Intelligent Channel Sensing and Scheduling in Future Wireless Networks

Özgür Erçetin

Sabanci University, Istanbul, Turkey.

Nov 22, 2017
UQAM

Supported by EC H2020-MSCA-RISE-2015, by Tubitak and by British Council Institutional Links Programmes.

Outline

1. Introduction
2. Intelligent Sensing and Transmission in Energy Harvesting Devices
   - System Model
   - Transmission Policy
   - Markov Decision Process Formulation
   - Numerical Results
   - Summary
3. Intelligent Probing and Scheduling in Cellular Networks with General Fading
   - System Model
   - Scheduling
   - Gaussian Process Regression
   - Numerical Results
   - Summary

Introduction

The Setup

- Contemporary wireless access schemes need full channel state information to function optimally.
- Kar, Luo, and Sarkar (2007, 2009) have shown that the outdated channel state information results in strictly suboptimal stability region, in contrast to outdated queue information which is known to be inconsequential in terms of stable rate region.

Hence, many follow up works have focused on the problem of joint probing/learning and scheduling under various network settings and guarantees. Berry (2004, 2009), Laourine and Tong (2010), etc.

Main Research Question

How can we design an intelligent probing and scheduling algorithm taking into account probing costs AND under limited probing capability?

Main Results

1. Sensing and Transmission with an Energy Harvesting (EH) Transmitter. (*IEEE TGCN 2017*)
   - Time-varying channel with memory (Gilbert-Elliot channel).
   - Formulate optimal communication problem as a POMDP.
   - Optimal policy is a threshold type policy on the belief state on the channel state.
   - Low complexity numerical solution using policy search (reinforcement learning technique).
2. Joint probing and scheduling for server allocation in a queuing system with random connectivity. (*ComNet 2016*)
   - Stationary or Non-Stationary general fading channel model.
   - A linear combination of queue-rate (weight) and mutual information is used to determine the collection of channels to be probed at any given time.
   - Gaussian process regression estimation method is used to arrive at channel estimates at any time.

Intelligent Sensing and Transmission in Energy Harvesting Devices

Motivation

- By 2020, **over 50 billion intelligent devices** will be connected to the Internet.
- **Disposal of 300 million batteries a day** across the globe.
- Early works focus on network lifetime maximization.
- Harvesting energy from ambient energy resources is an appealing solution.

Previous Works

Main Design Issue

Energy arrives at random times, in random and minuscule amounts.

- Significant past research on how to best manage networks with random energy sources:
  - Offline scheduling policies
    M.A. Antepli, E. Uysal-Biyikoglu,H. Erkal, "Optimal packet scheduling on an energy harvesting broadcast link," IEEE JSAC, 29(8), 2011.
  - Online scheduling policies
    Z. Wang, A. Tajer, Xiaodong Wang, "Communication of energy harvesting tags," IEEE TCOM, 60(4), 2012.
  - Pilot-based channel estimation
    Y. Cheng, W. Feng, R. Shi and N. Ge, "Pilot-Based Channel Estimation for AF Relaying Using Energy Harvesting," IEEE TVT, 66(8), 2017.
  - Opportunistic transmission
    J. Pradha and S. Kalamkar and A. Banerjee, "Energy Harvesting Cognitive Radio With Channel-Aware Sensing Strategy," IEEE COM. Letters, 18(7), 2014.

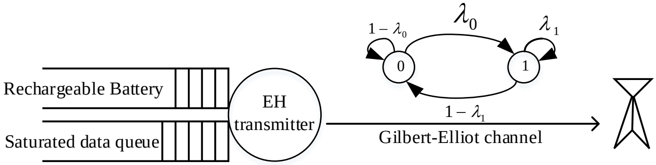
Challenge

Observation

Channel state information (CSI) acquisition improves performance significantly.

However, it consumes energy and time!

System Model



- Time-varying finite state (Gilbert-Elliot) channel.
- Channel state, G_t: Markov chain with two states: GOOD state (1) and BAD state (0).
- If $G_t = 1$: R bits per time slot; If $G_t = 0$: zero bits.
- E_t: binary harvested energy at time t; with $\Pr(E_t = 1) = q$.
- Transmission: unit energy; sensing: $0 < \tau < 1$ units.
- Finite size battery.
- ACK/NACK feedback after each transmission.

POMDP

- Partially Observable Markov Decision Process (POMDP).
  - Introduce belief state on channel state.
    - Continuous state MDP.
- System state: $S_t = (B_t, X_t)$
  - B_t: battery level at time t.
  - X_t: *belief* about channel state at time t.
    - Conditional probability of channel being in a GOOD state, given the history. $P[G_t = 1 | H_t] = p$, where H_t is history.

