Problems in the intersection of Information theory and Control

Achilleas Anastasopoulos
anastas@umich.edu

EECS Department
University of Michigan

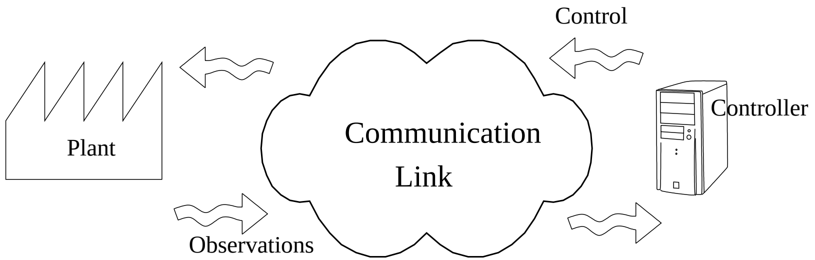
Dec 5, 2013

- Former PhD student: Junghuyn Bae
- Current PhD students: Deepanshu Vasal, Jui Wu

- Thanks to Demos Teneketzis and his former students Aditya Mahajan and Ashutosh Nayyar

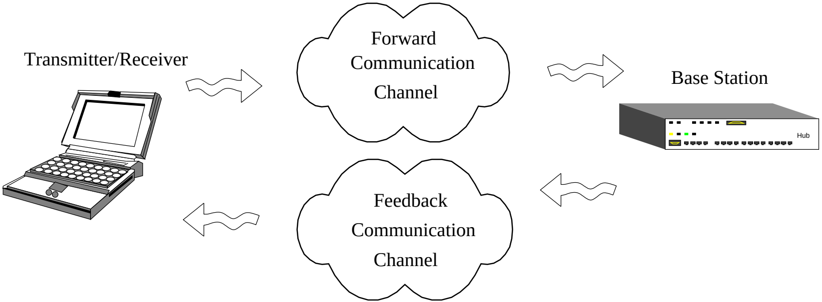
What this talk is NOT about

Controlling of a plant from a distance through a communication link



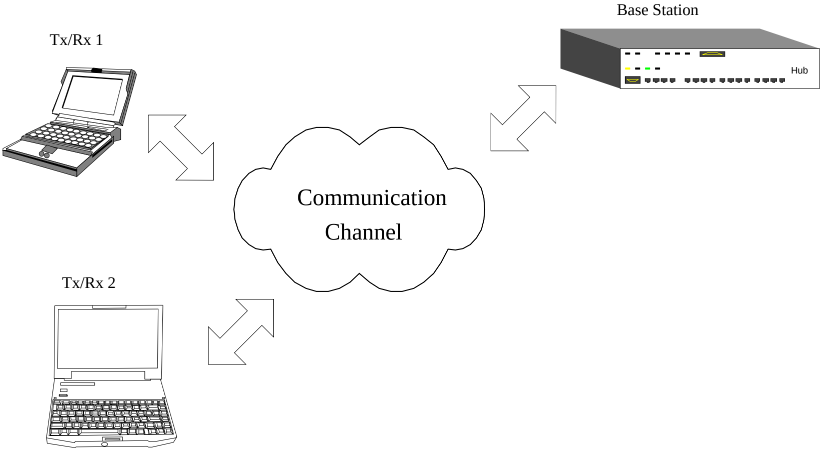
Instead, we are interested in

A) Viewing point-to-point communication as a Control problem



The act of transmitting a signal (partially) controls the overall communication system, with the hope of bringing it to a "desirable" state

B) Viewing multi-agent communications as a Control problem



Multiple agents (partially) control a communication network to bring it to a state beneficial for all (cooperatively/competitively)

C) More subtle: Viewing off-line optimization problems relevant to Information theory as control problems, e.g., Shannon capacity

$$C = \sup_{\{P_{X_t|X^{t-1},Y^{t-1}}(\cdot|\cdot,\cdot)\}_t} \frac{1}{T}\sum_{t=1}^{T} I(X_t \wedge Y_t|Y^{t-1})$$

Only connection to Communications is the problem origin.

No clear connection to Control either.
Where is the controller? where is the plant? what is the observation/control action?

Connections between Communication and Control

- Whenever there is feedback there is an intimate relation between Communication and Control
- One possible classification of problems
  1. Use Control techniques to design transmission schemes that achieve general performance measures (e.g., real-time coding[1] $\rightarrow$ examples A,B)
  2. Use Control techniques to design transmission schemes that achieve certain information-theoretic-inspired measures (e.g., capacity, error exponents[2] $\rightarrow$ examples A,B)
  3. Use Control techniques to evaluate information-theoretic quantities (e.g., capacity, error exponents[3] $\rightarrow$ example C)

---

[1] [Walrand and Varaiya, 1983, Mahajan and Teneketzis, 2009]

[2] [Horstein, 1963, Schalkwijk and Kailath, 1966, Shayevitz and Feder, 2008, Bae and Anastasopoulos, 2012]

[3] [Chen and Berger, 2005, Tatikonda and Mitter, 2009, Kavcic et al., 2009, Bae and Anastasopoulos, 2010]

Connections between Communication and Control

- Whenever there is feedback there is an intimate relation between Communication and Control
- One possible classification of problems
  1. Use Control techniques to design transmission schemes that achieve general performance measures (e.g., real-time coding[1] $\rightarrow$ examples A,B)
  2. Use Control techniques to design transmission schemes that achieve certain information-theoretic-inspired measures (e.g., capacity, error exponents[2] $\rightarrow$ examples A,B)
  3. Use Control techniques to evaluate information-theoretic quantities (e.g., capacity, error exponents[3] $\rightarrow$ example C)

---

[1] [Walrand and Varaiya, 1983, Mahajan and Teneketzis, 2009]

[2] [Horstein, 1963, Schalkwijk and Kailath, 1966, Shayevitz and Feder, 2008, Bae and Anastasopoulos, 2012]

[3] [Chen and Berger, 2005, Tatikonda and Mitter, 2009, Kavcic et al., 2009, Bae and Anastasopoulos, 2010]

Connections between Communication and Control

- Another possible classification of problems
  1. Problems involving a single controller (e.g., point-to-point transmission)
  2. Problems involving multiple controllers (e.g., multi-user transmission)
     - a) Agents act as members of a **team** trying to achieve a common goal
     - b) Agents act strategically having individual goals (**games**)

- Generally, dynamic problems with multiple agents and/or strategic interaction are more difficult: no standard solution methodology

- In this presentation we will discuss centralized/decentralized sequential team problems, and a sequential problem with strategic interaction (game)

Connections between Communication and Control

- Another possible classification of problems
  1. Problems involving a single controller (e.g., point-to-point transmission)
  2. Problems involving multiple controllers (e.g., multi-user transmission)
     - a) Agents act as members of a **team** trying to achieve a common goal
     - b) Agents act strategically having individual goals (**games**)

- Generally, dynamic problems with multiple agents and/or strategic interaction are more difficult: no standard solution methodology

- In this presentation we will discuss centralized/decentralized sequential team problems, and a sequential problem with strategic interaction (game)

