An explicit solution of a two user dynamic team

Aditya Mahajan

Dept of ECE
McGill University

September 30, 2010
Allerton

Is dynamic team theory useful?

Is dynamic team theory useful?

Hlyuchj and Gallager, 1981

Although the notion of a dynamic team problem has been around for over 25 years, the class of problems is of sufficient complexity that little progress has been made toward a general solution technique or even in finding general properties of optimal solutions. Hence its value to the multi-access problem does not go much beyond a conceptual level.

Is dynamic team theory useful?

Hlyuchj and Gallager, 1981

Although the notion of a dynamic team problem has been around for over 25 years, the class of problems is of sufficient complexity that little progress has been made toward a general solution technique or even in finding general properties of optimal solutions. Hence its value to the multi-access problem does not go much beyond a conceptual level.

What is the state of the art after 30 years?

Have we made any progress toward a general solution technique to be of any value to the problem that Hlyuchj and Gallager were interested in?

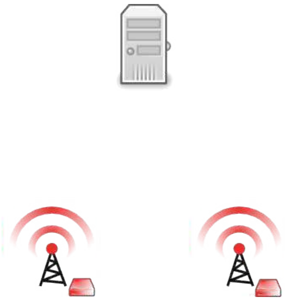
Problem Setup: Two-user multiple access broadcast

Two-users with single slot buffer

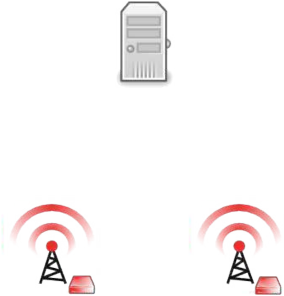
- $x_{i,t} \in \{0,1\}$: # packets in buffer
- $a_{i,t} \in \{0,1\}$: # new packet arrivals

$$a_{i,t} \sim \text{Ber}(p_i)$$

- $u_{i,t} \in \{0,1\}$: # transmitted packets



Problem Setup: Two-user multiple access broadcast



Two-users with single slot buffer

- $x_{i,t} \in \{0,1\}$: # packets in buffer
- $a_{i,t} \in \{0,1\}$: # new packet arrivals

$$a_{i,t} \sim \mathrm{Ber}(p_i)$$

- $u_{i,t} \in \{0,1\}$: # transmitted packets

Multiple access channel

Indicator of successful decoding: $z_t = u_{1,t} \oplus u_{2,t}$

$$x_{i,t+1} = (x_{i,t} - u_{i,t} z_t) \vee a_{i,t}$$

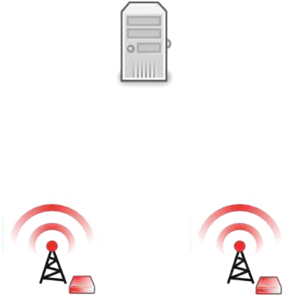
Problem Setup: Two-user multiple access broadcast

Two-users with single slot buffer

- $x_{i,t} \in \{0,1\}$: # packets in buffer
- $a_{i,t} \in \{0,1\}$: # new packet arrivals

$$a_{i,t} \sim \text{Ber}(p_i)$$

- $u_{i,t} \in \{0,1\}$: # transmitted packets



Multiple access channel

Indicator of successful decoding: $z_t = u_{1,t} \oplus u_{2,t}$

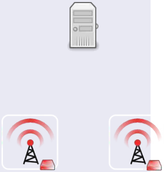
$$x_{i,t+1} = (x_{i,t} - u_{i,t} z_t) \vee a_{i,t}$$

Broadcast channel

z_t is available to the users after unit delay

Problem Setup: Two user multiple access broadcast

Problem (P1)



- *Given*: *arrival rates* p_1 *and* p_2
- *Choose*: *Transmission policies* $(\mathbf{g}_1, \mathbf{g}_2)$ *where* $\mathbf{g}_i = (g_{i,1}, g_{i,2}, \ldots, g_{i,T})$ *and*

$$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{i,1:t-1}, z_{1:t-1})$$

  - *Objective*: *Maximize*

$$\mathbb{E}^{\mathbf{g}_1,\mathbf{g}_2}\left\{\sum_{t=1}^{T} u_{1,t} \oplus u_{2,t}\right\} \quad or \quad \lim_{T\to\infty}\frac{1}{T}\mathbb{E}^{\mathbf{g}_1,\mathbf{g}_2}\left\{\sum_{t=1}^{T} u_{1,t} \oplus u_{2,t}\right\}$$

Literature Overview

Simplest canonical problem in multi-access networks.

