Coding for Channels with Rate-limited Side Information

Yossef Steinberg

Department of Electrical Engineering
Technion—Israel Institute of Technology
Haifa 32000, Israel

The 2006 Information Theory Workshop—ITW ‘06:
Punta del Este, Uruguay, March 2006
Outline

- Problem formulation
- Motivation:
  - Communication Systems
  - Watermarking
- Previous work
- Main result
- Extensions and future work
Memoryless channel $P_{Y|X,S}(y|x,s)$ and state $P_S(s)$
Problem Formulation

- Memoryless channel $P_{Y|X,S}(y|x,s)$ and state $P_S(s)$
- State sequence $S^n$ known a priori at the encoder
Problem Formulation

- Memoryless channel $P_{Y|X,S}(y|x,s)$ and state $P_S(s)$
- State sequence $S^n$ known a priori at the encoder
- A compressed version of $S^n$, with rate $\text{rate}(S^n) \leq R_d$, is provided to the decoder.
Problem Formulation

- Memoryless channel $P_{Y|X,S}(y|x,s)$ and state $P_S(s)$
- State sequence $S^n$ known a priori at the encoder
- A compressed version of $S^n$, with rate $R_s^n \leq R_d$, is provided to the decoder.

We are interested in the region of all achievable rates and input costs:

$$R = \frac{\log |\mathcal{M}|}{n}, \quad R_d = \frac{\log |\mathcal{T}|}{n}, \quad \Gamma = E\phi(X^n).$$
Motivation

Communication systems:

OFDM + coding, where coding is done across frequencies. The sender knows channels states (fading), and sends it via a wayside channel to the receiver.

Watermarking (WM) with compressed host at the decoder.
Motivation – WM (cont’d)

![Diagram of watermarks and channels]

- **Public Watermarking** – The host data $S^n$ is available only at the encoder.
Motivation – WM (cont’d)

- **Public Watermarking** – The host data $S^n$ is available only at the encoder.

- **Private Watermarking** – The host data $S^n$ is available at both, encoder and decoder.
Motivation – WM (cont’d)

- **Public Watermarking** – The host data $S^n$ is available only at the encoder.

- **Private Watermarking** – The host data $S^n$ is available at both, encoder and decoder.

- A bridge between the versions [Moulin & O’Sullivan] – A key $K^n$ is present at the encoder and decoder, with a given $P_{S,K}$. 

---

Encoder $\xrightarrow{m} X^n$ Channel $\xrightarrow{Y^n} \text{Decoder}$

$S^n$, $K^n$ $Ed(S^n, X^n) \leq D$ $P_{Y|X}$
Motivation – WM (cont’d)

- **Public Watermarking** – The host data $S^n$ is available only at the encoder.

- **Private Watermarking** – The host data $S^n$ is available at both, encoder and decoder.

- A bridge between the versions [Moulin & O’Sullivan] – A key $K^n$ is present at the encoder and decoder, with a given $P_{S,K}$.
  - $K^n$ is provided to the decoder at no cost
Motivation – WM (cont’d)

- **Public Watermarking** – The host data $S^n$ is available only at the encoder.

- **Private Watermarking** – The host data $S^n$ is available at both, encoder and decoder.

- A bridge between the versions [Moulin & O’Sullivan] – A key $K^n$ is present at the encoder and decoder, with a given $P_{S,K}$.
  - $K^n$ is provided to the decoder at no cost
  - How to choose $P_{K|S}$?
**Motivation – WM (cont’d)**

- **Public Watermarking** – The host data $S^n$ is available only at the encoder.

- **Private Watermarking** – The host data $S^n$ is available at both, encoder and decoder.

- A bridge between the versions [Moulin & O’Sullivan] – A key $K^n$ is present at the encoder and decoder, with a given $P_{S,K}$.
  - $K^n$ is provided to the decoder at no cost
  - How to choose $P_{K|S}$?

$\Rightarrow$ Quantify the decoder’s a priori knowledge by a rate-limit
Problem:

Characterize the region of all achievable \((R, R_d, D)\), where:

- \(R\) – Embedding rate,
- \(R_d\) – rate of compressed SI @ decoder
- \(D\) – distortion between host and input.
$S^n$ is known noncausally at the encoder $\Rightarrow$ channel coding part is related to the Gel’fand-Pinsker (GP) problem.

$Y^n$ depends statistically on $S^n$ and can serve as side information (SI) in retrieving the compressed state at the decoder $\Rightarrow$ coding of $S^n$ is related to the Wyner-Ziv (WZ) problem.
For the WZ problem, the SI $Y^n$ is not memoryless

There is no distortion constraint in retrieving $S^n$ at the decoder (instead, maximize capacity of the main channel)
Previous work

- Wyner & Ziv, 1976
- Gel’fand & Pinsker, 1980
- Heegard & El Gamal, 1983, "On the capacity of computer memory with defects." Introduced coding for state dependent channels with rate limited side information at both ends. Devised an achievable region.

