Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad-Hoc Networks: A POMDP Framework

Qing Zhao, Lang Tong, Anathram Swami, and Yunxia Chen

EE360 Presentation: Kun Yi
Stanford University

Main Challenges for Ad-hoc OSA Network

- 1. Cognitive radio nodes cannot sense the full spectrum
- Energy inefficient
- Hardware complexity

Main Challenges for Ad-hoc OSA Network

- 1. Cognitive radio nodes cannot sense the full spectrum
- Energy inefficient
- Hardware complexity

- 2. Transmitter-receiver synchronization
- Avoid unnecessary control messages

Main Challenges for Ad-hoc OSA Network

- 1. Cognitive radio nodes cannot sense the full spectrum
- Energy inefficient
- Hardware complexity

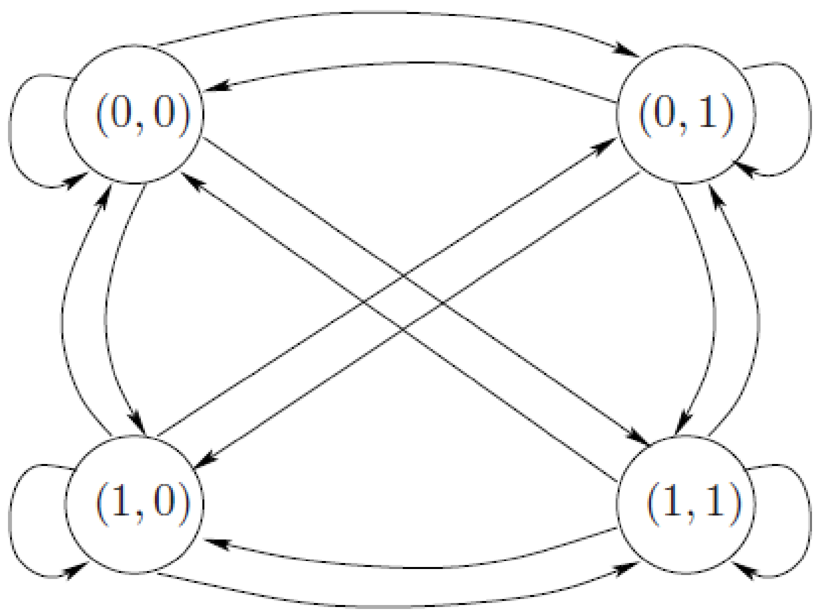
- 2. Transmitter-receiver synchronization
- Avoid unnecessary control messages

POMDP: Partially Observable Markov Process

Contents

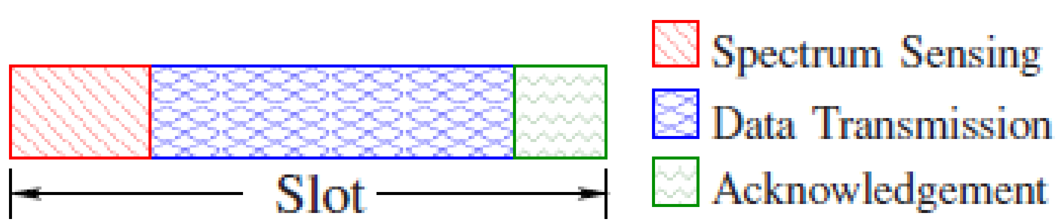
System Model: Network Model

- N channels , $M = 2^N$ states
- Each channel is either 0 (occupied) or 1 (idle)
- Channel state evolves as a discrete-time Markov process
- Assume knowledge of transition probabilities
- Usage statistics unchanged for T slots