- Slotted ALOHA and variants: Provide approximately optimal performance when the number of users is large. Huge literature . . .
- Collision incurs a cost but does not affect the dynamics Schoute, 76, Walrand Varaiya, 79,
- We are interested in the two-user problem in which collision affects the dynamics

Literature overview for Problem (P1)

Symmetric arrival rates

Asymmetric arrival rates

Literature overview for Problem (P1)

Symmetric arrival rates

- Hlyuchj Gallager 81:

- Ooi Wornell 96:

Asymmetric arrival rates

Literature overview for Problem (P1)

Symmetric arrival rates

- Hlyuchj Gallager 81:
  - Restrict to window protocols
  - Analytic soln.
  - lower bound
- Ooi Wornell 96:

Asymmetric arrival rates

Literature overview for Problem (P1)

Symmetric arrival rates

- Hlyuchj Gallager 81:
  - Restrict to window protocols
  - Analytic soln.
  - lower bound
- Ooi Wornell 96:
  - Genie reveals buffer state after a delay
  - Numerical soln
  - upper bound

Asymmetric arrival rates

Literature overview for Problem (P1)

Symmetric arrival rates

- Hlyuchj Gallager 81:
  - Analytic
  - lower bound
- Ooi Wornell 96:
  - Numerical
  - upper bound

lower and upper bounds match

Asymmetric arrival rates

Literature overview for Problem (P1)

Symmetric arrival rates

- Hlyuchj Gallager 81:
  - Analytic
  - lower bound
- Ooi Wornell 96:
  - Numerical
  - upper bound

lower and upper bounds match

Asymmetric arrival rates

- Lot of AI literature . . .
- Hansen et. al. 04
- Bernstein et. al. 05
- Szer Charpillet 06

Literature overview for Problem (P1)

Symmetric arrival rates

- Hlyuchj Gallager 81:
  - Analytic
  - lower bound
- Ooi Wornell 96:
  - Numerical
  - upper bound

lower and upper bounds match

Asymmetric arrival rates

- Lot of AI literature . . .
- Hansen et. al. 04
  - Numerical algorithm to find optimal soln
  - Out of memory for $T=5$
- Bernstein et. al. 05
  - Heuristic algorithm
  - Controller for size=8
- Szer Charpillet 06
  - Approx. algorithm
  - Out of memory for $T=5$

Literature overview for Problem (P1)

Symmetric arrival rates

- Hlyuchj Gallager 81:
  - Analytic
  - lower bound
- Ooi Wornell 96:
  - Numerical
  - upper bound

lower and upper bounds match

Asymmetric arrival rates

- Lot of AI literature ...

Approx algorithms ...
but can only solve the system until $T = 4$

Questions?

Symmetric arrival rates

Asymmetric arrival rates

Questions?

Symmetric arrival rates

- Optimal soln is known
- The proof is numerical
- Can we provide an analytic proof?

Asymmetric arrival rates

Questions?

Symmetric arrival rates

- Optimal soln is known
- The proof is numerical
- Can we provide an analytic proof?

Asymmetric arrival rates

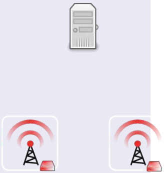
- Approx algorithms only work for small horizon
- Can we find algorithms that can solve large or infinite horizon problem?

Contributions of this paper

- Provide a dynamic programming decomposition
- The DP has countable state space and finite action space. Easy to use existing algorithms to find numerical solution for large or infinite horizon setups
- For symmetric arrival rates, find an analytic soln to the DP.

