Cognitive Medium Access: Exploration, Exploitation and Competition

by L. Lai, H. El Gamal, H. Jiang, and H.V. Poor
(submitted to IEEE Trans. Networking, October 2007)

proposed by Jan

CTG Reading Group

October 29, 2008

Cognitive Radio

- **Radio spectrum** has traditionally been organized according to fixed frequency plans defined through government licences

  $\Rightarrow$ Inefficient spectrum utilization
- **Cognitive Radio:** An intelligent radio that is aware of its environment and adapts accordingly
- Usually, focus on opportunistic **channel access** based on **spectrum sensing** capabilities

Three key issues

- **Exploration** of given channel resources to decide on availability
- **Exploitation** of available channels for own data transmission
- **Competition** with other cognitive users in multi-user scenario

Cognitive Radio

- **Radio spectrum** has traditionally been organized according to fixed frequency plans defined through government licences

  $\Rightarrow$ Inefficient spectrum utilization
- **Cognitive Radio:** An intelligent radio that is aware of its environment and adapts accordingly
- Usually, focus on opportunistic **channel access** based on **spectrum sensing** capabilities

Three key issues

- **Exploration** of given channel resources to decide on availability
- **Exploitation** of available channels for own data transmission
- **Competition** with other cognitive users in multi-user scenario

Focus of Paper

Side constraint: Availability probabilities of each channel **unknown**

- Lay **mathematical foundation** for cognitive medium access (MAC) using tools from
  - Reinforcement machine learning
  - Game theory
- Highlight **trade-off** between exploration and exploitation
- Design efficient **protocols** for cognitive MAC
- Derive theoretical **limits** & prove **optimality** of proposed schemes

Scenarios

- Single-user – single-channel setup
- Multi-user – single-channel setup
- Single-user – multiple-channel setup

Focus of Paper

Side constraint: Availability probabilities of each channel **unknown**

- Lay **mathematical foundation** for cognitive medium access (MAC) using tools from
  - Reinforcement machine learning
  - Game theory
- Highlight **trade-off** between exploration and exploitation
- Design efficient **protocols** for cognitive MAC
- Derive theoretical **limits** & prove **optimality** of proposed schemes

Scenarios

- Single-user – single-channel setup
- Multi-user – single-channel setup
- Single-user – multiple-channel setup

Problem Setup

- N primary channels given, each with bandwidth B
- Synchronous time-slotted communication (T time slots per block)
- θ_i: Probability of channel i being free; $\boldsymbol{\theta} := [\theta_1, ..., \theta_N]$
- θ block-wise constant and unknown *a priori* to cognitive user

  $\Rightarrow$ Balance between exploiting channel i using current knowledge of θ and exploring other channels to improve knowledge of θ

  $\Rightarrow$ Related to classic 'multi-armed bandit problem'
- Reward function: $W = \sum_{j=1}^{T} B\, Z_{S(j)}(j)$

  $S(j)$ channel chosen for sensing (and access), time slot j
  $Z_i(j) = 1$ if channel i free at time slot j; $Z_i(j) = 0$ otherwise

  $\Rightarrow$ Goal: Maximize expected throughput per block, $E\{W\}$

Problem Setup

- N primary channels given, each with bandwidth B
- Synchronous time-slotted communication (T time slots per block)
- θ_i: Probability of channel i being free; $\boldsymbol{\theta} := [\theta_1, ..., \theta_N]$
- $\boldsymbol{\theta}$ block-wise constant and unknown *a priori* to cognitive user

  $\Rightarrow$ Balance between exploiting channel i using current knowledge of $\boldsymbol{\theta}$ and exploring other channels to improve knowledge of $\boldsymbol{\theta}$

  $\Rightarrow$ Related to classic 'multi-armed bandit problem'

- Reward function: $W = \sum_{j=1}^{T} B\, Z_{S(j)}(j)$

  $S(j)$ channel chosen for sensing (and access), time slot j
  $Z_i(j) = 1$ if channel i free at time slot j; $Z_i(j) = 0$ otherwise