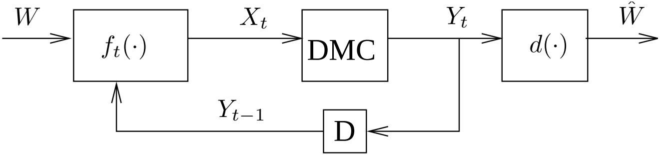
Overview

1. Point-to-point channels with memory and noiseless feedback
2. Multiple access channel with noiseless feedback
3. Cooperative communications in relay networks

Overview

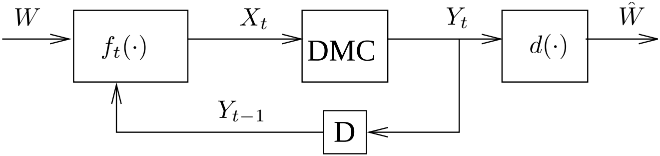
1. Point-to-point channels with memory and noiseless feedback
2. Multiple access channel with noiseless feedback
3. Cooperative communications in relay networks

Background: DMC with noiseless feedback



- Information message $W \in \{1, 2, \ldots, 2^{nR}\}$
- Transmitted symbols $X_t \in \mathscr{X}$, $t = 1, 2, \ldots, n$
- Received symbols $Y_t \in \mathscr{Y}$, $t = 1, 2, \ldots, n$
- Discrete-memoryless channel (DMC) defined by $Q(y_t|x_t)$
- Message estimate $\hat{W} \in \mathscr{W}$
- Encoding functions $X_t = f_t(W, Y^{t-1})$, $t = 1, 2, \ldots, n$
- Decoding function $\hat{W} = d(Y^n)$

Background: DMC with noiseless feedback



Fact[Shannon]: feedback does not increase capacity!

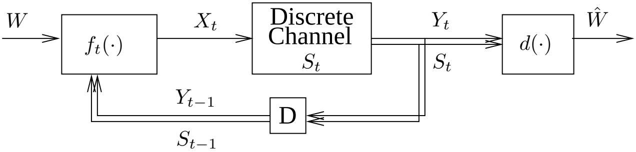
Capacity given by an **off-line, static, single-letter** optimization problem over distributions on $\mathscr{X}$

$$C = \sup_{P_X(\cdot)} \quad I(X_t \wedge Y_t),$$

with mutual information evaluated as

$$I(X_t \wedge Y_t) \stackrel{\text{def}}{=} \sum_x \sum_y P_X(x) Q(y|x) \log \frac{Q(y|x)}{\sum_{x'} P_X(x') Q(y|x')}$$

Channel with memory and noiseless feedback



- Information message $W \in \{1,2,\ldots,2^{nR}\}$
- Transmitted symbols $X_t \in \mathscr{X}$, $t = 1,2,\ldots,n$
- Channel state $S_t \in \mathscr{S}$, $t = 1,2,\ldots,n$
- Received symbols $Y_t \in \mathscr{Y}$, $t = 1,2,\ldots,n$ and channel state (perfect Rx CSI)
- Finite state channel (FSC) defined by $Q_y(y_t|x_t,s_t)$, $Q_s(s_{t+1}|s_t,x_t)$
- Message estimate $\hat{W} \in \mathscr{W}$
- Encoding functions $X_t = f_t(W,Y^{t-1},S^{t-1})$, $t = 1,2,\ldots,n$ (delayed Tx CSI)
- Decoding function $\hat{W} = d(Y^n,S^n)$

FSC capacity

Capacity of this channel is the result of the following **off-line** optimization problem[4] over infinitely many conditional distributions on $\mathscr{X}$

$$C = \sup_{\{P_{X_t|X_{t-1},S^{t-1},Y^{t-1}}\}_{t=1}^{\infty}} \liminf_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} I(X_t, X_{t-1} \wedge S_t, Y_t | S^{t-1}, Y^{t-1}).$$

- Observe: $P_{X_t|X_{t-1},S^{t-1},Y^{t-1}} \in \mathscr{X} \times \mathscr{S}^{t-1} \times \mathscr{Y}^{t-1} \to \mathscr{P}(\mathscr{X})$, so its domain increases with t
- **Q**: How can we utilize Control theory to solve this problem?

[4] [Tatikonda and Mitter, 2009, Bae and Anastasopoulos, 2010]

Parenthesis: Markov Decision Processes in 5 mins

A Markov Decision Process (MDP) is a random process with:

State $S_t \in \mathscr{S}$,

Control action $U_t \in \mathscr{U}$,

Instantaneous reward $R_t \in \mathscr{R}$,

defined by the following dynamics

$$P(S_{t+1}|S^t, U^t, R^t) = Q_s(S_{t+1}|S_t, U_t)$$
$$P(R_t|S^t, U^t, R^{t-1}) = Q_r(R_t|S_t, U_t)$$
$$U_t = g_t(S^t)$$

Problem: Design the sequence of mappings $g = \{g_t\}_t$ to maximize the average reward

$$J(g) \stackrel{\text{def}}{=} \mathbf{E}^g\{\sum_t R_t\}$$

Parenthesis: Markov Decision Processes in 5 mins

- A **single** controler observes **perfectly** the state and takes an action (centralized control with perfect state observation)

- **Solution**: Optimal control policy is Markov, i.e.,

$$U_t = g_t^*(S^t) = g_t^*(S_t)$$

- **Interpretation**: If state is perfectly observed by single controller, then it perfectly summarizes the entire history of observations

Parenthesis: MDPs and POMDPs in 5 mins

A Partially Observed MDP (POMDP) is a random process with:

State	$S_t \in \mathscr{S}$,
Observation	$Y_t \in \mathscr{Y}$,
Control action	$U_t \in \mathscr{U}$,
Instantaneous reward	$R_t \in \mathscr{R}$,

defined by the following dynamics

$$P(S_{t+1}|S^t, U^t, R^t, Y^t) = Q_s(S_{t+1}|S_t, U_t)$$
$$P(Y_t|S^t, U^{t-1}, R^{t-1}, Y^{t-1}) = Q_y(Y_t|S_t)$$
$$P(R_t|S^t, U^t, R^{t-1}, Y^t) = Q_r(R_t|S_t, U_t)$$
$$U_t = g_t(Y^t)$$

Problem: Design the sequence of mappings $g = \{g_t\}_t$ to maximize the average reward

$$J^g \stackrel{\text{def}}{=} \mathbf{E}^g\{\sum_t R_t\}$$

Parenthesis: MDPs and POMDPs in 5 mins

- **Solution**: Optimal control policy has the structure,

$$U_t = g_t^*(Y^t) = g_t^*(\Pi_t)$$

where $\Pi_t \in \mathscr{P}(\mathscr{S})$ and $\Pi_t(s) \stackrel{\text{def}}{=} Pr(S_t = s | U^{t-1}, Y^t) \quad \forall s \in \mathscr{S}$

- **Interpretation**: If state is imperfectly observed by controller, then the posterior belief of the state, Π_t, perfectly summarizes the entire history of observations
- **Takeaway:** MDPs/POMDPs are useful tools when we want to summarize the dependence of previous observations in our present decisions

Back to our problem: FSC capacity

$$C = \sup_{\{P_{X_t|X_{t-1},S^{t-1},Y^{t-1}}\}_{t=1}^{\infty}} \liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} I(X_t, X_{t-1} \wedge S_t, Y_t | S^{t-1}, Y^{t-1}).$$

