Two-Way Communication with Adaptive Data Acquisition

B. Ahmadi and O. Simeone, NJIT

ISIT, 2013
Motivation

Two-Way Communication

- Forward link

Node 1

Information + query

Node 2

Database server

B. Ahmadi and O. Simeone, NJIT
Motivation

Two-Way Communication

- Adaptive data acquisition

Cost-constrained actions

Inference at Node 2
Motivation

Two-Way Communication

- Backward link

Inference at Node 1

Goal: Characterize the achievable rate-distortion-cost region
Motivation

Two-Way Communication

- Backward link

Inference at Node 1

Goal: Characterize the achievable rate-distortion-cost region
Prior Work

- Adaptive data acquisition [Permuter and Weissman ’11]
- Two-way source coding [Kaspi ’85]
**Adaptive Data Acquisition**  
[Permuter and Weissman '11]

- Forward link + adaptive data acquisition: source coding with “controllable” side information

\[
X^n \xrightarrow{M} \widehat{X}^n
\]

- Node 1:
  - \(X^n \xrightarrow{p(y|a,x)} p(y|a,x)\)

- Node 2:
  - \(A^n\)
  - \(Y^n\)

- Action \(A^n(M)\) is cost constrained as \(\frac{1}{n} \sum_{i=1}^{n} E[\Lambda(A_i)] \leq \Gamma\)
Adaptive Data Acquisition
[Permuter and Weissman ’11]

Rate-distortion-cost function

\[ R(D, \Gamma) = \min_{p(a,u|x), f(U, Y)} I(X; A) + I(X; U|A, Y) \]

subject to \( E[\Lambda(A)] \leq \Gamma \) and \( E[d(X, f(U, Y))] \leq D_1 \)

Result extends to adaptive actions \( A^n(M, Y^{i-1}) \) [Choudhuri and Mitra ’12]
Adaptive Data Acquisition
[Permuter and Weissman ’11]

- Layered coding: first layer given by action sequence $A$, second layer using Wyner-Ziv (binning) via $U$
Adaptive Data Acquisition

Extensions

- [Kittichokechai et al. '11] Common reconstruction constraint ([Steinberg '09])
- [Chia et al. '11] [Ahmadi and Simeone '11] Heegard-Berger-Kaspi problem
- [Kittichokechai et al. '11] Secrecy constraints
- [Zhao et al. '12] Compression with actions
- [Ahmadi and Simeone '12] Distributed and cascade source coding problems
Two-Way Source-Coding

[Kaspi '85]

Generalization to any number of rounds in [Kaspi '85]
Two-Way Source-Coding

[Kaspi '85]

Theorem

The rate-distortion-cost region \( R(D_1, D_2, \Gamma) \) is given by the union of all rate pairs \( (R_1, R_2) \) that satisfy the conditions

\[
R_1 \geq I(X; U | Y)
\]

and
\[
R_2 \geq I(Y; V | X, U),
\]

with

\[
p(x, y, a, u, v) = p(x, y)p(u | x)p(v | u, y),
\]

such that

\[
E[d_1(X, Y, f_1(V, X))] \leq D_1,
\]

and
\[
E[d_2(X, Y, f_2(U, Y))] \leq D_2.
\]
Two-Way Source Coding with Adaptive Data Acquisition

System Model

- $X^n, Z^{n_i,i.d.} \sim p(x,y)$
- Node 1 observes a noisy version of $X^n$
- Action sequence $A^n(M_1)$ based on message $M_1$ of Node 1
- Side information $Y^n|\{A^n = a^n, X^n = x^n, Z^n = z^n\} \sim p(y^n|a^n, x^n, z^n)$
  $$p(y^n|a^n, x^n, z^n) = \prod_{i=1}^{n} p(y_i|a_i, x_i, z_i)$$
Action cost constraint: \( \frac{1}{n} \sum_{i=1}^{n} \mathbb{E} [\Lambda(A_i)] \leq \Gamma \)

Distortion constraints

\[
\frac{1}{n} \sum_{i=1}^{n} \mathbb{E} \left[ d_j(X_i, Y_i, Z_i, \hat{X}_{ji}) \right] \leq D_j \text{ for } j = 1, 2.
\]

Rate-distortion-cost region \( \mathcal{R}(D_1, D_2, \Gamma) \)
Two-Way Source Coding with Adaptive Data Acquisition

System Model

- $D_1 = D_{1,\text{max}}$: [Permuter and Weissman '11]
- $p(y|a, x, z) = p(y|x, z)$: [Kaspi '85]
The rate-distortion-cost region $\mathcal{R}(D_1, D_2, \Gamma)$ is given by the union of all rate pairs $(R_1, R_2)$ that satisfy the conditions

$$R_1 \geq I(Z; A) + I(Z; U|A, Y)$$
and
$$R_2 \geq I(Y; V|A, Z, U),$$

with

$$p(x, y, z, a, u, v) = p(x, z)p(a, u|z)p(y|a, x, z)p(v|a, u, y),$$

such that

$$E[d_1(X, Y, Z, f_1(V, Z))] \leq D_1,$$
$$E[d_2(X, Y, Z, f_2(U, Y))] \leq D_2,$$
and
$$E[\Lambda(A)] \leq \Gamma,$$
**Two-Way Source Coding with Adaptive Data Acquisition**