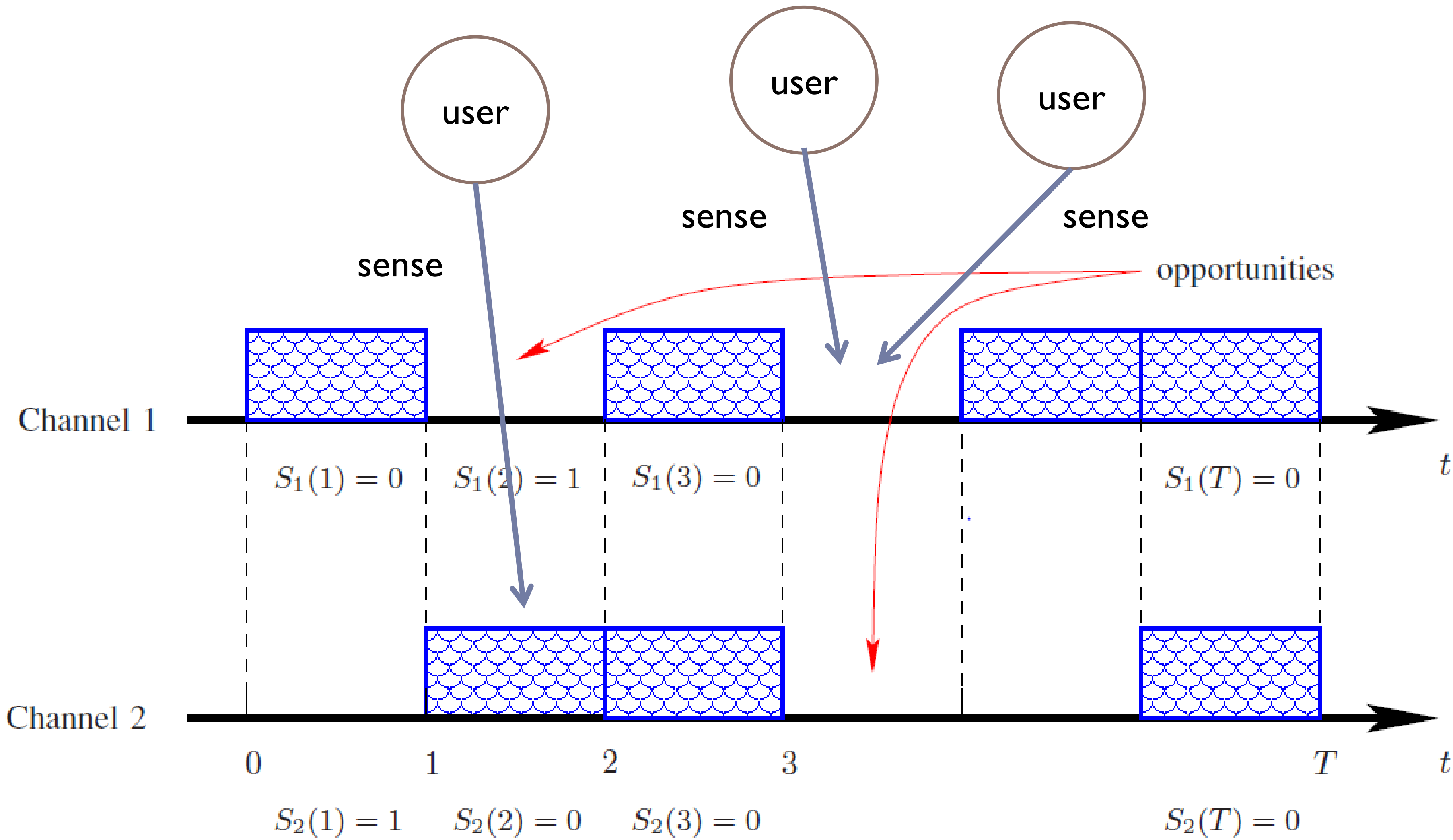


System Model: Cognitive Radio Network

- Secondary network seeking spectrum holes
- Join/exit network and transmit/receive independently
- Decentralized protocol
- Each user can only sense no more than L_1 channels, access no more than L_2 channels
- Slot structure:
  - 1. sense a set of channels, then access some of them
  - 2. randomly backoff, transmits if no other accesses channel
  - 3. acknowledges successful transmission