  $\Rightarrow$ Goal: Maximize expected throughput per block, $E\{W\}$

Problem Setup

- N primary channels given, each with bandwidth B
- Synchronous time-slotted communication (T time slots per block)
- θ_i: Probability of channel i being free; $\boldsymbol{\theta} := [\theta_1, ..., \theta_N]$
- $\boldsymbol{\theta}$ block-wise constant and unknown *a priori* to cognitive user

  $\Rightarrow$ Balance between exploiting channel i using current knowledge of $\boldsymbol{\theta}$ and exploring other channels to improve knowledge of $\boldsymbol{\theta}$

  $\Rightarrow$ Related to classic 'multi-armed bandit problem'
- Reward function: $$W = \sum_{j=1}^{T} B\, Z_{S(j)}(j)$$

  $S(j)$ channel chosen for sensing (and access), time slot j
  $Z_i(j) = 1$ if channel i free at time slot j; $Z_i(j) = 0$ otherwise

  $\Rightarrow$ Goal: Maximize expected throughput per block, $E\{W\}$

Optimal Bayesian Approach

- Assumption: PDF of θ known *a priori* $\rightarrow f(\theta)$
- Update PDF $f(\theta)$ with each new sensing result $z_i(j)$
- Optimal MAC strategy Γ^* depends on $f(\theta)$ and complete set $\Psi(j) = \{z_{s(1)}(1), \ldots, z_{s(j-1)}(j-1)\}$ of past sensing results
- Optimal trade-off between short-term gain (exploitation) and long-term gain (better knowledge of θ)
- Iterative solution stated in paper (plus some nice examples)
- Still, optimal strategy has prohibitive computational complexity
- Optimal Bayesian approach serves as ultimate theoretical limit

Optimal Bayesian Approach

- Assumption: PDF of θ known *a priori* $\rightarrow f(\theta)$
- Update PDF $f(\theta)$ with each new sensing result $z_i(j)$
- Optimal MAC strategy Γ^* depends on $f(\theta)$ and complete set $\Psi(j) = \{z_{s(1)}(1), ..., z_{s(j-1)}(j-1)\}$ of past sensing results
- Optimal trade-off between short-term gain (exploitation) and long-term gain (better knowledge of θ)
- Iterative solution stated in paper (plus some nice examples)
- Still, optimal strategy has prohibitive computational complexity
- Optimal Bayesian approach serves as ultimate theoretical limit

Suboptimum MAC Strategies

- Loss entailed by suboptimum MAC strategy Γ with respect to genie-aided strategy

$$L(\boldsymbol{\theta}, \Gamma) = \sum_{j=1}^{T} B\,\theta_{i*} - \sum_{j=1}^{T} B \sum_{i=1}^{N} \theta_i \,\mathrm{Pr}\{\Gamma(\Psi(j)) = i\}$$

where $\theta_{i*} = \max\{\boldsymbol{\theta}\}$

- Example: Consider MAC strategy which randomly selects channel i and sticks with it
  - $i = i^* \;\Rightarrow\; L(\theta, \Gamma) = 0$
  - $i \neq i^* \;\Rightarrow\; L(\theta, \Gamma)$ grows linearly with T (i.e., $L(\theta, \Gamma) \sim \mathcal{O}(T)$)
- Other examples show that there are several heuristic suboptimal MAC strategies that incur $L(\theta, \Gamma) \sim \mathcal{O}(T)$
- Goal: Design suboptimum MAC strategy which entails a smaller loss (if possible without requiring prior knowledge of $f(\theta)$)

Suboptimum MAC Strategies

- Loss entailed by suboptimum MAC strategy Γ with respect to genie-aided strategy

$$L(\boldsymbol{\theta}, \Gamma) = \sum_{j=1}^{T} B\,\theta_{i*} - \sum_{j=1}^{T} B \sum_{i=1}^{N} \theta_i \Pr\{\Gamma(\Psi(j)) = i\}$$

  where $\theta_{i*} = \max\{\boldsymbol{\theta}\}$

- Example: Consider MAC strategy which randomly selects channel i and sticks with it
  - $i = i^* \;\Rightarrow\; L(\boldsymbol{\theta}, \Gamma) = 0$
  - $i \neq i^* \;\Rightarrow\; L(\boldsymbol{\theta}, \Gamma)$ grows linearly with T (i.e., $L(\boldsymbol{\theta}, \Gamma) \sim \mathcal{O}(T)$)
- Other examples show that there are several heuristic suboptimal MAC strategies that incur $L(\boldsymbol{\theta}, \Gamma) \sim \mathcal{O}(T)$
- Goal: Design suboptimum MAC strategy which entails a smaller loss (if possible without requiring prior knowledge of $f(\boldsymbol{\theta})$)

