road Ahead

- Scalable source-channel coding for the fading channel.
  - Multi-antenna single user channel.
  - Relay channel.
- Scalable video multicasting with relays.
- Distortion-memory tradeoffs in cache-aided multicast.
So far we have used scalability of the source to send multiple layers:
- To mitigate fading.
- To improve multicast performance when channel qualities vary among users.

Next we consider cache-aided multicast:
- All channels can sustain the same rate.
- But users have memories (possibly of different sizes) and can prefetch content.

To maximize quality at each user:
- What to cache?
- What/how to deliver?
Cache-Aided Multicast: Motivation

- Storage is cheap.
- Asynchronous user demands.
- Prefetch popular content → Reduce peak traffic rates
- Two phases:
  - Caching (Placement) Phase
    - Constraint: Size of cache memories
  - Delivery Phase
    - Constraint: Rate required to serve requests
- Library contains $m$ files.
- $n$ receivers, receiver $i$ has cache of size $M_i$.
- Shared communication link.

**Diagram:**
- Sender $S_1$
- m files
- Capacity $R$
- n receivers $U_1, U_2, \ldots, U_n$
- Cache $M_i$

**Question:** How can we best use the shared link and caches?
Uncoded Delivery

- $m = 2$ files, $A$ and $B$
- $n = 2$ users
- Cache sizes $M = 1$. 

![Diagram representing uncoded delivery with nodes A1, A2, B1, B2, and connections A1, B1, A1, B1.]

Erkip 117/ 135
Uncoded Delivery: Same Demand

\[ d = (A, A) \]

\[ R = \frac{1}{2} \]
Uncoded Delivery: Different Demands

\[ d = (A, B) \]

\[ R = 1 \]

- Multicast advantages **only** possible for users with **same** demand.
Coded Delivery: Same Demand

\[ d = (A, A) \]

\[ R = \frac{1}{2} \]
Coded Delivery: Different Demands

\[ d = (A, B) \]

\[ R = \frac{1}{2} \]

- Multicast advantages possible for all demands.
Different demands are satisfied with a single coded multicast transmission.

Careful content placement → Multicasting opportunities.
- Prefetching uncoded raw bits and delivering linearly encoded messages.

Rate $R$ required in delivery phase as a function of $M$. 
Library consists of files that need to be delivered completely.

Goal: Minimize expected link rate.
  - Ji, Tulino, Llorca, Caire, Arxiv 2015.

Scalable video, link rate given.

Goal: Maximize reconstruction quality.
Library consists of \( m \) Gaussian sources with different variances. Each user has a different demand distribution and cache size. Total shared link rate \( R \).

Goal: Minimize total expected distortion.

- Averaged over user demands and summed over users.
Local Caching Aided Unicast (LC-U)

- Caching: Done locally by receivers.
  - Decentralized.

- Delivery:
  - Only depends on local cache.
  - Unicast transmissions.
  - Server jointly optimizes rate allocation.

- Solution at each phase: A variation of the reverse-waterfilling algorithm.
Caching: Done based on global knowledge by sender.
   - Random popularity based (RAP) caching.

Delivery:
   - Depends on aggregate cache.
   - Coded multicast + Unicast transmissions
   - Based on chromatic number index coding (CIC)
   - Greedy constrained coloring (GCC): Polynomial implementation

Joint optimization of caching and transmission rates.
For implementation details and analytical results see:
Distortion-Memory Tradeoff: Symmetry Across Users

- \( m = 100 \) files with \( \sigma^2 \in [0.6, 1.7] \).
- \( n = 20 \) users.
- Demand \( \sim \text{Zipf}(\alpha = 0.6) \).
- $m = 100$ files with $\sigma^2 \in [0.6, 1.7]$.
- $n = 20$ users.
- Demand $\sim \text{Zipf}(\alpha = 0.6)$. 
Distortion-Memory Tradeoff: Symmetry Across Users and Files

- $m = 100$ files with $\sigma^2 = 1.5$.
- $n = 20$ users.
- Uniform demand on files.
Caching in Video Multicast

- Caching → reducing source distortion → enhancing video quality.
- Exploits:
  - Coded multicast opportunities.
  - Scalability of video.
- Channel assumptions:
  - All users have the same quality.
  - No losses in the channel.
- How can scalable source coding help when the channel is not ideal?
- Can we utilize correlations across files?
Exploiting source scalability helps mitigate:
- Losses due to channel variations.
- Multicast bottlenecks.

It also helps design more efficient ways to cache and deliver information.
Lecture Summary

- Exploiting source correlation.
  - Why correlation matters.
  - Shannon’s source-channel separation theorem in networks.
  - Informational versus operational separation.
  - Broadcast channel, multiple access channel, compound multiple access channel, interference channel, relay channel.

- Scalable compression.
  - Communication over fading channels.
  - Multicasting.
  - Multicasting with user caches.
• Simplified, but very useful!
MANY THANKS TO:

Ozgu Alay
Andrea Goldsmith
Deniz Gunduz
Parisa Hassanzadeh
Pei Liu
Jaime Llorca
Collaborators

MANY THANKS TO:

Chris Ng  Shiv Panwar  Vince Poor

Antonia Tulino  Yao Wang