Illustrating the challenges

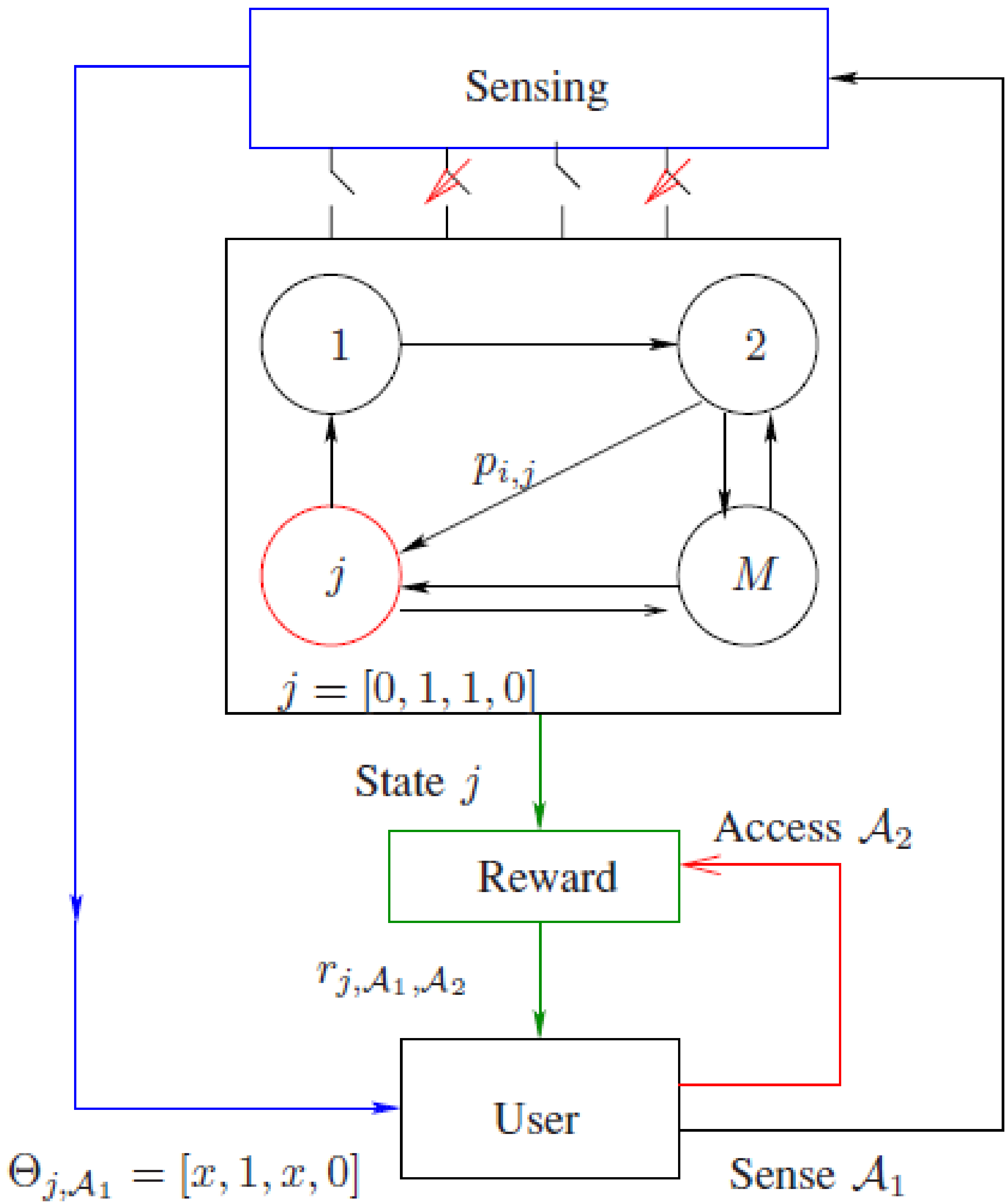


POMDP formulation

- S is set of states
- A is a set of actions
- Θ is a set of observations
- R reward function
- Λ belief vector, $\Lambda(t) = [\lambda_1(t), \ \ldots, \ \lambda_M(t)]$
  - $\lambda_i(t)$: conditional probability (given observation decision and history) that network state is i at beginning of slot t
  - sufficient statistic for the design of optimal action
- π policy: sequence of function $\Lambda(t) \rightarrow A(t)$
- Objective: maximize accumulated reward over a horizon

POMDP formulation:

- Beginning of slot, chooses a set A_1 channels to sense, $|A_1| \leq L_1$
- Choose subset of $A_2 \subset A_1$ channels to access (transmit/receive), $|A_2| \leq L_2$
- Given state is j, observes Θ_{j,A_1}
- Acknowledge successful/unsuccessful transmission
- Calculate reward r_{j,A_1,A_2}
  - $r_{j,A_1,A_2}(t) = \sum_{i \in A_2} S_i(t) B_i$, number of bits transmitted
- Objective: maximize sum(r) over T slots



Sensing
1
2
$p_{i,j}$
j
M
$j = [0, 1, 1, 0]$
State j
Access $\mathcal{A}_2$
Reward
$r_{j,\mathcal{A}_1,\mathcal{A}_2}$
User
$\Theta_{j,\mathcal{A}_1} = [x, 1, x, 0]$
Sense $\mathcal{A}_1$

Optimal Sensing/Access Strategy

- Assume error-free sensing, $L_1 = L_2 = 1$
- Choice of A_2: Transmit iff channel is sensed to be available
- $V_t\big(\Lambda(t)\big)$ denote the maximum expected remaining reward there can be from time t

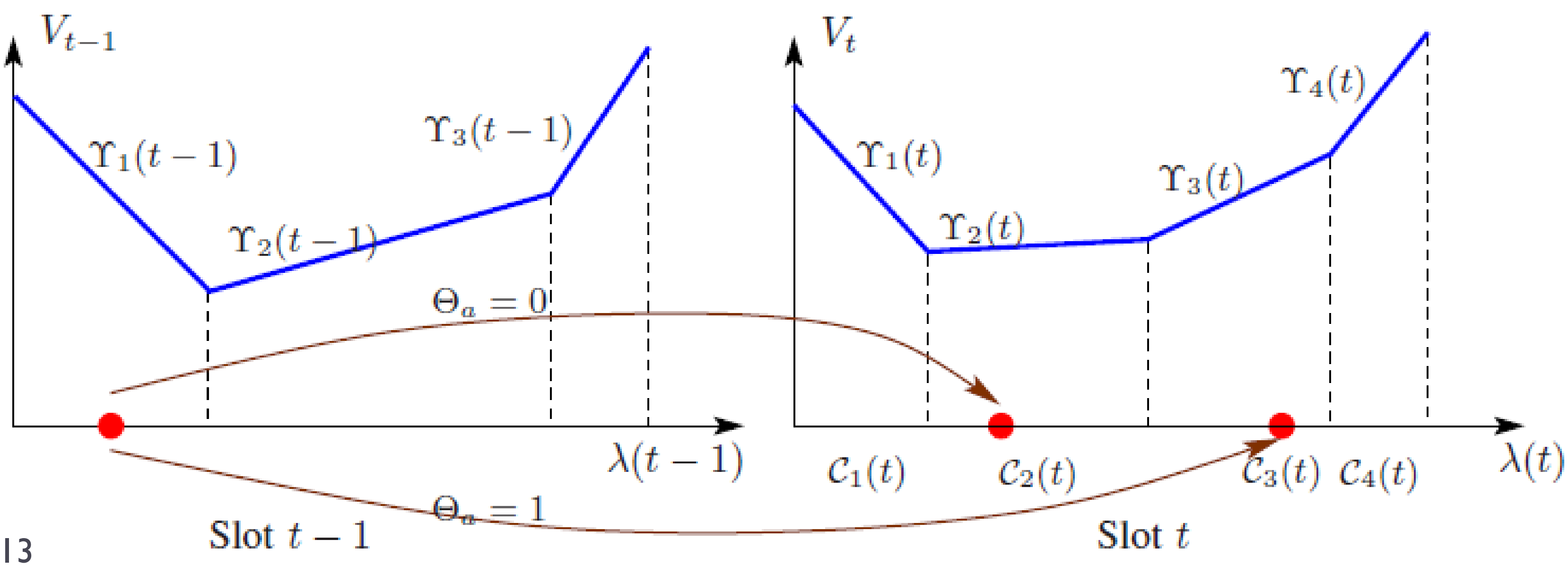
$$V_t(\Lambda(t)) = \max_{a=1,\cdots,N} \{ \sum_{i=1}^{M} \lambda_i \sum_{j=1}^{M} p_{i,j} \sum_{\theta=0}^{1} \Pr[\Theta_{j,a} = \theta] (\theta B_a + V_{t+1}(\mathcal{T}(\Lambda(t)|a,\theta))) \}$$