Problem Setup: Two user multiple access broadcast

Problem (P1)



- *Given*: *arrival rates* p_1 *and* p_2
- *Choose*: *Transmission policies* $(\mathbf{g}_1, \mathbf{g}_2)$ *where* $\mathbf{g}_i = (g_{i,1}, g_{i,2}, \ldots, g_{i,T})$ *and*

$$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{i,1:t-1}, z_{1:t-1})$$

  - *Objective*: *Maximize*

$$\mathbb{E}^{\mathbf{g}_1,\mathbf{g}_2}\left\{\sum_{t=1}^{T} u_{1,t} \oplus u_{2,t}\right\} \quad or \quad \lim_{T\to\infty} \frac{1}{T}\mathbb{E}^{\mathbf{g}_1,\mathbf{g}_2}\left\{\sum_{t=1}^{T} u_{1,t} \oplus u_{2,t}\right\}$$

Solution Outline

Transmission policy

$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{i,1:t-1}, z_{1:t-1})$

Solution Outline

Transmission policy

$$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{i,1:t-1}, z_{1:t-1})$$

Feedback ≡ control sharing

$$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{1,1:t-1}, u_{2,1:t-1})$$

Solution Outline

Transmission policy

$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{i,1:t-1}, z_{1:t-1})$

$x_{i,1:t-1}$ is redundant

$u_{i,t} = g_{i,t}(x_{i,t}, u_{1,1:t-1}, u_{2,1:t-1})$

Feedback $\equiv$ control sharing

$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{1,1:t-1}, u_{2,1:t-1})$

Solution Outline

Transmission policy

$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{i,1:t-1}, z_{1:t-1})$

$x_{i,1:t-1}$ is redundant

$u_{i,t} = g_{i,t}(x_{i,t}, u_{1,1:t-1}, u_{2,1:t-1})$

Feedback $\equiv$ control sharing

$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{1,1:t-1}, u_{2,1:t-1})$

Suff statistic for common info

$u_{i,t} = g_{i,t}(x_{i,t}, \pi_{1,t}, \pi_{2:t})$
where
$\pi_{i,t} = \Pr\left(x_{i,t} = 1 | u_{1,1:t-1}, u_{2,1:t-1}\right)$

Solution Outline (cont)

Dynamic Program

$$V_{T+1}(\pi_1, \pi_2) = 0$$

and for $t = T, T-1, \ldots, 1$

$$V_t(\pi_1, \pi_2) = \max\{W_{10,t}(\pi_1, \pi_2), W_{01,t}(\pi_1, \pi_2), W_{11,t}(\pi_1, \pi_2)\}$$

Solution Outline (cont)

Dynamic Program

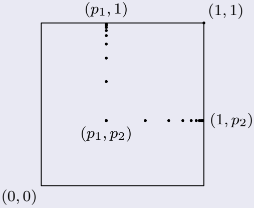
$$V_{T+1}(\pi_1, \pi_2) = 0$$

and for $t = T, T-1, \ldots, 1$

$$V_t(\pi_1, \pi_2) = \max\{W_{10,t}(\pi_1, \pi_2), W_{01,t}(\pi_1, \pi_2), W_{11,t}(\pi_1, \pi_2)\}$$

Reachability Analysis

The reachable set of (π_1, π_2) is countable.



Solution Outline

Transmission policy

$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{i,1:t-1}, z_{1:t-1})$

$x_{i,1:t-1}$ is redundant

$u_{i,t} = g_{i,t}(x_{i,t}, u_{1,1:t-1}, u_{2,1:t-1})$

Feedback $\equiv$ control sharing

$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{1,1:t-1}, u_{2,1:t-1})$

Suff statistic for common info

$u_{i,t} = g_{i,t}(x_{i,t}, \pi_{1,t}, \pi_{2:t})$
where
$\pi_{i,t} = \Pr(x_{i,t} = 1 | u_{1,1:t-1}, u_{2,1:t-1})$

Feedback $\equiv$ control sharing

- $z_t = u_{1,t} \oplus u_{2,t}$
- Thus,

$$u_{1,t} = z_t \oplus u_{2,t} \quad \textit{and} \quad u_{2,t} = z_t \oplus u_{1,t}$$

Feedback $\equiv$ control sharing

- $z_t = u_{1,t} \oplus u_{2,t}$
- Thus,

$$u_{1,t} = z_t \oplus u_{2,t} \quad \textit{and} \quad u_{2,t} = z_t \oplus u_{1,t}$$

- Hence,

$$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{1,1:t-1}, u_{2,1:t-1}, z_{1:t-1})$$

Feedback $\equiv$ control sharing

- $z_t = u_{1,t} \oplus u_{2,t}$
- Thus,

$$u_{1,t} = z_t \oplus u_{2,t} \quad \textit{and} \quad u_{2,t} = z_t \oplus u_{1,t}$$

- Hence,

$$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{1,1:t-1}, u_{2,1:t-1}, z_{1:t-1})$$

