Poster Number: 4335

Antenna Selection For Massive MIMO Systems Based On POMDP Framework

Sara Sharifi, Shahram ShahbazPanahi, and Min Dong

Dept. of Electrical, Computer, and Software Engineering, Ontario Tech University

Motivations and Challenges

- Motivation: Antenna selection technique can
  1. Reduce the hardware and computational complexity of the Massive MIMO systems.
  2. Benefits the diversity or beamforming gains.
- Challenge: Find an optimal policy to select a subset of antennas at each time slot such that
  1. The long term data rate is maximized.
  2. While only the partial CSI is available.

System Model

Let M and N be the number of BS antennas and radio frequency (RF) chains, respectively ($M > N$). We assume :

- Time division duplex (TDD) scheme.
- Single-antenna user.
- Channels evolve according to a Markov process.

Main goal: Select the best N antennas to **maximize the expected long-term data rate.**

POMDP Formulation

The partially observable Markov decision process (POMDP) frame-work represented as

- **State Space** denotes as $\mathcal{S}$, is a set of all possible state labeled as $\mathbf{s}_j$.
$$\mathbf{s}_j = \tilde{\mathbf{h}}_j \triangleq [\tilde{h}_{1j} \ \ \tilde{h}_{2j} \ \ \cdots \ \ \tilde{h}_{Mj}]^T,$$
where $\tilde{h}_{ij} \in \{\alpha_1, \alpha_2, \cdots, \alpha_Q\}$,
- **Action:** Selecting N out of M antennas. $\mathsf{a}_t \triangleq [\mathsf{a}_1 \ \ \mathsf{a}_2 \ \ \cdots \ \ \mathsf{a}_M]^T$. Here $\mathsf{a}_i \in \{0, 1\}$.
- **Transition Probability**, T is a matrix where the (i, j) element is $T_{ij} = \Pr(\mathbf{s}_t = \mathbf{s}_j | \mathbf{s}_{t-1} = \mathbf{s}_i)_{i=1}^{Q^M}$.
- **Observation:** $\mathbf{o}_t \triangleq \mathrm{diag}(\mathbf{a}_t)\mathbf{s}_t$.
- **Observation probability**:
$$\mathbf{O}(\mathbf{o}, \mathbf{a}) = \mathrm{diag}\Big(\Pr\Big\{\mathbf{o}_t = \mathbf{o} | \mathbf{s}_t = \mathbf{s}_i, \mathbf{a}_t = \mathbf{a}\Big\}_{i=1}^{Q^M}\Big)$$
- **Reward function:** The MISO data rate
$$R(\mathbf{s}, \mathbf{a}) = \log_2\left(1 + \frac{P\|\mathbf{o}\|^2}{\sigma^2}\right)$$
- **Belief** at time t is $\mathbf{b}_t \triangleq [b_{1,t} \ \ b_{2,t} \ \cdots \ b_{|\mathcal{S}|,t}]^T$ where $b_{j,t} = \Pr\{\mathbf{s}_t = \mathbf{s}_j | \mathcal{H}_{t-1}\}$. Here $\mathcal{H}_{t-1} \triangleq \{\mathbf{o}_{t-1}, \mathbf{a}_{t-1}, \mathcal{H}_{t-2}\}$.

Policy and Objective Function

- **Policy:** At time t policy maps the belief vector $\mathbf{b}_t$ to the action $\mathbf{a}_t$ such that $\mathbf{a}_t = \pi(\mathbf{b}_t)$.
- **Objective Function:**
$$J_\pi(\mathbf{b}_0) = E_{\{\mathbf{s}_t\}}\left\{\sum_{t=0}^{\infty} R(\mathbf{s}_t, \mathbf{a}_t) \Big| \mathbf{b}_0\right\}$$
The main goal: Find the optimal policy, π^* as $\pi^* = \arg\max_\pi J_\pi(\mathbf{b}_0)$, for any $\mathbf{b}_0$.

