Opportunistic Spectrum Access in Self Similar Primary Traffic

Keqin Liu, Xiangyang Xiao, Qing Zhao

Department of Electrical and Computer Engineering
University of California, Davis, CA 95616

Supported by NSF and ARL-CTA.
Multi-Channel Opportunistic Access

- Opportunistic Access: adapt to time-varying channel quality.
- Channel quality: “good (1)” or “bad (0)”
- Applications: Cognitive Radios, Downlink Scheduling in cellular network, Opportunistic transmission, Jamming/Anti-jamming
- Limited Sensing: can only sense and access $K$ out of $N$ channels in each slot.

Which channels to sense in each slot?
Existing Results on Markovian Channel Model

- Channel Quality in Cognitive Radio: “idle” (no primary traffic)/“busy”

- Channel quality evolves as an Markov chain

- Formulated as a restless multi-armed bandit process (Zhao&Etal’08)

- Established indexability and optimality of Whittle’s index policy (Liu&Zhao'08)

---

\[ p_{01} \]

\[ p_{00} \]

\[ p_{10} \]

\[ p_{11} \]
Myopic Policy for Markovian Channel Model

- For *homogeneous* channels, Whittle’s index policy is equivalent to the myopic policy.

- A semi-universal structure (Zhao&Krishnamachari’07)

- The optimality of the myopic policy

  - $N = 2$ (Zhao&Krishnamachari’07)
  - $N > 2$ when each channel has positive memory (Ahmad&etal’08)
  - $K = N - 1$ (Liu&Zhao’08)

---


Structure of the Myopic Policy: Positive Memory

- Stay with idle channels and leave busy ones to the end of the queue.

\[ K(t) \]

\[ K(t + 1) \]

---

© Keqin Liu, Qing Zhao. Presentation at MILCOM. Nov, 2008.


Structure of the Myopic Policy: Negative Memory

- Stay with busy channels and leave idle ones to the end of the queue.
- Reverse the order of unobserved channels.

\[ \mathcal{K}(t) \]

\[ \mathcal{K}(t + 1) \]
Robustness of the Myopic Policy

- No need to know the transition probabilities except the order of $p_{11}$ and $p_{01}$.
- Automatically tracks model variations.

Model Variation

$p_{11}=0.6$, $p_{01}=0.1$ (T≤5); $p_{11}=0.9$, $p_{01}=0.4$ (T>5)
Extend the Channel Model

How to extend to non-Markovian channel model with long memory?

► How to formulate the problem under a long-memory channel model?

► Does myopic policy still have a simple and robust structure?
  □ Robustness becomes more crucial under a more complex channel model

► Does myopic policy still have a strong performance?
Outline

- Self Similar Traffic
- Multi-time Scale Hierarchical Markovian Channel Model
- Restless Multi-armed Bandit Formulation and Optimal Policy
- Structure and Performance of Myopic Policy
- Conclusion
Self Similar Traffic

- Long range dependency of channel quality

- The channel quality $Q(t)$ evolves as a self similar process.

- Scale-invariant behavior

\[ (Q(at_1), \ldots, Q(at_k)) \overset{d}{=} (a^HQ(t_1), \ldots, a^HQ(t_k)) \]
Self Similar Primary Traffic

Autocorrelation decays polynomially $\Rightarrow$ long range dependency
Outline

- Self Similar Traffic
- Multi-time Scale Hierarchical Markovian Channel Model
- Restless Multi-armed Bandit Formulation and Optimal Policy
- Structure and Performance of Myopic Policy
- Conclusion
Multi-time Scale Hierarchical Markovian Channel Model

Multi-time scale hierarchical Markovian on-off process (Misra & Gong’98)

- $L$-level, each level is a two-state Markov process

\begin{align*}
&\text{Higher Level} \\
&\begin{array}{c}
1 \\
\longrightarrow \quad p_{10}^{(1)} \\
\longrightarrow \quad p_{01}^{(1)} \\
1
\end{array}
\quad \begin{array}{c}
0 \\
\longrightarrow \quad p_{01}^{(2)} \\
\longrightarrow \quad p_{10}^{(2)} \\
0
\end{array}
\quad \text{Lower Level}
\end{align*}