- Think of C as the total average reward of some "fictitious" MDP with appropriate states, actions, instantaneous rewards, etc
- Hint:

$$I(X_t, X_{t-1} \wedge S_t, Y_t | S^{t-1}, Y^{t-1}) =$$

$$= \mathbf{E}\{\log \frac{P(S_t, Y_t | X_t, X_{t-1}, S^{t-1}, Y^{t-1})}{P(S_t, Y_t | S^{t-1}, Y^{t-1})}\}$$

$$= \mathbf{E}\{\log \frac{Q_y(Y_t|S_t, X_t) Q_s(S_t|S_{t-1}, X_{t-1})}{\sum_{x_t, x_{t-1}} Q_y(Y_t|S_t, x_t) Q_s(S_t|S_{t-1}, x_{t-1}) P(x_t|x_{t-1}, S^{t-1}, Y^{t-1}) P(x_{t-1}|S^{t-1}, Y^{t-1})}\}$$

Back to our problem: FSC capacity

Instantaneous reward depends on

a) some current variables, e.g., $S_{t-1}, S_t, X_{t-1}, X_t, Y_t$

b) the input distribution $P(x_t|x_{t-1}, S^{t-1}, Y^{t-1})$

c) the quantity $P(x_{t-1}|S^{t-1}, Y^{t-1})$

- Define the Control action $U_t \in \mathscr{X} \to \mathscr{P}(\mathscr{X})$ (conditional distribution of X_t given X_{t-1})
- Allow U_t to be a deterministic function of S^{t-1}, Y^{t-1}, $U_t = g_t(S^{t-1}, Y^{t-1})$
- Meaning: $U_t(x_{t-1})(\cdot) = g_t(S^{t-1}, Y^{t-1})(x_{t-1})(\cdot) = P(X_t = \cdot|X_{t-1} = x_{t-1}, S^{t-1}, Y^{t-1})$

FSC capacity

- Further define the r.v. (information state) $\Theta_t \in \mathscr{P}(\mathscr{X})$ with $\Theta_t(x) \stackrel{\text{def}}{=} Pr(X_t = x | S^t, Y^t), \quad \forall x \in \mathscr{X}$
- Average instantaneous reward becomes

$$I(X_t, X_{t-1} \wedge S_t, Y_t | S^{t-1}, Y^{t-1}) =$$

$$= \mathbf{E}\{\log \frac{Q_y(Y_t|S_t, X_t) Q_s(S_t|S_{t-1}, X_{t-1})}{\sum_{x_t, x_{t-1}} Q_y(Y_t|S_t, x_t) Q_s(S_t|S_{t-1}, x_{t-1}) U_t(x_{t-1})(x_t) \Theta_{t-1}(x_{t-1})}\}$$

- Observe: dependence on S^{t-1}, Y^{t-1} is "hidden" in the **generation** of $U_t = g_t(S^{t-1}, Y^{t-1})$ and the **evolution** of Θ_{t-1}

FSC capacity

Theorem ([Bae and Anastasopoulos, 2010])

The original optimization problem is equivalent to an MDP with

- *State* $\{(S_{t-1}, \Theta_{t-1})\}_t$
- *Control action* U_t

Markov policies are optimal, i.e., optimal actions can be of the form

$$U_t = g_t^*(S_{t-1}, \Theta_{t-1}) \Leftrightarrow P^*_{X_t|X_{t-1}, S^{t-1}, Y^{t-1}} = P^*_{X_t|X_{t-1}, S_{t-1}, \Theta_{t-1}}$$

Capacity is now simplified to a **single-letter** *expression*

$$C = \sup_{P_{X_t|X_{t-1}, S_{t-1}, \Theta_{t-1}}} I(X_t, X_{t-1} \wedge S_t, Y_t | S_{t-1}, \Theta_{t-1})$$

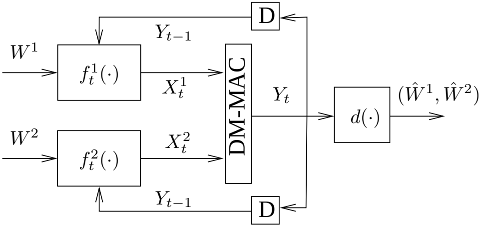
Lessons learned

- Complex optimization problems in Information theory can be translated to simple centralized control problems
- Starting point: some multi-letter capacity expression
- Methodology: appropriately define a control system to unveil an MDP/POMDP
- These ideas can also help in designing actual on-line capacity-achieving transmission schemes (not in this talk)

Overview

1. Point-to-point channels with memory and noiseless feedback
2. Multiple access channel with noiseless feedback
3. Cooperative communications in relay networks

System model: the information theoretic setup



- Messages $W^i \in \{1, 2, \ldots, 2^{nR^i}\}$, $i = 1, 2$
- Transmitted symbols $X_t^i \in \mathscr{X}^i$, $i = 1, 2$, $t = 1, 2, \ldots, n$
- Received symbols $Y_t \in \mathscr{Y}$, $t = 1, 2, \ldots, n$
- Discrete-memoryless MAC (DM-MAC) $Q(y_t | x_t^1, x_t^2)$
- Message estimates $(\hat{W}^1, \hat{W}^2) \in \mathscr{W}^1 \times \mathscr{W}^2$
- Encoding functions $X_t^i = f_t^i(W^i, Y^{t-1})$, $i = 1, 2$, $t = 1, 2, \ldots, n$
- Decoding function $(\hat{W}^1, \hat{W}^2) = d(Y^n)$

NLF-MAC capacity

- Feedback provides an enlargement of capacity region for the MAC [Gaarder and Wolf, 1975, Cover and Leung, 1981, Ozarow, 1984, Bross and Lapidoth, 2005, Venkataramanan and Pradhan, 2009]
- For the general NLF-MAC the capacity is not known as a single-letter expression!
- Multi-letter expression was developed by [Kramer, 2003]

NLF-MAC capacity: multi-letter expression

Fact ([Kramer, 2003], [Salehi, 1978])

The problem of evaluating the NLF-MAC capacity region is equivalent to solving the following optimization problem for every $\underline{\lambda} = (\lambda_1, \lambda_2, \lambda_3) \geq 0$

$$J_{\underline{\lambda}} = \sup \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \{\lambda_1 I(X_t^1 \wedge Y_t | X^{2,t}, Y^{t-1}) + \lambda_2 I(X_t^2 \wedge Y_t | X^{1,t}, Y^{t-1}) + \lambda_3 I(X_t^1, X_t^2 \wedge Y_t | Y^{t-1})\},$$

and the supremum is over all input distributions of the form

$$\{P(X_t^1 | X^{1,t-1}, Y^{t-1}),\ P(X_t^2 | X^{2,t-1}, Y^{t-1})\}_t$$

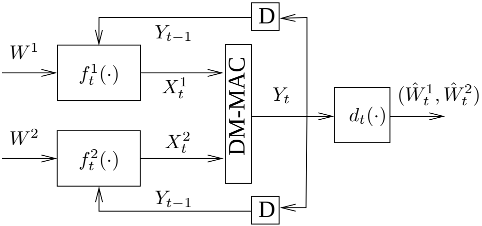
Why is the NLF-MAC capacity still an open problem?