Two-state Channel Model

Gilbert-Elliot channel model ($Q = 2$): $\tilde{h}_{ij} \in \{\alpha, \beta\}$, for $j = 1, 2, \ldots 2^M$ and $i = 1, 2, \ldots, M$, where $|\alpha| > |\beta|$.

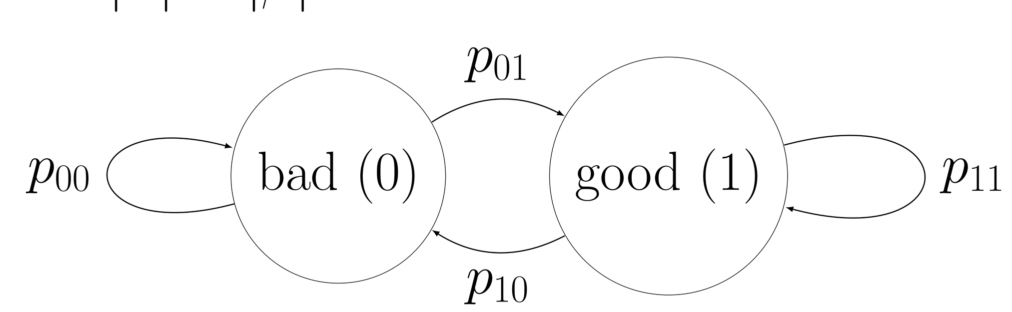

Figure 1: A two-state Markov chain model.

The channel state is **positively correlated,** i.e, $p_{11} > p_{01}$.

Optimality of the Myopic Policy for Two-state Channel Model

- Equivalent belief vector: $\boldsymbol{\omega}_t \triangleq [\omega_{1,t} \ \ \ldots \ \ \omega_{M,t}]^T$ where
$$\omega_{i,t} \triangleq \Pr(\mathsf{c}_{i,t} = 1 | \mathcal{H}_{t-1}).$$
  - $\mathsf{c}_{i,t} = 1$, means the channel between the i-th antenna and the user is in good state.
- $\boldsymbol{\omega}_{t+1}$ is updated as
$$\omega_{i,t+1} = \begin{cases} p_{11} & \text{if } a_{i,t} = 1, \ \ c_{i,t} = 1; \\ p_{01} & \text{if } a_{i,t} = 1, \ \ c_{i,t} = 0; \\ \omega_{i,t} p_{11} + (1 - \omega_{i,t}) p_{01} & \text{if } a_{i,t} = 0. \end{cases} \quad (1)$$
- Myopic policy is selecting those N antennas with the corresponding N largest entries in vector $\boldsymbol{\omega}_t$ at the current time slot t.

Optimality of Myopic Policy

For the two-state positively correlated channel model, in our POMDP-based antenna selection problem, **Myopic policy is optimal**.

Simulation Results

Time-average data rate: $\bar{R}_t \triangleq \frac{1}{t}\sum_{\tau=0}^{t} R(\mathbf{s}_\tau, \mathbf{a}_\tau)$. Results are averaged over 100 Monte Carlo runs.

Perfect Two-state Channel Model

- Value iteration algorithm provides the optimal solution.
- $p_{01} = 0.2$, $p_{11} = 0.8$, $\{\alpha_1, \alpha_2\} = \{\sqrt{0.1}, \sqrt{10}\}$.