immediate reward

future reward

Optimal Sensing/Access Strategy

- $V_t\big(\Lambda(t)\big)$ denote the maximum expected remaining reward there can be from time t
- Convex and piecewise linear
- Calculate backwards:
  - $\Upsilon_i(t+1) \;\rightarrow \textit{optimal action and } \Upsilon_i(t)$



Optimal Sensing/Access Strategy

- Assume error-free sensing, $L_1 = L_2 = 1$
- Choice of A_2: Transmit iff channel is sensed to be available
- $V_t\big(\Lambda(t)\big)$ denote the maximum expected remaining reward there can be from time t

$$V_t(\Lambda(t)) = \max_{a=1,\cdots,N}\{\sum_{i=1}^{M}\lambda_i\sum_{j=1}^{M}p_{i,j}\sum_{\theta=0}^{1}\Pr[\Theta_{j,a}=\theta](\theta B_a + V_{t+1}(\mathcal{T}(\Lambda(t)|a,\theta)))\}$$

immediate reward

future reward

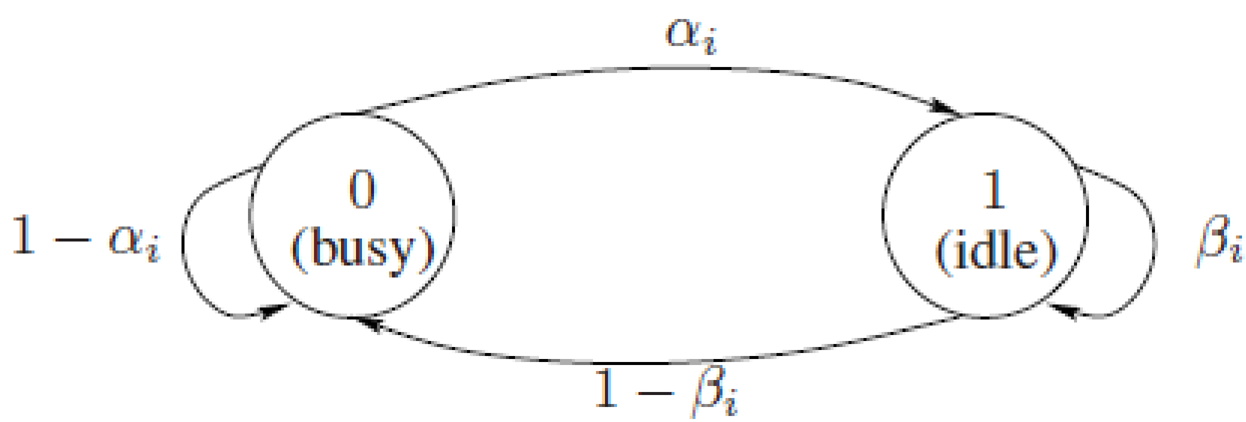
$$< \Lambda(t+1), \Upsilon_{i_{\Lambda(t+1)}}(t+1) >$$

Sub-optimal Strategy

- Reason: dimension of Λ grows exponentially
- Assume each channel independent, then can reduce the dimension of the sufficient statistics from 2^N to N
- New sufficient statistic: belief of each channel

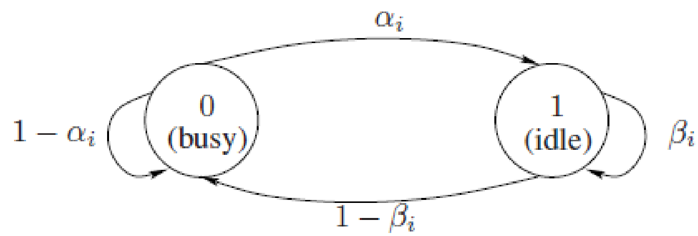
$$\Omega(t) = [\omega_1(t), \ldots, \omega_N(t)]$$