- Three main difficulties
  1. The optimal input distributions on X_t^i depend on entire history $X^{i,t-1}$ and Y^{t-1}
  2. The optimization problem involves two controllers with different observations (decentralized control!)
  3. The per-stage rewards (mutual info expressions) are complicated functions of the involved random variables
- Claim: we can address the first two of the above three difficulties
- Solve a slightly different problem: real-time communication over the NLF-MAC

Why is the NLF-MAC capacity still an open problem?

- Three main difficulties
  1. The optimal input distributions on X_t^i depend on entire history $X^{i,t-1}$ and Y^{t-1}
  2. The optimization problem involves two controllers with different observations (decentralized control!)
  3. The per-stage rewards (mutual info expressions) are complicated functions of the involved random variables
- Claim: we can address the first two of the above three difficulties
- Solve a slightly different problem: real-time communication over the NLF-MAC

System model: real-time communication



Same model as before, except

- Message estimates for each time t, $(\hat{W}_t^1, \hat{W}_t^2) \in \mathscr{W}^1 \times \mathscr{W}^2$
- Decoding functions $(\hat{W}_t^1, \hat{W}_t^2) = d_t(Y^t)$, $t = 1, 2, \ldots, n$
- Instantaneous reward function $\rho_t(W^1, W^2, \hat{W}_t^1, \hat{W}_t^2)$
- Find a set of encoding/decoding functions $g \stackrel{\text{def}}{=} \{f_t^1, f_t^2, d_t\}_t$ that maximize

$$J(g) = \mathbf{E}^g\{\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \rho_t(W^1, W^2, \hat{W}_t^1, \hat{W}_t^2)\},$$

Problem statement: discussion

- Many reasonable choices for reward functions $\rho_t(\cdot)$, e.g.,

$$\rho_t(W^1, W^2, \hat{W}_t^1, \hat{W}_t^2) = \mathbf{1}_{W^1=\hat{W}_t^1 \text{ and } W^2=\hat{W}_t^2} \Rightarrow$$
$$\mathbf{E}\rho_t(W^1, W^2, \hat{W}_t^1, \hat{W}_t^2) = Pr(W^1 = \hat{W}_t^1 \text{ and } W^2 \neq \hat{W}_t^2)$$

- Focus on structural properties of the communication system that are common regardless of these choices.

- Salient features of the problem:
  1. Domain of encoding functions $X_t^i = f_t^i(W^i, Y^{t-1})$ increases with time.
  2. Existence of common information at encoders (Y^{t-1} at time t) and private information (W^i)
  3. Decentralized, non-classical information structure (this is **not** a MDP/POMDP-like problem!)

Problem statement: discussion

- Many reasonable choices for reward functions $\rho_t(\cdot)$, e.g.,

$$\rho_t(W^1, W^2, \hat{W}_t^1, \hat{W}_t^2) = \mathbf{1}_{W^1=\hat{W}_t^1 \text{ and } W^2=\hat{W}_t^2} \Rightarrow$$
$$\mathbf{E}\,\rho_t(W^1, W^2, \hat{W}_t^1, \hat{W}_t^2) = Pr(W^1 = \hat{W}_t^1 \text{ and } W^2 \neq \hat{W}_t^2)$$

- Focus on structural properties of the communication system that are common regardless of these choices.

- Salient features of the problem:
  1. Domain of encoding functions $X_t^i = f_t^i(W^i, Y^{t-1})$ increases with time.
  2. Existence of common information at encoders (Y^{t-1} at time t) and private information (W^i)
  3. Decentralized, non-classical information structure (this is **not** a MDP/POMDP-like problem!)

Introduction of pre-encoder[5]

- Equivalent encoder description:
  Each user's transmission $X_t^i = f_t^i(W^i, Y^{t-1})$ can be thought of as a two-stage process

  1. Based on available feedback Y^{t-1} select encoding functions
  $$E_t^i : \mathscr{W}^i \to \mathscr{X}^i, \qquad i = 1, 2,$$
  through a pre-encoder mapping
  $$(E_t^1, E_t^2) = h_t(Y^{t-1}).$$
  2. Generate transmitted signals by evaluating the encoding functions at W^i, i.e.,
  $$X_t^i = E_t^i(W^i), \qquad i = 1, 2.$$

[5] [Walrand and Varaiya, 1983, Nayyar and Teneketzis, 2008]

Introduction of pre-encoder[5]

- Equivalent encoder description:
  Each user's transmission $X_t^i = f_t^i(W^i, Y^{t-1})$ can be thought of as a two-stage process

  1. Based on available feedback Y^{t-1} select encoding functions

  $$E_t^i : \mathscr{W}^i \to \mathscr{X}^i, \qquad i = 1, 2,$$

  through a pre-encoder mapping

  $$(E_t^1, E_t^2) = h_t(Y^{t-1}).$$

  2. Generate transmitted signals by evaluating the encoding functions at W^i, i.e.,

  $$X_t^i = E_t^i(W^i), \qquad i = 1, 2.$$

[5][Walrand and Varaiya, 1983, Nayyar and Teneketzis, 2008]

Introduction of pre-encoder[5]

- Equivalent encoder description:
  Each user's transmission $X_t^i = f_t^i(W^i, Y^{t-1})$ can be thought of as a two-stage process

  1. Based on available feedback Y^{t-1} select encoding functions

  $$E_t^i : \mathscr{W}^i \to \mathscr{X}^i, \qquad i = 1, 2,$$

  through a pre-encoder mapping

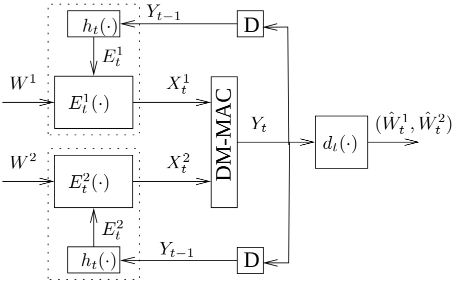
  $$(E_t^1, E_t^2) = h_t(Y^{t-1}).$$

  2. Generate transmitted signals by evaluating the encoding functions at W^i, i.e.,

  $$X_t^i = E_t^i(W^i), \qquad i = 1, 2.$$

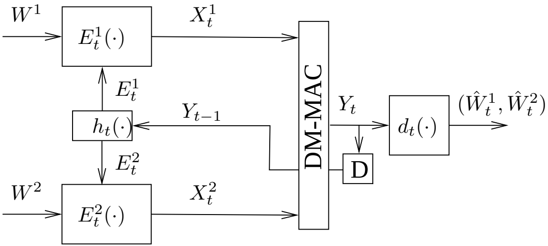
[5] [Walrand and Varaiya, 1983, Nayyar and Teneketzis, 2008]

Introduction of pre-encoder



- Decentralization of information is imposed by design (h_t only uses the common information Y^{t-1} available to both encoders)
- Both encoders can evaluate each-other's encoding functions through $(E_t^1, E_t^2) = h_t(Y^{t-1})$ (can be thought of as a *fictitious* coordinator)

Transforming to a centralized control problem



- The control problem boils down to selecting encoding functions $(E_t^1, E_t^2) = h_t(Y^{t-1})$. Generation of X_t^i is a "dumb" function evaluation $X_t^i = E_t^i(W^i)$
- New equivalent design $g \stackrel{\text{def}}{=} \{f_t^1, f_t^2, d_t\}_t \Rightarrow \tilde{g} \stackrel{\text{def}}{=} \{h_t, d_t\}_t$
- Above transformation still suffers from increasing domain $\mathscr{Y}^{t-1}$ of the pre-encoder h_t, i.e., $(E_t^1, E_t^2) = h_t(Y^{t-1})$.