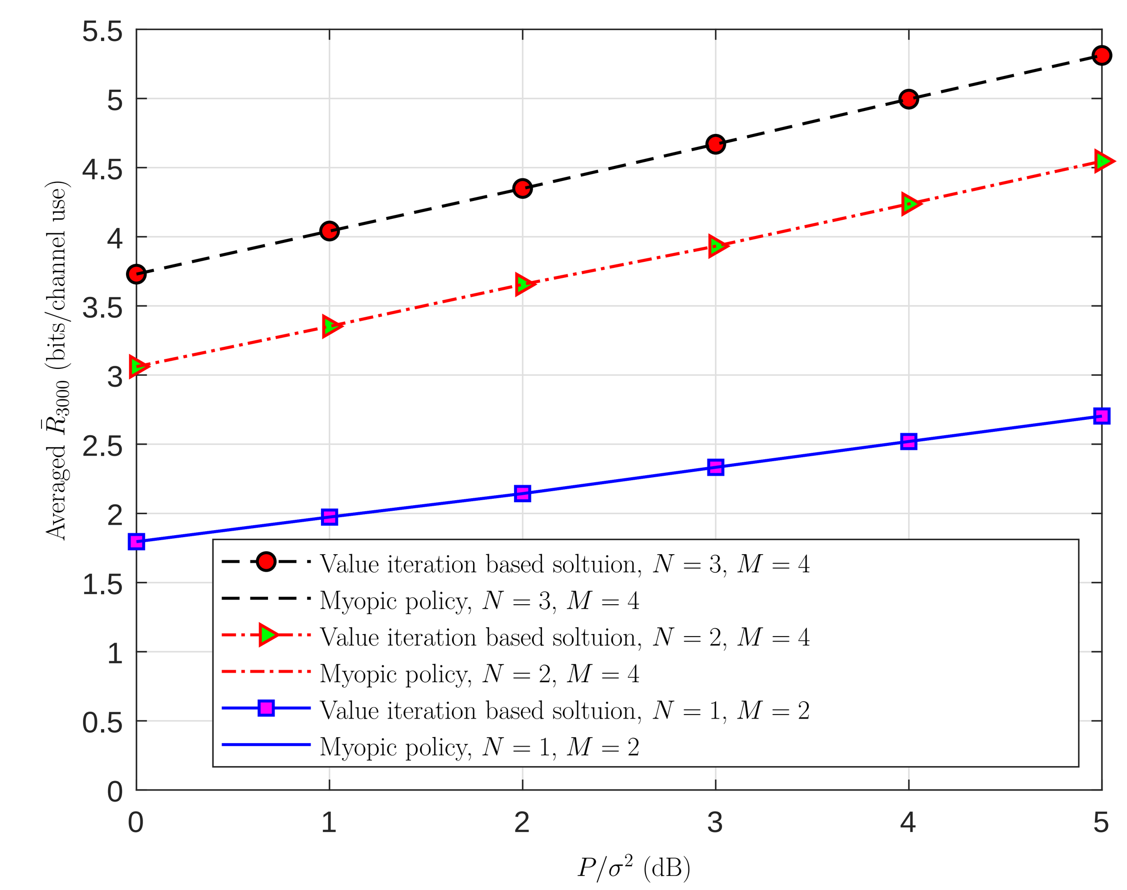

Figure 2: Time-average rate $\bar{R}_{3000}$ versus P/σ^2.

Gauss-Markov Channel Model

The first-order Gauss-Markov channel model $\mathsf{h}_{i,t} \triangleq \xi \mathsf{h}_{i,t-1} + \sqrt{1 - \xi^2}\mathsf{z}_{i,t}$, for $i = 1, \ldots, M$.

- $M = 200$, $P/\sigma^2 = 5$ dB, $\mathsf{h}_{i,t} \sim \mathcal{CN}(0, \mathbf{I})$, $\mathsf{z}_{i,t} \sim \mathcal{CN}(0, \mathbf{I})$, and $\sigma_h^2 = 1$.