**Sketch of Proof for Achievability**

- **Node 1**: Layered coding with first layer given by action $A$ and second layer using Wyner-Ziv (binning) via $U$
- **Node 2**: Wyner-Ziv (binning) via $V$
- **Decoder (Node 1)**: estimate $\hat{X}_1^n$ via sample by sample function $\hat{X}_{1i} = \hat{x}_1(V_i, Z_i)$
- **Decoder (Node 2)**: estimate $\hat{X}_2^n$ via sample by sample function $\hat{X}_{2i} = \hat{x}_2(U_i, Y_i)$
Case Study

- $X \sim \text{Bern}(0.5)$ (e.g., high/low concentration of a chemical)
- $Z_i = e$ with probability $\epsilon$, and $Z_i = X_i$ with probability $1 - \epsilon$ (e.g., $e$ represents malfunctioning of unusual events)

- Side information $Y_i$ is such that

\[
Y = \begin{cases} 
X & \text{for } A = 1 \\
\phi & \text{for } A = 0 
\end{cases}
\]

with $\Lambda(A) = A$
Case Study

- $X \sim \text{Bern}(0.5)$ (e.g., high/low concentration of a chemical)
- $Z_i = e$ with probability $\epsilon$, and $Z_i = X_i$ with probability $1 - \epsilon$ (e.g., $e$ represents malfunctioning of unusual events)

Side information $Y_i$ is such that

$$Y = \begin{cases} X & \text{for } A = 1 \\ \phi & \text{for } A = 0 \end{cases} ,$$

with $\Lambda(A) = A$
Node 1 wishes to reconstruct the source $X^n$

Node 2 is interested in recovering $Z^n$

E.g., Node 2 has the double purpose of: i) monitoring the operation of Node 1; ii) assisting Node 1 in case of measurement failures (erasures)
Case Study

\[ D_1 = D_{1,\text{max}} \text{ and } D_2 = 0 \]

- E.g., only monitoring of Node 1 by Node 2

\[
R_1 \geq H_2(\epsilon) + (1 - \epsilon - \Gamma)^+ \\
\text{and } R_2 \geq 0
\]

- Rate \( H_2(\epsilon) \) to describe the erasures process \( E^n \) with \( E_i = 1\{Z_i = e\} \)
- Rate \( (1 - \epsilon - \Gamma)^+ \) to describe a fraction of the \( n(1 - \epsilon) \) non-erased values of \( Z^n \)
- The remaining \( n_{\text{min}}(\Gamma, 1 - \epsilon) \) are measured by Node 2
Case Study

\( D_1 = 0 \) and \( D_2 = D_{2,\text{max}} \)

- E.g., only measurement of \( X^n \) at Node 1
- If \( \Gamma \geq \epsilon \)

\[
R_1 \geq H_2(\epsilon) - \Gamma H_2\left(\frac{\epsilon}{\Gamma}\right)
\]

and \( R_2 \geq \epsilon \)

- If \( \Gamma < \epsilon \), lossless reconstruction of \( X \) at Node 1 is not feasible
- Time-sharing strategy is suboptimal
Case Study

\( D_1 = D_2 = 0 \)

- If \( \Gamma \geq \epsilon \), we have
  \[ R_1 \geq H_2(\epsilon) + (1 - \Gamma) \]
  and \( R_2 \geq \epsilon \)

- Rate \( H_2(\epsilon) \) to describe the erasures process \( E^n \) with \( E_i = 1_{\{Z_i = e\}} \)
- Rate \( (1 - \Gamma) \) to describe a fraction of the \( n(1 - \Gamma) \leq n(1 - \epsilon) \) non-erased values of \( Z^n \)
- The remaining \( n\text{mi}(\Gamma - \epsilon) \) are measured by Node 2
- Rate \( \epsilon \) to compress the \( n\epsilon \) erased samples of \( X_i \)
Case Study

\[ D_1 = D_{1,\text{max}}, \quad D_2 = 0 \]

\[ D_1 = 0, \quad D_2 = D_{2,\text{max}} \] (time-sharing)

\[ D_1 = 0, \quad D_2 = 0 \]
Conclusions

- Two-way source coding problem with adaptive data acquisition
- Information transfer, query and data acquisition
- Rate-distortion characterization extending [Kaspi '85] [Permuter Weissman '11]
- Open problems: multiple rounds; multiple sources or multiple receivers with separate/interactive data acquisition