Introduction of information state

- We would like to summarize Y^{t-1} in a quantity (state) with time invariant domain
- Related attempts:
  1. Introduction of auxiliary variables in information theory (e.g., [Cover and Leung, 1981, Bross and Lapidoth, 2005])
  2. Form a *graph* describing the correlation structure of the messages after receiving Y^{t-1} [Venkataramanan and Pradhan, 2009]
- A more direct approach: define the random quantities

$$\Pi_t \in \mathscr{P}(\mathscr{W}^1 \times \mathscr{W}^2), \qquad t = 0, 1, 2, \ldots$$

as

$$\Pi_t(w^1, w^2) \stackrel{\text{def}}{=} Pr(W^1 = w^1, W^2 = w^2 | Y^t),$$

i.e., the posterior distribution of the message pair given the observation.

Introduction of information state

- We would like to summarize Y^{t-1} in a quantity (state) with time invariant domain
- Related attempts:
  1. Introduction of auxiliary variables in information theory (e.g., [Cover and Leung, 1981, Bross and Lapidoth, 2005])
  2. Form a *graph* describing the correlation structure of the messages after receiving Y^{t-1} [Venkataramanan and Pradhan, 2009]
- A more direct approach: define the random quantities

$$\Pi_t \in \mathscr{P}(\mathscr{W}^1 \times \mathscr{W}^2), \qquad t = 0, 1, 2, \ldots$$

as

$$\Pi_t(w^1, w^2) \stackrel{\text{def}}{=} Pr(W^1 = w^1, W^2 = w^2 | Y^t),$$

i.e., the posterior distribution of the message pair given the observation.

Introduction of information state

- We would like to summarize Y^{t-1} in a quantity (state) with time invariant domain
- Related attempts:
  1. Introduction of auxiliary variables in information theory (e.g., [Cover and Leung, 1981, Bross and Lapidoth, 2005])
  2. Form a *graph* describing the correlation structure of the messages after receiving Y^{t-1} [Venkataramanan and Pradhan, 2009]
- A more direct approach: define the random quantities

$$\Pi_t \in \mathscr{P}(\mathscr{W}^1 \times \mathscr{W}^2), \qquad t = 0, 1, 2, \ldots$$

as

$$\Pi_t(w^1, w^2) \stackrel{\text{def}}{=} Pr(W^1 = w^1, W^2 = w^2 | Y^t),$$

i.e., the posterior distribution of the message pair given the observation.

Introduction of information state

Lemma

1. *The quantity* Π_t *can be recursively updated as*

$$\Pi_t = \Phi(\Pi_{t-1}, E_t^1, E_t^2, Y_t), \qquad t = 1, 2, \ldots$$

2. $(\Pi_t)_t$ *is a controlled Markov process with control action* (E_t^1, E_t^2)
3. *The optimal decoder function at time* t *is only a function of* Π_t

$$(\hat{W}_t^1, \hat{W}_t^2) = d_t^*(Y^t) = d_t^*(\Pi_t)$$

4. *The average instantaneous costs are functions of* Π_{t-1}, E_t^1, E_t^2*, i.e.,*

$$\mathbf{E}\{\rho_t(W^1, W^2, \hat{W}_t^1, \hat{W}_t^2)\} = \mathbf{E}\{\Psi_t(\Pi_{t-1}, E_t^1, E_t^2)\}.$$

where Ψ_t *are known functions.*

Main structural result

Theorem

The optimal communication system for the NLF-MAC consists of

1. *Encoders of the form* $X_t^i = E_t^i(W^i)$, $i = 1, 2$, *where*

$$(E_t^1, E_t^2) = h_t(\Pi_{t-1})$$

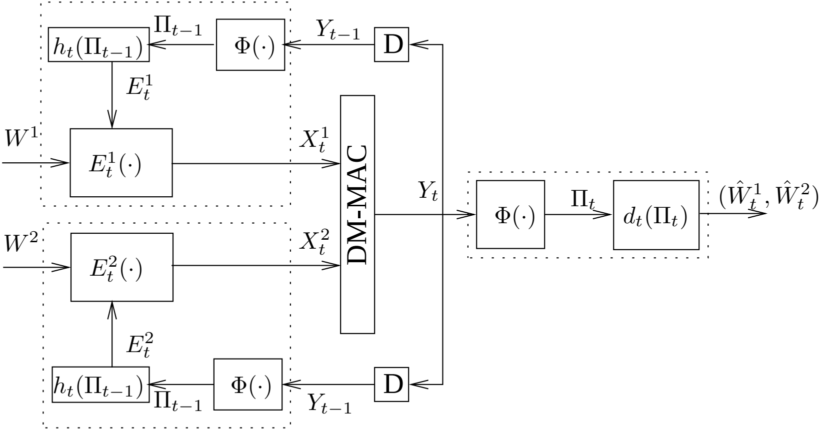
2. *A receiver that generates message estimates as*

$$(\hat{W}_t^1, \hat{W}_t^2) = d_t(\Pi_t),$$

   where d_t *is a known function.*

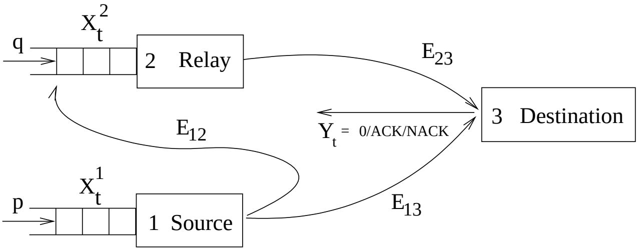
3. *The optimal* h_t *can be determined as the solution of a fix-point equation (dynamic program)*

Equivalent optimal communication system



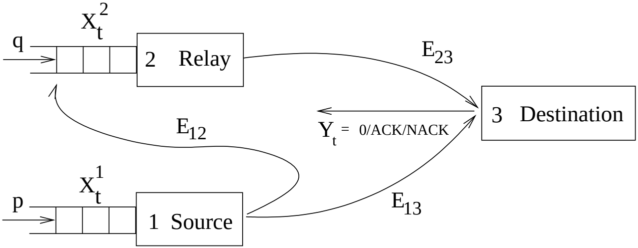
Overview

Single-relay network



- Bernoulli arrivals at Source (w.p. p) and at Relay (w.p. q)
- Packets waiting at Source's and Relay's queues $X_t = (X_t^1, X_t^2) \in \mathbb{N} \times \mathbb{N}$
- Actions $U_t^1 \in \mathscr{U}^1 \stackrel{\text{def}}{=} \{0, E_{13}, E_{12}\}$, $U_t^2 \in \mathscr{U}^2 \stackrel{\text{def}}{=} \{0, E_{23}\}$
- Simple collision model. Feedback $Y_t \in \{\emptyset, ACK, NACK\}$

Single-relay network



- Instantaneous costs are functions of energy and "delay"

$$C_t^i = \rho^i(X_t^i, U_t^i) \qquad (\text{e.g., } = X_t^i + U_t^i), \qquad i = 1,2$$