- Since $z_t = u_{1,t} \oplus u_{2,t}$,

$$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{1,1:t-1}, u_{2,1:t-1})$$

Solution Outline

Transmission policy

$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{i,1:t-1}, z_{1:t-1})$

$x_{i,1:t-1}$ is redundant

$u_{i,t} = g_{i,t}(x_{i,t}, u_{1,1:t-1}, u_{2,1:t-1})$

Feedback $\equiv$ control sharing

$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{1,1:t-1}, u_{2,1:t-1})$

Suff statistic for common info

$u_{i,t} = g_{i,t}(x_{i,t}, \pi_{1,t}, \pi_{2:t})$
where
$\pi_{i,t} = \Pr(x_{i,t} = 1 | u_{1,1:t-1}, u_{2,1:t-1})$

$x_{i,1:t-1}$ is redundant

- Arbitrarily fix the transmission policy of user 2
- $(x_{1,t}, u_{1,1:t-1}, u_{2,1:t-1})$ is a controlled Markov chain with control action $u_{1,t}$

$x_{i,1:t-1}$ is redundant

- Arbitrarily fix the transmission policy of user 2
- $(x_{1,t}, u_{1,1:t-1}, u_{2,1:t-1})$ is a controlled Markov chain with control action $u_{1,t}$
- Conditioned on the controls, the dynamics are independent

$$x_{1,1:t} \leftrightarrow (u_{1,1:t-1}, u_{2,1:t-1}) \leftrightarrow x_{2,1:t}$$

$x_{i,1:t-1}$ is redundant

- Arbitrarily fix the transmission policy of user 2
- $(x_{1,t}, u_{1,1:t-1}, u_{2,1:t-1})$ is a controlled Markov chain with control action $u_{1,t}$
- Conditioned on the controls, the dynamics are independent

$$x_{1,1:t} \leftrightarrow (u_{1,1:t-1}, u_{2,1:t-1}) \leftrightarrow x_{2,1:t}$$

- Thus, conditional expected reward

$$\mathbb{E}[u_{1,t} \oplus u_{2,t} | x_{1,1:t}, u_{1,1:t-1}, u_{2,1:t-}] \\ = \mathbb{E}[u_{1,t} \oplus u_{2,t} | x_{1,t}, u_{1,1:t-1}, u_{2,1:t-}]$$

$x_{i,1:t-1}$ is redundant

- Arbitrarily fix the transmission policy of user 2
- $(x_{1,t}, u_{1,1:t-1}, u_{2,1:t-1})$ is a controlled Markov chain with control action $u_{1,t}$
- Conditioned on the controls, the dynamics are independent

$$x_{1,1:t} \leftrightarrow (u_{1,1:t-1}, u_{2,1:t-1}) \leftrightarrow x_{2,1:t}$$

- Thus, conditional expected reward

$$\mathbb{E}[u_{1,t} \oplus u_{2,t} | x_{1,1:t}, u_{1,1:t-1}, u_{2,1:t-}] = \mathbb{E}[u_{1,t} \oplus u_{2,t} | x_{1,t}, u_{1,1:t-1}, u_{2,1:t-}]$$

- Thus,

$$u_{i,t} = g_{i,t}(x_{i,t}, u_{1,1:t-1}, u_{2,1:t-1})$$

Solution Outline

Transmission policy

$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{i,1:t-1}, z_{1:t-1})$

$x_{i,1:t-1}$ is redundant

$u_{i,t} = g_{i,t}(x_{i,t}, u_{1,1:t-1}, u_{2,1:t-1})$

Feedback $\equiv$ control sharing

$u_{i,t} = g_{i,t}(x_{i,1:t}, u_{1,1:t-1}, u_{2,1:t-1})$

Suff statistic for common info

$u_{i,t} = g_{i,t}(x_{i,t}, \pi_{1,t}, \pi_{2:t})$
where
$\pi_{i,t} = \Pr\left(x_{i,t} = 1 | u_{1,1:t-1}, u_{2,1:t-1}\right)$

Sufficient statistic for common information

$$u_{i,t} = g_{i,t}(x_{i,t}, u_{1,1:t-1}, u_{2,1:t-1})$$

- Common information: $(u_{1,1:t-1}, u_{2,1:t-1})$
- Private information: $x_{i,t}$

A special case of Mahajan, Nayyar, Teneketzis, 2008

Same solution approach (using the notion of a coordinator) applies

Sufficient statistic for common information (cont)