- Channel model



Sub-optimal Strategy

- Channel:



- Simple recurrence

$$\Omega(t+1) = [\omega_1(t+1), \cdots, \omega_N(t+1)] \triangleq \mathcal{T}(\Omega(t)|a_*(t), \Theta_{a_*}(t)),$$

$$\omega_i(t+1) = \begin{cases} 1 & \text{if } a_*(t) = i, \Theta_{a_*}(t) = 1 \\ 0 & \text{if } a_*(t) = i, \Theta_a(t) = 0 \\ \omega_i(t)\beta_i + (1-\omega_i(t))\alpha_i & \text{if } a_*(t) \neq i \end{cases}$$

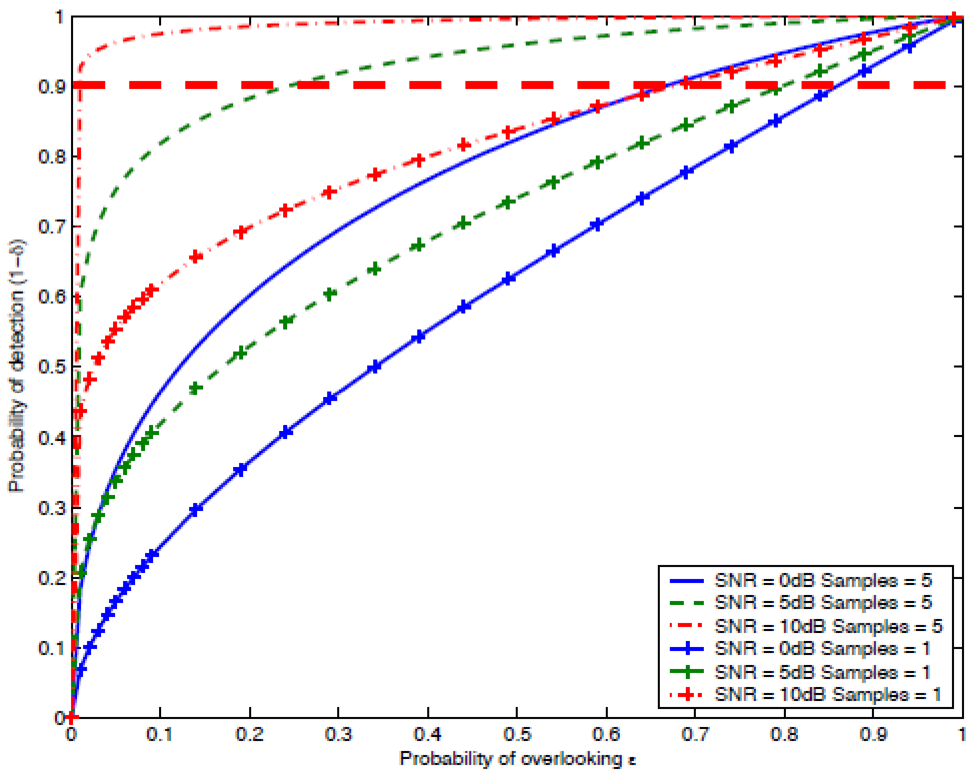
- Strategy: maximize immediate reward
- (probability of sensing a current idle channel)

$$a_*(t) = \arg \max_{a=1,\cdots,N} (\omega_a(t)\beta_a + (1-\omega_a(t))\alpha_a)B_a.$$

Problem: Sensing Error

- Does the optimal strategy change?
- Keep the collision probability below the constraint
  - $miss\ detection\ chance\ \delta \leq \zeta$

Receiver Operating Characteristics for different spectrum sensors



choose

$$\delta^* = \zeta$$

operating on the threshold!!