Suboptimum MAC Strategies

- Loss entailed by suboptimum MAC strategy Γ with respect to genie-aided strategy

$$L(\boldsymbol{\theta}, \Gamma) = \sum_{j=1}^{T} B\,\theta_{i*} - \sum_{j=1}^{T} B \sum_{i=1}^{N} \theta_i \,\mathrm{Pr}\{\Gamma(\Psi(j)) = i\}$$

  where $\theta_{i*} = \max\{\boldsymbol{\theta}\}$

- Example: Consider MAC strategy which randomly selects channel i and sticks with it
  - $i = i^* \;\Rightarrow\; L(\boldsymbol{\theta}, \Gamma) = 0$
  - $i \neq i^* \;\Rightarrow\; L(\boldsymbol{\theta}, \Gamma)$ grows linearly with T (i.e., $L(\boldsymbol{\theta}, \Gamma) \sim \mathcal{O}(T)$)
- Other examples show that there are several heuristic suboptimal MAC strategies that incur $L(\boldsymbol{\theta}, \Gamma) \sim \mathcal{O}(T)$
- Goal: Design suboptimum MAC strategy which entails a smaller loss (if possible without requiring prior knowledge of $f(\boldsymbol{\theta})$)

Order Optimal MAC Strategy

- It is shown that $L(\theta, \Gamma)$ scales at least with $L(\theta, \Gamma) \sim \mathcal{O}(\ln T)$ if $f(\theta)$ is not known, as we need at least $\mathcal{O}(\ln T)$ time slots to sample each channel and get a reliable estimate of θ
- Proposed order optimal MAC strategy:
  - At the beginning of each block sense each channel once
  - At the beginning of time slot j calculate estimate

$$\hat{\theta}_i(j) = X_i(j)/Y_i(j)$$

$Y_i(j)$: Number of time slots in which channel i was already sensed
$X_i(j)$: Number of time slots in which channel i was found free

- Assign index to channel i

$$\Lambda_i(j) := \hat{\theta}_i(j) + \sqrt{2 \ln j / Y_i(j)}$$

- In time slot $(j+1)$ choose channel with largest index $\Lambda_i(j)$ to sense
- Correction term in $\Lambda_i(j)$ makes sure that channel i^* is sensed many times before it is declared best

Order Optimal MAC Strategy

- It is shown that $L(\boldsymbol{\theta}, \Gamma)$ scales at least with $L(\boldsymbol{\theta}, \Gamma) \sim \mathcal{O}(\ln T)$ if $f(\boldsymbol{\theta})$ is not known, as we need at least $\mathcal{O}(\ln T)$ time slots to sample each channel and get a reliable estimate of $\boldsymbol{\theta}$
- Proposed order optimal MAC strategy:
  - At the beginning of each block sense each channel once
  - At the beginning of time slot j calculate estimate

$$\hat{\theta}_i(j) = X_i(j)/Y_i(j)$$

  $Y_i(j)$: Number of time slots in which channel i was already sensed
  $X_i(j)$: Number of time slots in which channel i was found free

- Assign index to channel i

$$\Lambda_i(j) := \hat{\theta}_i(j) + \sqrt{2 \ln j / Y_i(j)}$$

- In time slot $(j+1)$ choose channel with largest index $\Lambda_i(j)$ to sense
- Correction term in $\Lambda_i(j)$ makes sure that channel i^* is sensed many times before it is declared best

Multiuser Case

- With respect to exploitation/exploration trade-off same goals for each individual cognitive user
- Additionally, minimize collision probability, i.e., different cognitive users should sense different channels
  - Optimal distributed MAC protocol proposed which is based on symmetric rule (θ known)
  - Game-theoretic approach investigated to operate at Nash equilibrium