- The $k$-th level varies much slower than the $K+1$-th level

\[ p_{ii}^{(k)} \gg p_{ii}^{(m)} \quad \text{and} \quad p_{ij}^{(k)} \ll p_{ij}^{(m)} \quad \text{for} \quad i \neq j \]

- Higher-level Markov process has positive memory

\[ V. \text{Misra and W. Gong, “A hierarchical model for teletraffic,” in Proc. of the 37th Annual CDC, pp. 1674 – 1679, 1998.} \]
Multi-time Scale Hierarchical Markovian Channel Model

The channel is busy if all levels are in state 0

Events at different time scales: Establishing a session, Releasing a message

The quality of the channel evolves as a self similar process (Misra&Gong’98)

---

Outline

- Self Similar Traffic
- Multi-time Scale Hierarchical Markovian Channel Model
- Restless Multi-armed Bandit Formulation and Optimal Policy
- Structure and Performance of Myopic Policy
- Conclusion
The Sensing Policy

- **Sensing Policy** $\pi$
  - Choose the set $I(t)$ of $K$ channels to sense in each slot $t$.

- Consider $K = 1$ for the ease of presentation

- **Immediate Reward**
  - If sensed channel $i$ is idle, $B_i$ units of reward is accrued
  - If a sensed channel is busy, no reward; wait until the next slot
  - $R(t) = Q_{I(t)}(t)B_i$

- **Objective:** Maximize the expected total reward over a finite horizon $T$

$$\max_{\pi}\{\mathbb{E}[\sum_{t=1}^{T} R(t)]\}.$$
Optimal Sensing Policy for Opportunity Tracking

- Use entire observation history.
- Learn from the observation history.
- Foresighted planning: maximize total remaining reward.
- Optimal action: Gaining immediate reward vs. Gaining spectrum information.
A restless bandit Formulation

- The channel state is given by the state of the augmented Markov chain

- State of channel $n$ in slot $t$:

$$S_n(t) = (S_n^{(1)}(t), S_n^{(2)}(t), \cdots, S_n^{(L)}(t))$$

- $S_n^{(k)} \in \{0, 1\}$ represents the state of the $k$-th level Markov process for channel $n$

- Channel quality $Q_n(t) \in \{0, 1\}$:

$$Q_n(t) = 0 \text{ iff } S_n^{(k)} = 0 \text{ for all } 1 \leq k \leq L$$

- Given sensing action $I(t)$: immediate reward $R_{I(t)} = Q_{I(t)}(t)B_{I(t)}$
Sufficient Statistic for Optimal Action Making

- Sensing action $I(t)$ should be based on all observation history.
- Sufficient statistic: the a posteriori distribution $\lambda_n^{(k)}(t)$ (for all $1 \leq n \leq N$ and $1 \leq k \leq L$) that exploits the entire observation history.

$$
\lambda_n^{(k)}(t) = \Pr(S_n^{(k)}(t) = 1|\{a(i), O_{a(i)}(i)\}_{i=1}^{t-1})
$$

- Belief vector for channel $n$:

$$
\Lambda_n(t) = [\lambda_n^{(1)}(t), \lambda_n^{(2)}(t), \ldots, \lambda_n^{(L)}(t)].
$$

- System belief vector: $\Lambda(t) = [\Lambda_1(t), \Lambda_2(t), \ldots, \Lambda_N(t)].$

- Sensing Policy $\pi = \{\pi_t\}_{t=1}^T$: mappings from $\Lambda(t)$ to $a(t)$ for each $1 \leq t \leq T$.

Sequential stochastic control problem:

$$
\pi^* = \arg \max_{\pi} \mathbb{E}_\pi \left[ \sum_{t=1}^{T} R_{\pi_t}(\Lambda(t))(t) \bigg| \Lambda(1) \right].
$$
The Transition of the Belief Vector

- The belief vector transits according to Markov processes.