Transmission Policy: Deferring

At time slot t, Tx takes action $A_t \in \{D, O, T\}$:

- 1) **Defer transmission (D)**
  - No transmission
  - No feedback.
  - Belief is updated as $J(p) = p\lambda_1 + (1-p)\lambda_0$.

Transmission Policy: Sensing

- 2) **Sense the channel and transmit opportunistically (O)**
  - if channel is GOOD:
    - Transmit $(1-\tau)R$ bits in the remainder of the slot.
    - Consume one energy unit in total.
    - $p = \lambda_1$
  - if channel is BAD:
    - Remains silent.
    - Saves $1-\tau$ units of energy.
    - $p = \lambda_0$

Transmission Policy: Transmit without sensing

- 3) **Transmit without sensing (T)**
  - Transmit R bits without sensing.
  - Learns channel state thanks to the feedback.
  - Receive ACK: $p = \lambda_1$; receive NACK: $p = \lambda_0$.

MDP Formulation

- State of the system: $S_t = (B_t, X_t)$

Discounted Reward (β: discount factor)

$$V^{\pi}(b,\ p) = \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t R(S_t, A_t) | S_0 = (b,\ p)\right],$$

for all $b \in \{0, \tau, \ldots, B_{max}\}$ and $p \in [0,\ 1]$, where

$$R(S_t, A_t) = \begin{cases} X_t R & \text{if } A_t = T \text{ and } B_t \geq 1, \\ (1-\tau) X_t R & \text{if } A_t = O \text{ and } B_t \geq 1, \\ 0 & A_t = D. \end{cases}$$

Action-Value Function

- $V_A(b,\ p)$: Action-value function
  - expected infinite-horizon discounted reward of taking action A at state $(b,\ p)$.

Bellman optimality equation

$$V(b,\ p) = \max_{A \in \{D,O,T\}} \{V_A(b,\ p)\}.$$

Value Function Properties

Lemma (Increasing)

- *Value function is **increasing** in b, i.e., $V(b_1, p) \geq V(b_0, p)$ when $b_1 > b_0$.*
- *Value function is **increasing** in p, i.e., $V(b, p_1) \geq V(b, p_0)$ when $p_1 > p_0$.*

Lemma (Convex)

For any given $b \geq 0$, V(b, p) is convex in p.

Structure of the Optimal Policy

Theorem

For any $p \in [0,\ 1]$ and $b \geq 0$, there exist thresholds $0 \leq \rho_1(b) \leq \rho_2(b) \leq \rho_3(b) \leq 1$, such that, for $b \geq 1$

$$\pi^*(b,\ p) = \begin{cases} D, & \textit{if}\ \ 0 \leq p \leq \rho_1(b)\ \textit{or}\ \rho_2(b) \leq p \leq \rho_3(b) \\ O, & \textit{if}\ \ \rho_1(b) \leq p \leq \rho_2(b), \\ T, & \textit{if}\ \ \rho_3(b) \leq p \leq 1, \end{cases} \tag{1}$$

and for $\tau \leq b < 1$,

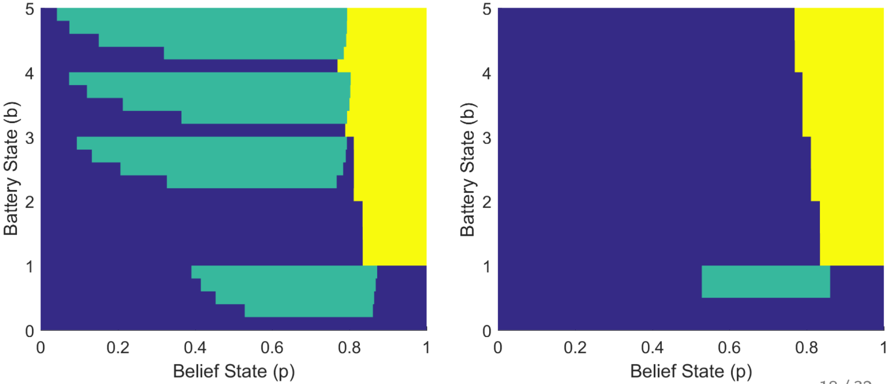
$$\pi^*(b,p) = \begin{cases} D, & \textit{if}\ \ 0 \leq p \leq \rho_1(b)\ \textit{or}\ \rho_2(b) \leq p \leq 1, \\ O, & \textit{if}\ \ \rho_1(b) \leq p \leq \rho_2(b). \end{cases} \tag{2}$$

- For $b \geq 1$, at most three thresholds.
- For $b < 1$, at most two thresholds.