Coordinator of the two users

- Observation of coordinator: common information

$$(u_{1,1:t-1}, u_{2,1:t-1})$$

- Action of the coordinator: partial functions $(\gamma_{1,t}, \gamma_{2,t})$ s.t.

$$u_{i,t} = \gamma_{i,t}(x_{i,t})$$

Sufficient statistic for common information (cont)

Coordinator of the two users

- Observation of coordinator: common information

$$(u_{1,1:t-1}, u_{2,1:t-1})$$

- Action of the coordinator: partial functions $(\gamma_{1,t}, \gamma_{2,t})$ s.t.

$$u_{i,t} = \gamma_{i,t}(x_{i,t})$$

- For ease of notation, let $\varphi_{i,t} = \gamma_{i,t}(1)$. Then

$$u_{i,t} = \varphi_{i,t} x_{i,t}$$

Sufficient statistic for common information (cont)

Coordinator of the two users

- Observation of coordinator: common information

$$(u_{1,1:t-1}, u_{2,1:t-1})$$

- Action of the coordinator: partial functions $(\gamma_{1,t}, \gamma_{2,t})$ s.t.

$$u_{i,t} = \gamma_{i,t}(x_{i,t})$$

- For ease of notation, let $\varphi_{i,t} = \gamma_{i,t}(1)$. Then

$$u_{i,t} = \varphi_{i,t} x_{i,t}$$

- Think of $(\varphi_{1,t}, \varphi_{2,t})$ as the control action of the coordinator.

Optimization problem at the coordinator

Problem (P2)

- *Given: arrival rates p_1 and p_2*
- *Choose: Coordination policy $\mathbf{h} = (h_1, h_2, \dots, h_T)$ where*

$$(\varphi_{1,t}, \varphi_{2,t}) = h_t(u_{1,1:t-1}, u_{2,1:t-1}, \varphi_{1,1:t-1}, \varphi_{2,1:t-1})$$

- *Objective: Maximize*

$$\mathbb{E}^{\mathbf{h}}\left\{ \sum_{t=1}^{T} u_{1,t} \oplus u_{2,t} \right\} \quad \text{or} \quad \lim_{T\to\infty} \frac{1}{T} \mathbb{E}^{\mathbf{h}}\left\{ \sum_{t=1}^{T} u_{1,t} \oplus u_{2,t} \right\}$$

Sufficient statistic for common information (cont)

Proposition

Problem (P1) and (P2) are equivalent.

Proof.

- Any transmission policy $(\mathbf{g}_1, \mathbf{g}_2)$ for (P1) can be implemented in (P2) by choosing

$$\varphi_{i,t} = g_{i,t}(1, u_{1,1:t-1}, u_{2,1:t-1})$$

resulting in identical realization of all system variables.

- Any coordination policy $\mathbf{h}$ for (P2) can be implemented in (P1) by choosing

$$g_{i,t}(x_{i,t}, u_{1,1:t-1}, u_{2,1:t-1}) = \varphi_{i,t} x_{i,t}$$

where $\varphi_{i,t}$ is recursively chosen according to

$$(\varphi_{1,t}, \varphi_{2,t}) = h_t(u_{1,1:t-1}, u_{2,1:t-1}, \varphi_{1,1:t-1}, \varphi_{2,1:t-1})$$

□

Sufficient statistic for common information (cont)

Definition

$$\pi_{i,t} = \Pr\left(x_{i,t} = 1 \,\middle|\, \begin{array}{c} u_{1,1:t-1}, u_{2,1:t-1} \\ \varphi_{1,1:t-1}, \varphi_{2,1:t-1} \end{array}\right)$$

Proposition

In (P2), restricting attention to coordination policies of the form

$$(\varphi_{1,t}, \varphi_{2,t}) = h_t(\pi_{1,t}, \pi_{2,t})$$

is without loss. Therefore, in (P1) restricting attention to transmission policies of the form

$$u_{i,t} = g_{i,t}(x_{i,t}, \pi_{1,t}, \pi_{2,t}$$

is without loss.

Sufficient statistic for common information (cont)

Proof.