- Reasonable assumptions: $E_{12} + E_{23} < E_{13}$, $p + q < 1$, units either receive or transmit

Three scenaria of interest: Scenario A

- **Centralized** control of queues with **perfect** observation

$$(U_t^1, U_t^2) = f_t(X^{1,t}, X^{2,t}, Y^{t-1})$$

- Find a set of policies $f \stackrel{\text{def}}{=} \{f_t\}_t$ that minimize

$$J(f) = \mathbf{E}^f\{\sum_t \rho^1(X_t^1, U_t^1) + \rho^2(X_t^2, U_t^2)\}$$

- **Solution**: Centralized stochastic control problem. Can be formulated as an MDP. $(U_t^1, U_t^2) = f_t^*(X_t^1, X_t^2)$

Three scenaria of interest: Scenario A

- **Centralized** control of queues with **perfect** observation

$$(U_t^1, U_t^2) = f_t(X^{1,t}, X^{2,t}, Y^{t-1})$$

- Find a set of policies $f \stackrel{\text{def}}{=} \{f_t\}_t$ that minimize

$$J(f) = \mathbf{E}^f \{\sum_t \rho^1(X_t^1, U_t^1) + \rho^2(X_t^2, U_t^2)\}$$

- **Solution**: Centralized stochastic control problem. Can be formulated as an MDP. $(U_t^1, U_t^2) = f_t^*(X_t^1, X_t^2)$

Three scenaria of interest: Scenario B

- Decentralized control of queues: each agent i observes only his own queue length X_t^i and both agents have a common goal (team problem)

$$U_t^i = f_t^i(X^{i,t}, Y^{t-1}), \qquad i = 1, 2$$

Find a set of policies $f \stackrel{\text{def}}{=} (f_t^1, f_t^2)_t$ that minimize

$$J(f) = \mathbf{E}^f \{ \sum_t \rho^1(X_t^1, U_t^1) + \rho^2(X_t^2, U_t^2) \},$$

- Salient features of the problem:
  1. Domain of control mappings $U_t^i = f_t^i(X^{1,t}, Y^{t-1})$ increases with time.
  2. Presence of common information (Y^{t-1} at time t) and private information ($X^{i,t}$ at time t for agent i)
  3. Decentralized, non-classical information structure (this is **not** a MDP/POMDP-like problem!)

Three scenaria of interest: Scenario B

- Decentralized control of queues: each agent i observes only his own queue length X_t^i and both agents have a common goal (team problem)

$$U_t^i = f_t^i(X^{i,t}, Y^{t-1}), \qquad i = 1, 2$$

Find a set of policies $f \stackrel{\text{def}}{=} (f_t^1, f_t^2)_t$ that minimize

$$J(f) = \mathbf{E}^f\{\sum_t \rho^1(X_t^1, U_t^1) + \rho^2(X_t^2, U_t^2)\},$$

- Salient features of the problem:
  1. Domain of control mappings $U_t^i = f_t^i(X^{1,t}, Y^{t-1})$ increases with time.
  2. Presence of common information (Y^{t-1} at time t) and private information ($X^{i,t}$ at time t for agent i)
  3. Decentralized, non-classical information structure (this is **not** a MDP/POMDP-like problem!)

Structural results for the team problem

Lemma ([Vasal and Anastasopoulos, 2012])

1. *Knowledge of Y^{t-1} and $U^{i,t-1}$ reveals $U^{t-1} = (U^{1,t-1}, U^{2,t-1})$, so U^{t-1} is common knowledge (Y^{t-1} is not needed further)*

$$U_t^i = f_t^i(X^{i,t}, U^{t-1}), \qquad i = 1,2$$

2. *Optimal policy depends only on the current private state X_t^i*

$$U_t^i = f_t^i(X_t^i, U^{t-1}), \qquad i = 1,2$$

Still we have not addressed the decentralization issue and the expanding domain of f_t^i issue.

Introduction of pre-encoder[6]

Equivalent controller description:
Each agent's decision $U_t^i = f_t^i(X_t^i, U^{t-1})$ can be thought of as a two-stage process

1. Based on common info U^{t-1} select "prescription" functions $\Gamma_t^i : \mathbb{N} \to \mathscr{U}^i, \quad i = 1,2$ through the pre-encoder mapping

$$(\Gamma_t^1, \Gamma_t^2) = h_t(U^{t-1})$$

2. The actions U_t^i are determined by evaluating Γ_t^i at the private information X_t^i, i.e.,

$$U_t^i = \Gamma_t^i(X_t^i), \qquad i = 1,2$$

[6] [Walrand and Varaiya, 1983, Nayyar and Teneketzis, 2008]

Introduction of pre-encoder[6]

Equivalent controller description:
Each agent's decision $U_t^i = f_t^i(X_t^i, U^{t-1})$ can be thought of as a two-stage process

1. Based on common info U^{t-1} select "prescription" functions $\Gamma_t^i : \mathbb{N} \to \mathscr{U}^i, \quad i = 1,2$ through the pre-encoder mapping

$$(\Gamma_t^1, \Gamma_t^2) = h_t(U^{t-1})$$

2. The actions U_t^i are determined by evaluating Γ_t^i at the private information X_t^i, i.e.,

$$U_t^i = \Gamma_t^i(X_t^i), \qquad i = 1,2$$

[6] [Walrand and Varaiya, 1983, Nayyar and Teneketzis, 2008]

Introduction of pre-encoder[6]

Equivalent controller description:
Each agent's decision $U_t^i = f_t^i(X_t^i, U^{t-1})$ can be thought of as a two-stage process

1. Based on common info U^{t-1} select "prescription" functions $\Gamma_t^i : \mathbb{N} \to \mathscr{U}^i, \quad i = 1,2$ through the pre-encoder mapping

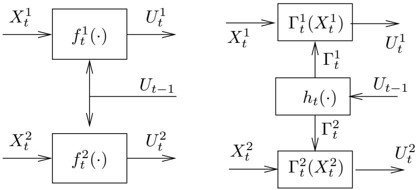
$$(\Gamma_t^1, \Gamma_t^2) = h_t(U^{t-1})$$

2. The actions U_t^i are determined by evaluating Γ_t^i at the private information X_t^i, i.e.,

$$U_t^i = \Gamma_t^i(X_t^i), \qquad i = 1,2$$

[6] [Walrand and Varaiya, 1983, Nayyar and Teneketzis, 2008]

Transformation to a centralized problem



- Generation of U_t^i is a "dumb" function evaluation $U_t^i = \Gamma_t^i(X_t^i)$
- The control problem boils down to selecting prescription functions $h \stackrel{\text{def}}{=} \{h_t\}_t$,
- Both agents can evaluate each-other's prescription functions through $(\Gamma_t^1, \Gamma_t^2) = h_t(U^{t-1})$ (can be thought of as a *fictitious* controller)
- The decentralized control problem has been transformed to a centralized control problem
- Last issue to address: increasing domain $\mathscr{U}^{t-1}$ of the pre-encoder mappings h_t.