The current model is a special case of Heegard & El Gamal’s model.
The Heegard & El Gamal model:

Heegard & El Gamal devised an achievable region, tight for the cases:

1. $R_e = 0, \quad R_d = 0$
2. $R_e = H(S), \quad R_d = H(S|Y)$ \hspace{1cm} (both sides fully informed)
3. $R_e = H(S), \quad R_d = 0$ \hspace{1cm} (the GP model)
4. $R_e$ arbitrary, $R_d = H(S|Y)$ \hspace{1cm} (rate-limited SI @ encoder, fully informed decoder).

Case 4 was treated also by Rosenzweig et al, 2005. Dual to the problem treated here.
Previous work (cont’d)

The Heegard & El Gamal model:

Case 4. $R_e$ arbitrary, $R_d = H(S|Y)$ (rate-limited SI @ encoder, fully informed decoder).

$$R \leq I(X;Y|S, S_e)$$
$$R_e \geq I(S : S_e)$$

for some $S_e$ such that $X \leftarrow S_e \leftarrow S \quad S_e \leftarrow (S, X) \leftarrow Y$
Previous work (cont’d)

Works related to WM: (very partial list)


- Willems & Kalker, 2002 – WM system without attack channel. Two new ingredients:
  - The host $S^n$ is reconstructed within distortion $D_2$ at the decoder
  - Composite rate limit: a rate limit is put on the data set $X^n$. (Huffman code.)

- Maor & Merhav, 2005a, 2005b – Extended Willems & Kalker work: (a) general lossless codes, (b) attack channel.
Main result

\( \mathcal{R}^* \) – collection of all \((R, R_d, \Gamma)\) satisfying

\[
R \leq I(U; Y|S_d) - I(U; S|S_d)
\]
\[
R_d \geq I(S; S_d) - I(Y; S_d)
\]
\[
\Gamma \geq \mathbb{E}\phi(X)
\]

for some \((U, S_d)\) such that \((U, S_d) \looparrowright (S, X) \looparrowright Y\). Then

**Theorem:** For any discrete memoryless state-dependent channel, with full noncausal SI at the transmitter, and rate-limited SI at the receiver, a triple \((R, R_d, \Gamma)\) is achievable if and only if \((R, R_d, \Gamma) \in \mathcal{R}^*\).
Main result (cont’d)

\( \mathcal{R}^* \) – collection of all \((R, R_d, \Gamma)\) satisfying

\[
R \leq I(U; Y|S_d) - I(U; S|S_d)
\]
\[
R_d \geq I(S; S_d) - I(Y; S_d)
\]
\[
\Gamma \geq \mathbb{E}\phi(X)
\]

for some \((U, S_d)\) such that \((U, S_d) \circ (S, X) \circ Y\).

- \( S_d \) – A WZ rv, represents the compressed state \( S^n \). Fully decoded, with \( Y^n \) as SI.

- \( U \) – A GP rv, represents the encoded message. Fully decoded conditioned on \( S_d \) in both sides.
Main result (cont’d)

\( \mathcal{R}^* \) – collection of all \((R, R_d, \Gamma)\) satisfying

\[
\begin{align*}
R & \leq I(U; Y|S_d) - I(U; S|S_d) \\
R_d & \geq I(S; S_d) - I(Y; S_d) \quad (*) \\
\Gamma & \geq \mathbb{E}\phi(X)
\end{align*}
\]

for some \((U, S_d)\) such that \((U, S_d) \circledast (S, X) \circledast Y\).

- \((U, S_d) \circledast (S, X) \circledast Y\) does not imply \(S_d \circledast S \circledast Y\). Therefore \((*)\) is not equivalent to

\[
R_d \geq I(S; S_d|Y),
\]

full duality with GP.

- In classical WZ, \(S_d \circledast S \circledast Y\) is needed to guarantee joint typicality of \(S_d\) and \(Y\). Here it is guaranteed due to the channel.
Main result (cont’d)

$\mathcal{R}^*$ – collection of all $(R, R_d, \Gamma)$ satisfying

\[
\begin{align*}
R & \leq I(U; Y|S_d) - I(U; S|S_d) \\
R_d & \geq I(S; S_d) - I(Y; S_d) \\
\Gamma & \geq \mathbb{E}\phi(X)
\end{align*}
\]

for some $(U, S_d)$ such that $(U, S_d) \varnothing (S, X) \varnothing Y$.

Properties of $\mathcal{R}^*$

- $\mathcal{R}^*$ is convex
- $X = f(U, S_d, S)$, $f$ deterministic, suffices to exhaust $\mathcal{R}^*$. 
A typical \((R, R_d)\) curve

A typical \((R, R_d)\) curve, for fixed \(\Gamma\):

\[
\max \left[ I(U; Y) - I(U; S) \right]
\]

(Gelfand-Pinsker capacity with \(E\phi(X) \leq \Gamma\))
The rate allocated to provide the decoder with SI, is always \textit{at least as high} as the gain in the forward rate.

Provide SI to the decoder when the wayside channel cannot be used to transmit data – e.g.

- Remotely located physical channel
- WM, where a compressed host is kept in memory at the decoder, for future use.
Future work

- Extensions to networks
  - MAC, BC, etc
  - Ad hoc networks. Part of the users are silent, and can transmit SI at low cost.
- Specific models. Coding schemes.
- Computational algorithms.