- $(\pi_{1,t}, \pi_{2,t})$ is a controlled Markov process with control action $(\varphi_{1,t}, \varphi_{2,t})$.
- Expected conditional reward

$$
\begin{aligned}
&\mathbb{E}[u_{1,t} \oplus u_{2,t} | u_{1,1:t-1}, u_{2,1:t-}, \varphi_{1,1:t}, \varphi_{2,1:t}] \\
&\quad = \pi_{1,t}\varphi_{1,t}(1 - \pi_{2,t}\varphi_{2,t}) + (1 - \pi_{1,t}\varphi_{1,t})\pi_{2,t}\varphi_{2,t} \\
&\quad = \mathbb{E}[u_{1,t} \oplus u_{2,t} | \pi_{1,t}, \pi_{2,t}, \varphi_{1,t}, \varphi_{2,t}]
\end{aligned}
$$

□

Solution Outline (cont)

Dynamic Program

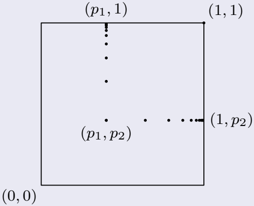
$$V_{T+1}(\pi_1, \pi_2) = 0$$

and for $t = T, T-1, \ldots, 1$

$$V_t(\pi_1, \pi_2) = \max\{W_{10,t}(\pi_1, \pi_2), W_{01,t}(\pi_1, \pi_2), W_{11,t}(\pi_1, \pi_2)\}$$

Reachability Analysis

The reachable set of (π_1, π_2) is countable.



Dynamic Program

- DP follows immediately from the fact that $(\pi_{1,t}, \pi_{2,t})$ is a controlled Markov process.
- By the same argument, the DP naturally extends to infinite horizon setup.

Reachability Analysis

- Let A_i be an operator from $[0,1]$ to $[0,1]$ such that for any $\pi \in [0,1]$
$$A_i\pi = 1 - (1 - p_i)(1 - \pi)$$

Evolution of info state

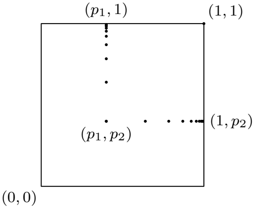
1. When $(\varphi_{1,t}, \varphi_{2,t}) = (0,0)$, $(\pi_{1,t+1}, \pi_{2,t+1}) = (A_1\pi_{1,t}, A_2\pi_{2,t})$.
2. When $(\varphi_{1,t}, \varphi_{2,t}) = (1,0)$, $(\pi_{1,t+1}, \pi_{2,t+1}) = (p_1, A_2\pi_{2,t})$.
3. When $(\varphi_{1,t}, \varphi_{2,t}) = (0,1)$, $(\pi_{1,t+1}, \pi_{2,t+1}) = (A_1\pi_{1,t}, p_2)$.
4. When $(\varphi_{1,t}, \varphi_{2,t}) = (1,1)$,
$$(\pi_{1,t+1}, \pi_{2,t+1}) = \begin{cases} (1,1) & \text{if } x_{1,t} = x_{2,t} = 1 \\ (p_1, p_2) & \text{otherwise} \end{cases}$$

Reachability Analysis (cont)

Reachable Set

Suppose the system starts in state $(\pi_1, \pi_2) = (p_1, p_2)$. Then the reachable set of (π_1, π_2) is

$$S = \{(1,1), (p_1, 1), (1, p_2), (p_1, p_2)\} \bigcup \{(A_1^n p_1, p_2), (p_1, A_2^n p_2), : n \in \mathbb{N}\}$$



Reachability Analysis (cont)

- The reachable set of $(\pi_{1,t}, \pi_{2,t})$ is countable.
- Thus, the inifnite horizon DP has countable state space and finite action space
- Stanard techniques to numerically solve such DP (e.g. Sennot, 97 , Leizarowitz Schwartz, 07)
- Contrast this with earlier attempt to obtain a numerical solution for this problem.

Symmetric arrivals

- Optimal coordination policy is symmetric $h(\pi_1, \pi_2) = h(\pi_2, \pi_1)$

Some definitions

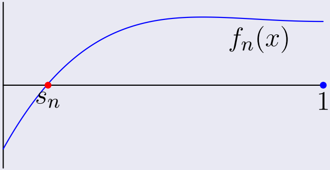
- Let $\tau \approx 0.38196$ be the root of $x = (1 - x)^2$ that lies in $[0, 1]$.