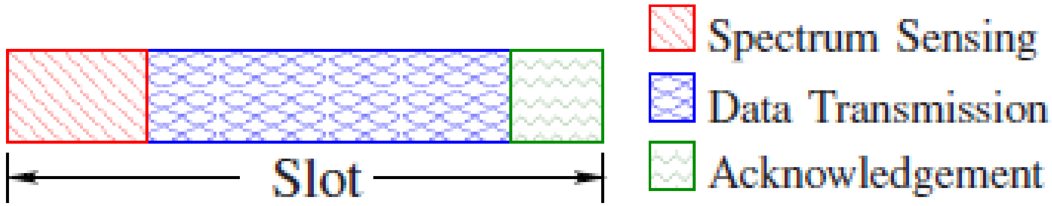
Decentralized Cognitive MAC: Accessing

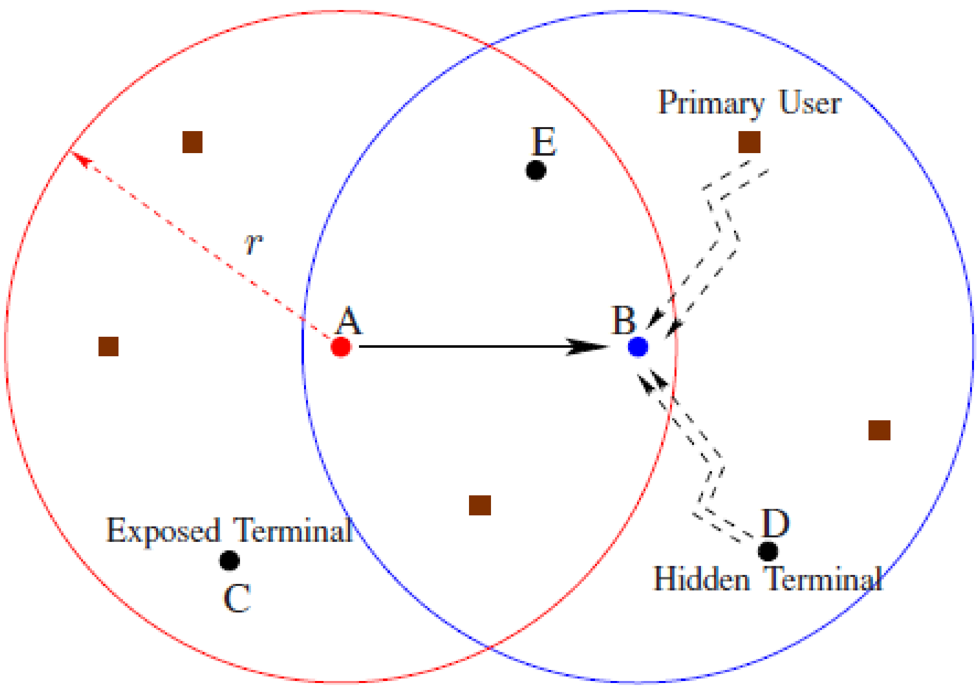
- Key: if transmitter and receiver carry out the same algorithm based on the same information, their belief vector should be the same!! (Always synchronized)
- If sensing error, receiver don't have observation Θ_{a^*}
- Update belief according to
  - a^*: decision on which channel to sense
  - K_{a^*}: whether an acknowledgement sent/received
- Assume error-free ACK (what if not?)

$$\omega_i(t+1) \triangleq \Pr[S_i(t)=1|\Omega(t), a_*, K_{a_*}]$$
$$= \begin{cases} \omega_i(t)\beta_i + (1-\omega_i(t))\alpha_i & \text{if } a_* \neq i \\ 1 & \text{if } a_* = i, K_{a_*} = 1 \\ \frac{\epsilon(\omega_{a_*}\beta_{a_*} + (1-\omega_{a_*})\alpha_{a_*})}{\epsilon(\omega_{a_*}\beta_{a_*} + (1-\omega_{a_*})\alpha_{a_*}) + (\omega_{a_*}(1-\beta_{a_*}) + (1-\omega_{a_*})(1-\alpha_{a_*}))} & \text{if } a_* = i, K_{a_*} = 0 \end{cases}$$

Problem: Hidden/Exposed Terminals

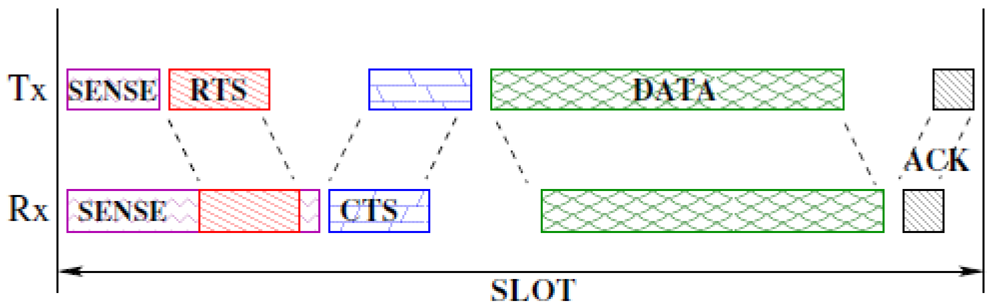
- Recall basic slot protocol





Revised MAC

- Revised MAC: sending RTS/CTS to alleviate hidden/exposed terminals



- Combined with sensing/accessing choice between slots: Complete Cognitive MAC protocol