Introduction of information state

- We would like to summarize U^{t-1} in a quantity (state) with time invariant domain
- Define the random quantities

$$\Pi_t \in \mathscr{P}(\mathbb{N} \times \mathbb{N}), \qquad t = 0, 1, 2, \ldots$$

as

$$\Pi_t(x_t^1, x_t^2) \stackrel{\text{def}}{=} P(X_t^1 = x_t^1, X_t^2 = x_t^2 | U^{t-1})$$

i.e., the posterior distribution of the queue lengths given the observation.

Introduction of information state

- We would like to summarize U^{t-1} in a quantity (state) with time invariant domain
- Define the random quantities

$$\Pi_t \in \mathscr{P}(\mathbb{N} \times \mathbb{N}), \qquad t = 0, 1, 2, \dots$$

as

$$\Pi_t(x_t^1, x_t^2) \stackrel{\text{def}}{=} P(X_t^1 = x_t^1, X_t^2 = x_t^2 | U^{t-1})$$

i.e., the posterior distribution of the queue lengths given the observation.

Main structural result

Theorem ([Vasal and Anastasopoulos, 2012])

The original decentralized control problem is equivalent to an MDP with

- *State* Π_t
- *Control actions* $\Gamma_t \stackrel{def}{=} (\Gamma_t^1, \Gamma_t^2)$
- *Instantaneous costs* $\mathbf{E}\{\rho_t^1(X_t^1, U_t^1) + \rho_t^2(X_t^2, U_t^2) | \Pi_t, \Gamma_t\}$

Markov policies are optimal, i.e., optimal actions can be of the form

$$\Gamma_t = h_t^*(\Pi_t) \qquad \Rightarrow \qquad U_t^i = f_t^{i*}(\Pi_t, X_t^i)$$

It turns out there is a further simplification: instead of joint posterior distributions, we can use the two marginals! A general version of this result in [Nayyar et al., 2011]

Main structural result

Theorem ([Vasal and Anastasopoulos, 2012])

The original decentralized control problem is equivalent to an MDP with

- *State* Π_t
- *Control actions* $\Gamma_t \stackrel{def}{=} (\Gamma_t^1, \Gamma_t^2)$
- *Instantaneous costs* $\mathbf{E}\{\rho_t^1(X_t^1, U_t^1) + \rho_t^2(X_t^2, U_t^2) | \Pi_t, \Gamma_t\}$

Markov policies are optimal, i.e., optimal actions can be of the form

$$\Gamma_t = h_t^*(\Pi_t) \qquad \Rightarrow \qquad U_t^i = f_t^{i*}(\Pi_t, X_t^i)$$

It turns out there is a further simplification: instead of joint posterior distributions, we can use the two marginals! A general version of this result in [Nayyar et al., 2011]

Three scenaria of interest: Scenario C

- The Source/Relay act strategically: they want to minimize their own average costs over the given time horizon

$$J^i(f) = \mathbf{E}^f\{\sum_t \rho^i(X_t^i, U_t^i)\}, \qquad i = 1,2$$

- Enlarge action space for Relay (to allow acceptance/rejection of Source packet)

$$U_t^2 \in \mathscr{U}^2 = \{0a, 0r, E_{23}\}$$

- One can study the resulting dynamic game and find Nash/sub-game perfect equilibria
- Unfortunately the equilibria of this game do not coincide with the optimal centralized solution of scenario A! (a.k.a., price of *anarchy*) Example: if optimal centralized action was $(E_{12}, 0a)$ this can never be a NE, because Relay is better off playing $0r$ (reject packet from source)

Three scenaria of interest: Scenario C

- The Source/Relay act strategically: they want to minimize their own average costs over the given time horizon

$$J^i(f) = \mathbf{E}^f\{\sum_t \rho^i(X_t^i, U_t^i)\}, \qquad i = 1,2$$

- Enlarge action space for Relay (to allow acceptance/rejection of Source packet)

$$U_t^2 \in \mathscr{U}^2 = \{0a, 0r, E_{23}\}$$

- One can study the resulting dynamic game and find Nash/sub-game perfect equilibria
- Unfortunately the equilibria of this game do not coincide with the optimal centralized solution of scenario A! (a.k.a., price of *anarchy*) Example: if optimal centralized action was $(E_{12}, 0a)$ this can never be a NE, because Relay is better off playing $0r$ (reject packet from source)

Three scenaria of interest: Scenario C

- The Source/Relay act strategically: they want to minimize their own average costs over the given time horizon

$$J^i(f) = \mathbf{E}^f\{\sum_t \rho^i(X_t^i, U_t^i)\}, \qquad i = 1,2$$

- Enlarge action space for Relay (to allow acceptance/rejection of Source packet)

$$U_t^2 \in \mathscr{U}^2 = \{0a, 0r, E_{23}\}$$

- One can study the resulting dynamic game and find Nash/sub-game perfect equilibria
- Unfortunately the equilibria of this game do not coincide with the optimal centralized solution of scenario A! (a.k.a., price of *anarchy*) Example: if optimal centralized action was $(E_{12}, 0a)$ this can never be a NE, because Relay is better off playing $0r$ (reject packet from source)

Three scenaria of interest: Scenario C

- The Source/Relay act strategically: they want to minimize their own average costs over the given time horizon

$$J^i(f) = \mathbf{E}^f\{\sum_t \rho^i(X_t^i, U_t^i)\}, \qquad i = 1,2$$

- Enlarge action space for Relay (to allow acceptance/rejection of Source packet)

$$U_t^2 \in \mathscr{U}^2 = \{0a, 0r, E_{23}\}$$

- One can study the resulting dynamic game and find Nash/sub-game perfect equilibria
- Unfortunately the equilibria of this game do not coincide with the optimal centralized solution of scenario A! (a.k.a., price of *anarchy*) Example: if optimal centralized action was $(E_{12}, 0a)$ this can never be a NE, because Relay is better off playing $0r$ (reject packet from source)

Incentivizing cooperation

- Our approach: Devise a protocol that provides incentives to Source/Relay so that the resulting dynamic game has equilibria that coincide with the solutions of the optimal centralized problem (Scenario A)
- Introduce a state/action-dependent monetary transfer $c(X_t, U_t)$ between agents

$$\hat{\rho}^1(X_t, U_t) = \rho^1(X_t^1, U_t^1) + c(X_t, U_t)$$
$$\hat{\rho}^2(X_t, U_t) = \rho^2(X_t^2, U_t^2) - c(X_t, U_t)$$

- Observe: the total societal cost is the same as in the centralized problem

$$\hat{\rho}^1(X_t, U_t) + \hat{\rho}^2(X_t, U_t) = \rho^1(X_t^1, U_t^1) + \rho^2(X_t^2, U_t^2)$$

Incentivizing cooperation

- Our approach: Devise a protocol that provides incentives to Source/Relay so that the resulting dynamic game has equilibria that coincide with the solutions of the optimal centralized problem (Scenario A)
- Introduce a state/action-dependent monetary transfer $c(X_t, U_t)$ between agents

$$\hat{\rho}^1(X_t, U_t) = \rho^1(X_t^1, U_t^1) + c(X_t, U_t)$$
$$\hat{\rho}^2(X_t, U_t) = \rho^2(X_t^2, U_t^2) - c(X_t, U_t)$$