Symmetric arrivals

- Optimal coordination policy is symmetric $h(\pi_1, \pi_2) = h(\pi_2, \pi_1)$

Some definitions

- Let $\tau \approx 0.38196$ be the root of $x = (1-x)^2$ that lies in $[0,1]$.
- Let $f_n(x) = 1 + (1-x)^2 - (3+x)(1-x)^{n+1}$



  and s_n denote the root of $f_n(x)$ that is between $[0,1]$.
- $s_0 > \tau > s_1 > s_2 > \cdots 0$

Symmetric arrivals

Theorem

An optimal policy of the infinite horizon variant of (P2) is:

- round-robin policy for $p \geq \tau$

$$h^*(\pi_1, \pi_2) = \begin{cases} (1,0) & \text{if } \pi_1 > \pi_2, \\ (0,1) & \text{if } \pi_1 < \pi_2, \\ (1,0) \text{ or } (0,1) & \text{if } \pi_1 = \pi_2. \end{cases}$$

- transmit if you have a packet policy for $p < \tau$

$$h^*(\pi_1, \pi_2) = \begin{cases} (1,1) & \text{if } \pi_1 \leq A^m p, \pi_2 \leq A^m p, \\ (1,0) & \text{if } \pi_1 > \pi_2, \pi_1 > A^m p \\ (0,1) & \text{if } \pi_1 < \pi_2, \pi_2 > A^m p \\ (1,0) \text{ or } (0,1) & \text{if } \pi_1 = \pi_2 = 1. \end{cases}$$

where m is s.t. $s_{m+1} \leq p \leq s_m$.

Symmetric arrivals

Theorem

The average reward per unit time for the infinite horizon variant of (P2) is

$$J^* = \begin{cases} p[1-(2p^2-1)/D(p)] & \text{if } p \leq s_1, \\ (1-\bar{p}^2) & \text{if } s_1 \leq p; \end{cases}$$

where $\bar{p} = 1 - p$ and $D(p) = 1 + p^2 + p^3$.

Symmetric arrivals

Proof

Guess the form of the value function and verify!

1. When $p \geq s_1$,

$$\begin{aligned} v(p, A^n p) = v(A^n p, p) &= (1 - \bar{p}^{n+1}), \quad n > 1 \\ v(p, 1) = v(1, p) &= 1, \\ v(1,1) &= (1 + \bar{p}^2), \\ v(p,p) &= p \end{aligned}$$

Symmetric arrivals

Proof (cont)

Guess the form of the value function and verify!

2. When $s_{m+1} \leq p < s_m$, $m \in \mathbb{N}$

$$v(p,1) = v(1,p) = p[1 - f_0(p)/D(p)],$$
$$v(1,1) = 1,$$
$$v(p,p) = f_1(p)/D(p),$$
$$v(A^n p, p) = v(p, A^n p) = \begin{cases} c_*(n) & \text{if } n \leq m, \\ c^*(n) & \text{if } n > m \end{cases}$$

where

$$c_*(n) = \frac{\bar{p}}{p}(1 - \bar{p}^n)J^* + \bar{p}^{n+1} - \bar{p} + v(p,p),$$
$$c^*(n) = (1 - \bar{p}^{n+1}) + c_*(1) - v(1,p)$$

Symmetric arrivals

Proof

Guess the form of the value function and verify!

- Rest is just a matter of elementary (but tedious) algebra.
- The important point is that once we have a dynamic program, optimality of a particular policy can be checked systematically.

Proof

Guess the form of the value function and verify!

- Rest is just a matter of elementary (but tedious) algebra.
- The important point is that once we have a dynamic program, optimality of a particular policy can be checked systematically.
- We also need to guess the differential reward functions for the non-optimal actions. In general, this can be difficult. But, we exploit the symmetry and the fact that state space is countable.

Conclusion

Contributions

- An interesting example of two-user dynamic team that can be solved explicitly.
- For symmetric arrivals, identified the optimal policy analytically. The previous proof of optimality involved numerically solving a genie aided upper bound.
- For asymmetric arrivals, identified a DP with countable state space and finite action space. Earlier attempts for a numerical solution could only solve finite horizon problems with $T = 4$.

Conclusion

Future work

- We are missing a structural result:
  Each user gets a transmission opportunity $\varphi_{i,t} = 1$, at least once in two consecutive time slots
- The optimal policy satisfies this property.
- If we can prove this upfront, the DP will be much simpler (finite state and finite action spaces).

Thank You