Simulation Result

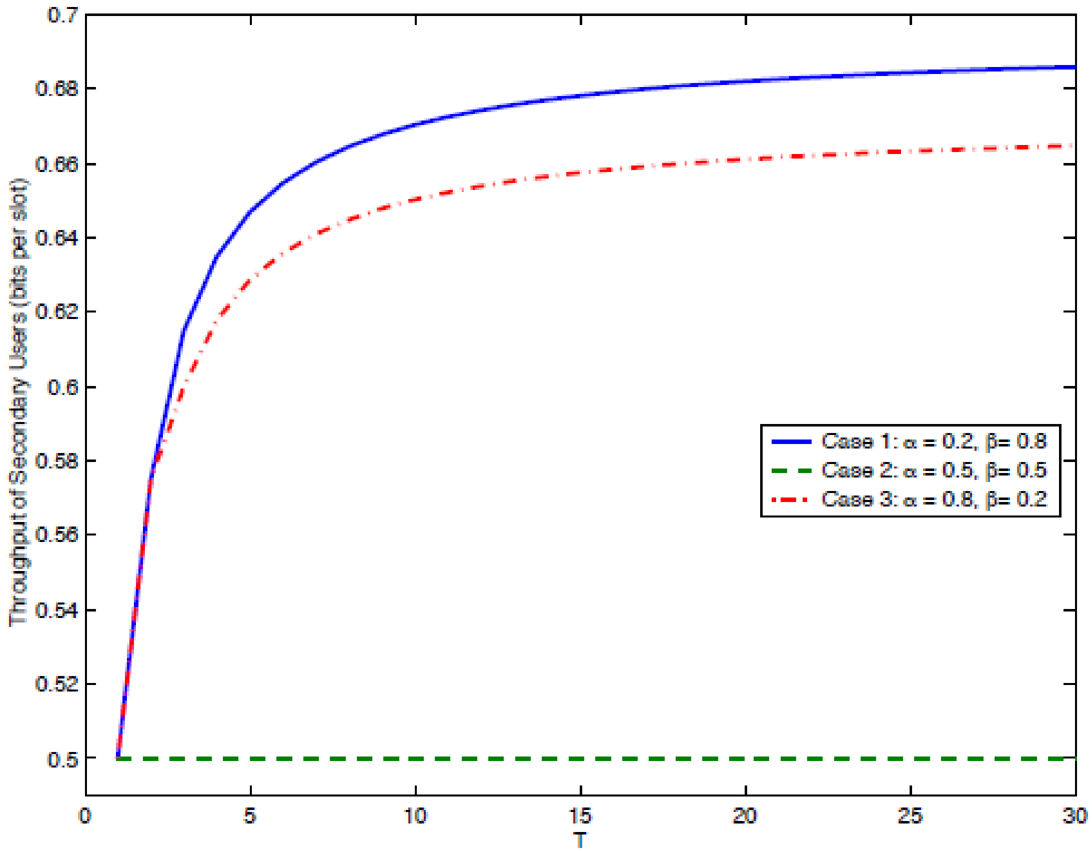


Fig. 12. Performance of the optimal cognitive MAC protocol under different spectrum occupancy statistics ($N = 3$ independent channels with the same bandwidth $B = 1$ and transition probabilities $\{\alpha, \beta\}$).

Comparing Greedy and Optimal

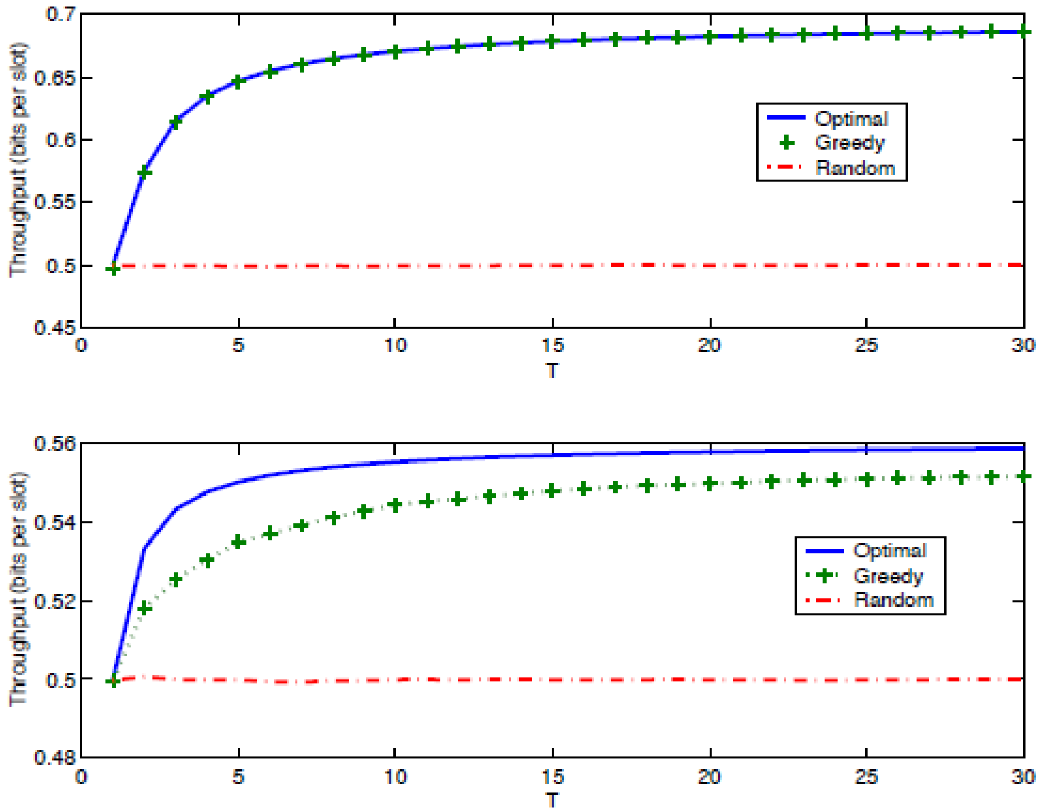


Fig. 13. Performance comparison of the greedy approach and the optimal strategy (in the upper plot, we have $N = 3$ independent channels with the same bandwidth $B = 1$ and transition probabilities $\{\alpha = 0.2, \beta = 0.8\}$; in the lower plot, $N = 3$, $\vec{\alpha} = [0.8, 0.6, 0.4], \vec{\beta} = [0.6, 0.4, 0.2], \vec{B} = [3/4, 1, 3/2]$).