- Observe: the total societal cost is the same as in the centralized problem

$$\hat{\rho}^1(X_t, U_t) + \hat{\rho}^2(X_t, U_t) = \rho^1(X_t^1, U_t^1) + \rho^2(X_t^2, U_t^2)$$

Incentivizing cooperation: basic result

- Important assumption: users know each others cost functions $\hat{\rho}^i$ (strategic behaviour does not manifest itself in desire for privacy/untrouthful revelation of cost structure)

Theorem ([Vasal and Anastasopoulos, 2013])

There exist monetary transfers $c(\cdot,\cdot)$ such that the unique Nash (sub-game perfect) equilibrium of the resulting dynamic game is exactly the optimal solution of the centralized control problem

- Implication: Source and Relay are incentivized to to behave in a way that coincides with the optimal centralized solution

Incentivizing cooperation: basic result

- Important assumption: users know each others cost functions $\hat{\rho}^i$ (strategic behaviour does not manifest itself in desire for privacy/untrouthful revelation of cost structure)

Theorem ([Vasal and Anastasopoulos, 2013])

There exist monetary transfers $c(\cdot,\cdot)$ such that the unique Nash (sub-game perfect) equilibrium of the resulting dynamic game is exactly the optimal solution of the centralized control problem

- Implication: Source and Relay are incentivized to to behave in a way that coincides with the optimal centralized solution

Incentivizing cooperation: basic result

- Important assumption: users know each others cost functions $\hat{\rho}^i$ (strategic behaviour does not manifest itself in desire for privacy/untrouthful revelation of cost structure)

Theorem ([Vasal and Anastasopoulos, 2013])

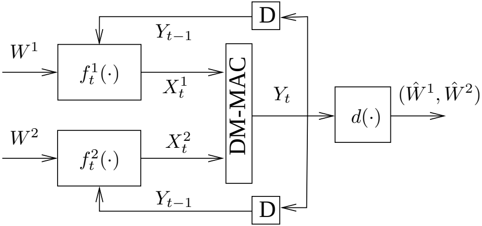
There exist monetary transfers $c(\cdot, \cdot)$ such that the unique Nash (sub-game perfect) equilibrium of the resulting dynamic game is exactly the optimal solution of the centralized control problem

- Implication: Source and Relay are incentivized to to behave in a way that coincides with the optimal centralized solution

Conclusions

- A number of communications problems can be viewed as centralized/decentralized control systems
- Using ideas from Control we can derive structural results and simplify the solution of these problems
- Can handle: dynamics; cooperation (team problems); and to some extent competition (games)
- Still a lot of open problems in this area
    - Capacity-achieving / Error exponent-achieving actual communication systems (single/multi-user)
    - Single-letter capacity for MAC with feedback...

Open problem: Capacity of Multiple Access Channel with noiseless feedback



- Messages $W^i \in \{1, 2, \ldots, 2^{nR^i}\}$, $i = 1, 2$
- Transmitted symbols $X_t^i \in \mathscr{X}^i$, $i = 1, 2$, $t = 1, 2, \ldots, n$
- Received symbols $Y_t \in \mathscr{Y}$, $t = 1, 2, \ldots, n$
- Discrete-memoryless MAC (DM-MAC) $Q(y_t | x_t^1, x_t^2)$
- Message estimates $(\hat{W}^1, \hat{W}^2) \in \mathscr{W}^1 \times \mathscr{W}^2$
- Encoding functions $X_t^i = f_t^i(W^i, Y^{t-1})$, $i = 1, 2$, $t = 1, 2, \ldots, n$
- Decoding function $(\hat{W}^1, \hat{W}^2) = d(Y^n)$

Thank you!

Bae, J. H. and Anastasopoulos, A. (2010).
The capacity of Markov channels with noiseless output and state feedback.
In *Information Theory and Applications*, San Diego, CA.

Bae, J. H. and Anastasopoulos, A. (2012).
A posterior matching scheme for finite-state channels with feedback.
IEEE Trans. Information Theory.
(Submitted).

Bross, S. I. and Lapidoth, A. (2005).
An improved achievable region for the discrete memoryless two-user multiple-access channel with noiseless feedback.
Information Theory, IEEE Transactions on, 51(3):811–833.

Chen, J. and Berger, T. (2005).
The capacity of finite-state Markov channels with feedback.
Information Theory, IEEE Transactions on, 51(3):780–798.

Cover, T. and Leung, C. (1981).

An achievable rate region for the multiple-access channel with feedback.
IEEE Trans. Information Theory, 27(3):292–298.

Gaarder, N. and Wolf, J. (1975).
The capacity region of a multiple-access discrete memoryless channel can increase with feedback (corresp.).
IEEE Trans. Inform. Theory, 21(1):100–102.

Horstein, M. (1963).
Sequential transmission using noiseless feedback.
IEEE Trans. Inform. Theory, 9(3):136–143.

Kavcic, A., Mandic, D., Huang, X., and Ma, X. (2009).
Upper bounds on the capacities of non-controllable finite-state channels using dynamic programming methods.
In *Proc. International Symposium on Information Theory*, pages 2346–2350.

Kramer, G. (2003).

Capacity results for the discrete memoryless network.
IEEE Trans. Information Theory, 49(1):4–21.

Mahajan, A. and Teneketzis, D. (2009).
Optimal design of sequential real-time communication systems,.
IEEE Trans. Inform. Theory.
(accepted for publication).

Nayyar, A., Mahajan, A., and Teneketzis, D. (2011).
Optimal control strategies in delayed sharing information structures.
IEEE Trans. Automatic Control, 56(7):1606–1620.

Nayyar, A. and Teneketzis, D. (2008).
On globally optimal real-time encoding and decoding strategies in multi-terminal communication systems.
In *Proc. IEEE Conf. on Decision and Control*, pages 1620–1627, Cancun, Mexico.

Ozarow, L. (1984).

The capacity of the white Gaussian multiple access channel with feedback.
IEEE Trans. Information Theory, 30(4):623–629.

Salehi, M. (1978).
Cardinality bounds on auxiliary variables in multiple-user theory via the method of ahlswede and korner.
Technical Report 33, Stanford Univ., Stanford, CA.

Schalkwijk, J. and Kailath, T. (1966).
A coding scheme for additive noise channels with feedback–I: No bandwidth constraint.
IEEE Trans. Inform. Theory, 12(2):172–182.

Shayevitz, O. and Feder, M. (2008).
The posterior matching feedback scheme: Capacity achieving and error analysis.
In *Proc. International Symposium on Information Theory*, pages 900–904.

Tatikonda, S. and Mitter, S. (2009).
The capacity of channels with feedback.
IEEE Trans. Information Theory, 55(1):323–349.

Vasal, D. and Anastasopoulos, A. (2012).
Achieving socially optimal solution through payments in a dynamic game for the relay channel.
In *Proc. Allerton Conf. Commun., Control, Comp.*

Vasal, D. and Anastasopoulos, A. (2013).
Incentive design in dynamic games for cooperative communications.
In *Information Theory and Applications*, San Diego, CA.

Venkataramanan, R. and Pradhan, S. S. (2009).
A new achievable rate region for the discrete memoryless multiple-access channel with feedback.
In *Proc. International Symposium on Information Theory*.

Walrand, J. and Varaiya, P. (1983).
Optimal causal coding - decoding problems.

IEEE Trans. Information Theory, 29(6):814–820.