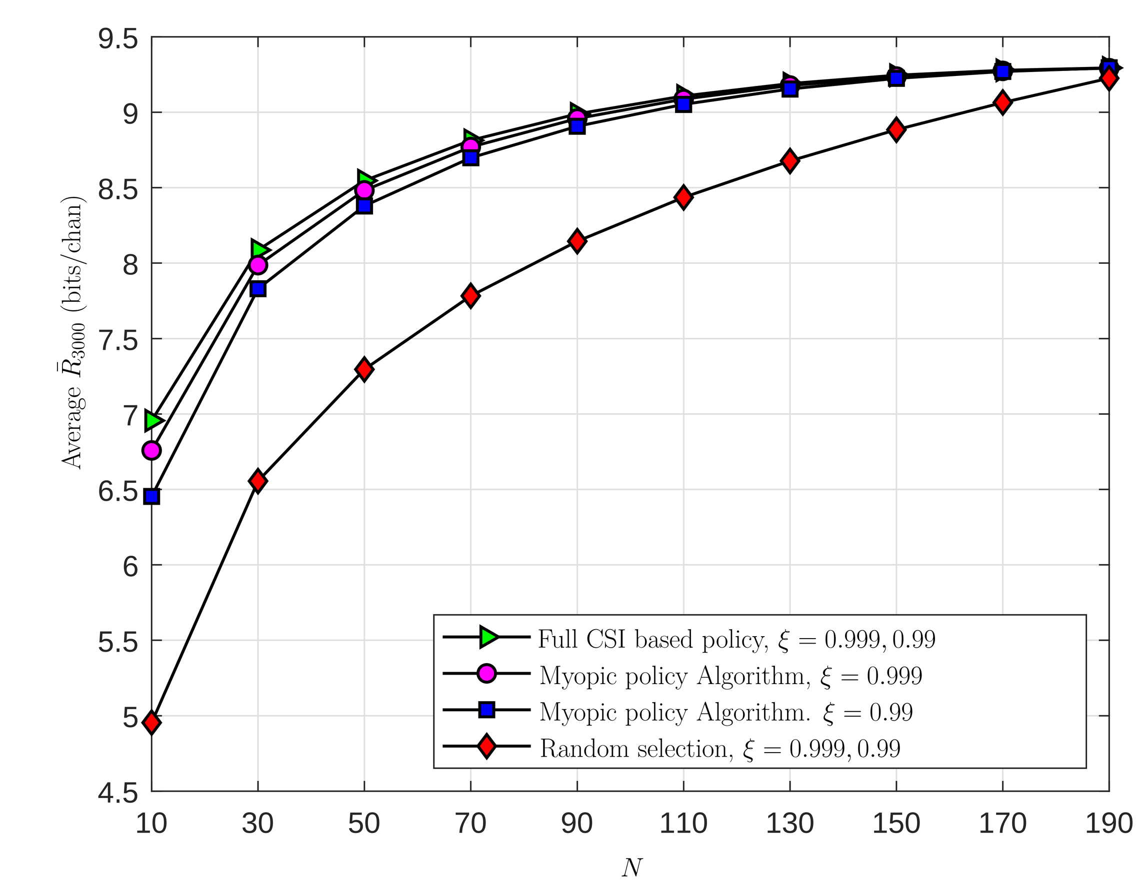

Figure 3: Average $\bar{R}_{3000}$ versus N.

Main Results

1. Figure 2: Myopic solution is optimal for the POMDP-based antenna selection problem.
2. Figure 3: Although only partial quantized CSI is available to the myopic policy, the performance of this policy is significantly close to full un-quantized CSI antenna selection policy.

Myopic Policy Algorithm

Algorithm 1 Myopic policy-based antenna selection

Inputs: Set v based on ξ and σ_h^2 .

At each time slot t:

Input: $\mathbf{o}_t$

1: Quantize the elements of $\mathbf{o}_t$ into α and β and update the elements of $\mathbf{c}_t$.

2: Update $\boldsymbol{\omega}_{t+1}$ (1).

3: For $i = 1, 2, \cdots, M$, choose the i-th entry of $\mathbf{a}_{t+1}$ as
$$a_{i,t+1} = \begin{cases} 1, & \text{if } \omega_{i,t+1} \text{ among the largest } N \text{ entries of } \boldsymbol{\omega}_{t+1}, \\ 0, & \text{otherwise.} \end{cases}$$

Output: $\mathbf{a}_{t+1}$

Our Contributions

- Formulating the massive antenna selection problem as a POMDP framework.
- Showing the optimality of Myopic policy for our antenna selection POMDP-based problem, for positively correlated two-state channel model.
- Proposing computationally affordable myopic policy-based algorithm for massive antenna selection problem.
- Applying the proposed myopic policy algorithm to the Rayleigh fading channel model to maximize the expected long-term downlink data rate.