Performance in Sensing Error

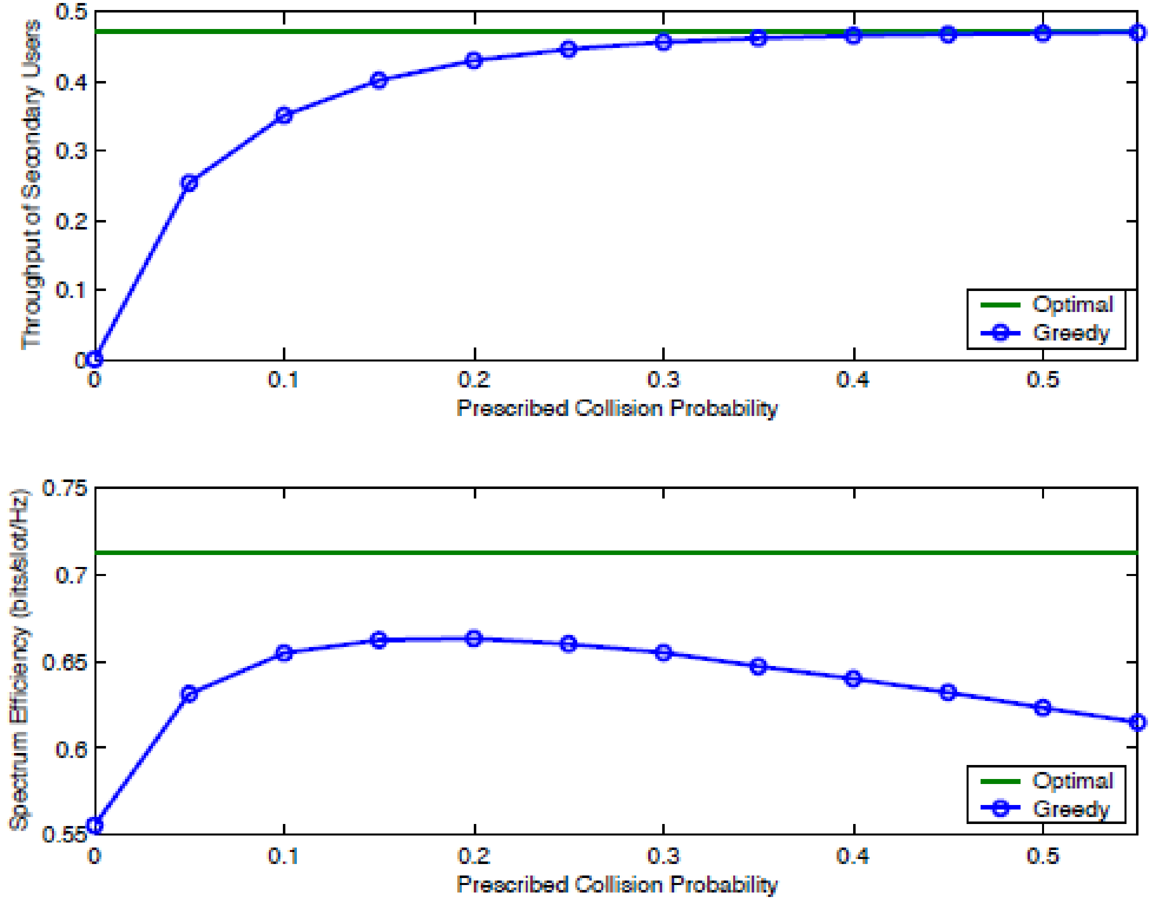


Fig. 14. OSA performance in the presence of sensing error ($N = 3$ independent channels with the same bandwidth $B = 1$ and transition probabilities $\{\alpha = 0.4, \beta = 0.5\}$).

Discussion and Extension

- What assumptions do the formulation make?
  - Known transit probabilities
  - Slotted structure
  - Error-free control message
  - Channel fading/shadowing not taken into account
- Algorithms for Dynamic Spectrum Access With Learning for Cognitive Radio, Unnikrishnan et al.
  - Argues hard to maintain synchronization
  - Centralized control with learning
  - Can deal with unknown transition probabilities
  - Don't need ACK

Summary

- Cognitive radio nodes may not be able to sense the whole spectrum
- Ad-hoc nature requires decentralized protocol
- Synchronization obtained without dedicated control
- Combination of theory from different fields
- Make certain assumptions/simplifications to separate the problems

References

- Qing Zhao, Lang Tong , Ananthram Swami , Yunxia Chen. Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework, *Selected Areas in Communications, IEEE Journal on* , vol.25, no.3, pp.589-600, April 2007.

- R. Smallwood and E. Sondik, The optimal control of partially ovservable Markov processes over a finite horizon, *Operations Research*, pp. 1071–1088, 1971.

- J. Unnikrishnan and V.V. Veeravalli, Algorithms for dynamic spectrum access with learning for cognitive radio, IEEE Trans. Signal Process., vol. 58, no. 2, pp. 750-760, Feb. 2010.