- $\Lambda(t+1)$ can be updated from $\Lambda(t)$, $I(t)$, and $O(I(t))(t)$ via Bayes rule.

\[
\lambda_{n}(k)(t+1) = \begin{cases} 
    p_{01}^{(n,k)} + \frac{\lambda_{n}(k)(t)(p_{11}^{(n,k)} - p_{01}^{(n,k)})}{1 - \prod_{i=1}^{L}(1 - \lambda_{n}^{(i)}(t))}, & a(t) = n, O_{n}(t) = 1 \\
    p_{01}^{(n,k)}, & a(t) = n, O_{n}(t) = 0 \\
    \lambda_{n}(k)(t)(p_{11}^{(n,k)} - p_{01}^{(n,k)}) + p_{01}^{(n,k)}, & a(t) \neq n
\end{cases}
\]
Complexity of Solving the Optimal Sensing Strategy

- Solving for the optimal policy: PSPACE-hard

- A heuristic policy: Myopic Policy—maximize expected immediate reward

The myopic action is given by

\[
\hat{a}(t) = \arg \max_{a=1,\ldots,N} \Pr[O_a(t) = 1|\Lambda(t)]B_a
\]

\[
= \arg \max_{a=1,\ldots,N} (1 - \prod_{k=1}^{L} (1 - \lambda_a^{(k)}(t)))B_a.
\]

For homogeneous channels, the myopic policy has a simple and robust structure that achieves the near-optimal performance.
Outline

► Self Similar Traffic

► Multi-time Scale Hierarchical Markovian Channel Model

► Restless Multi-armed Bandit Formulation and Optimal Policy

► Structure and Performance of Myopic Policy

► Conclusion
The Structure of the Myopic Policy for homogeneous Channels

- The myopic policy has a simple round-robin structure under the following two conditions.

Condition on Channel Model: The Markov processes at all levels have positive memory, i.e., \( p_{11}^{(k)} > p_{01}^{(k)} \) for all \( 1 \leq k \leq L \).

- Channel quality is more like to be the same rather than changing

- Without loss of generality, assume \( \lambda_{1}^{(1)}(1) \geq \lambda_{2}^{(1)}(1) \geq \cdots \geq \lambda_{N}^{(1)}(1) \).

Condition on Initial Belief Vector: The order of the initial belief values at all levels are the same, i.e., \( \lambda_{1}^{(k)}(1) \geq \lambda_{2}^{(k)}(1) \geq \cdots \geq \lambda_{N}^{(k)}(1) \) for all \( 1 \leq k \leq L \).

- Satisfied when each system belief starts from the stationary distribution.
The Structure of the Myopic Policy for homogeneous Channels

- Stay in the same channel when it is idle
- Switch to the channel visited the longest time ago when it is busy
Robustness of the Myopic Policy

- No need to know the transition probabilities $p_{11}^{(k)}$ and $p_{01}^{(k)}$.

- Automatically tracks model variations.
The Performance of the Myopic Policy

- Achieves the near-optimal performance
Relaxations on the Conditions for the Structure

- Consider two-level hierarchical Markovian channel models \((L = 2)\)

**Relaxation of Initial Belief Vector** The orders of belief values can be different

For any two channels \(i\) and \(j\) with \(\lambda_i^{(1)}(1) \geq \lambda_j^{(1)}(1)\), \(\lambda_i^{(2)}(1)\) can be smaller than \(\lambda_j^{(2)}(1)\) while the following two equations hold.

\[
\prod_{k=1}^{2} (1 - \lambda_i^{(k)}(1)) \leq \prod_{k=1}^{2} (1 - \lambda_j^{(k)}(1)),
\]

\[
\lambda_i^{(2)}(1) - \lambda_j^{(2)}(1) \geq \frac{(1 - P_{11}^{(2)})(P_{11}^{(1)} - P_{01}^{(1)})(\lambda_i^{(1)}(1) - \lambda_j^{(1)}(1))}{(1 - P_{11}^{(1)})(P_{11}^{(2)} - P_{01}^{(2)})}.
\]

- Initial belief values have the same channel ordering \(\implies\) The above inequalities are trivially satisfied

- Under the relaxation, the round robin structure still holds.
Conclusion

Main Results

► Hierarchical Markovian channel model to approximate self similar traffic

► Restless bandit formulation for solving the optimal sensing policy

► Simple and robust structure of the myopic policy for homogeneous channels

► The near-optimal performance of the myopic policy for homogeneous channels