Numerical Results: Effect of Sensing Duration

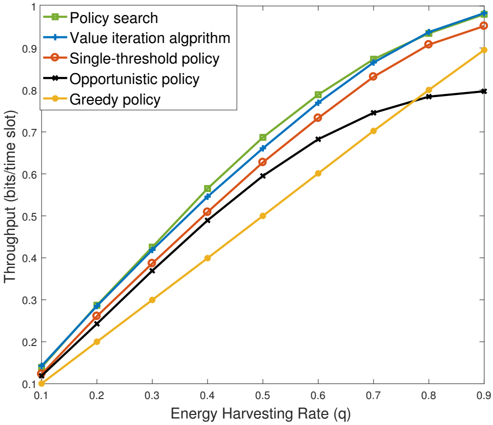
Optimal thresholds for taking actions D (blue), O (green), T (yellow) for $B_{max} = 5$, $\beta = 0.98$, $\lambda_1 = 0.9$, $\lambda_0 = 0.6$, $R = 3$ and $q = 0.1$.

- $\tau = 0.2$.
- $\tau = 0.5$.



Numerical Results: Throughput

$B_{max} = 5,\ \tau = 0.2,\ \beta = 0.999,\ \lambda_1 = 0.8,\ \lambda_0 = 0.2,\ R = 2.$



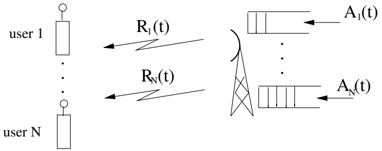
- Policy search: Optimal thresholds.
- Value iteration algorithm: Bellman optimality equations.
- Single-threshold policy: only defer or transmit.
- Opportunistic policy: always sense the channel.
- Greedy policy: transmit whenever there is energy.

Summary

- Intelligent channel sensing improves performance of EH transmitters.
- Proved that optimal policy is a battery-dependent threshold policy on belief state.
- Calculated optimal threshold values numerically using value iteration and policy search algorithms.
- Extension to multi-state generalized Gilbert-Elliot channels.
- Variable sensing cost controls channel sensing accuracy.
- **Estimating the unknown statistics using machine learning techniques.**

Intelligent Probing and Scheduling in Cellular Networks with General Fading

Motivation



- The seminal work of Tassiulas and Ephremides on server allocation in a queuing system with random connectivity showed that MaxWeight policy achieves a %100 throughput.
- Practically infeasible to obtain the full CSI.
- Outdated CSI results in strict suboptimality whereas outdated queue information inconsequential in terms of stable rate region.

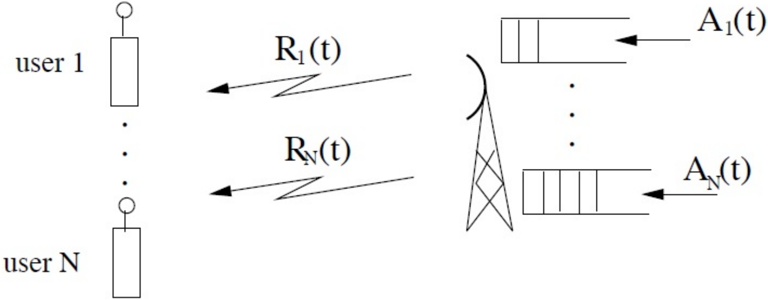
Previous Work

Main Design Issue

Channel statistics unavailable in general.

- Known underlying channel dynamics
  - A. Gopalan, C. Caramanis and S. Shakkottai. Low-delay Wireless Scheduling with Partial Channel-State Information. Proceedings of IEEE Infocom, Orlando, FL, 2012.
  - A. Duel-Hallen. Fading channel prediction and estimation for mobile radio adaptive transmission systems. Proc. of IEEE, 2007.
  - W. Ouyang, S. Murugesan, A. Eryilmaz, and N. B. Shroff. Exploiting channel memory for joint estimation and scheduling in downlink networks. Proceedings of INFOCOM, 2011.
- Gaussian process regression in wireless networks
  - D. Gu and H. Hu. Spatial gaussian process regression with mobile sensor networks. IEEE Trans. Neural Netw. Learn. Syst., 2012.
  - A. Krause, A. Singh, and C. Guestrin. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. J. Mach. Learn. Res., 2008.
  - F. Perez-Cruz, et al. Gaussian processes for nonlinear signal processing. IEEE Signal Process. Mag., 2013.

System model



- N quasi-static downlink channels.
- $c_n(t)$: channel gain, stays constant for one slot; but changes *continuously* from one slot to another.
- $R_n(t)$: achievable rate, upper bounded by Shannon rate.
- $Q_n(t)$: queue length.
- Unsaturated queues with arrivals $A_n(t)$: amount of data arriving with $\lambda_n = \mathbb{E}(A_n(t))$

Optimal Scheduling

- Λ_h: Capacity region of the network, set of all possible $(\lambda_1, \ldots, \lambda_N)$, that can be stabilized by the network.
- Max weight scheduling (MWS) achieves Λ_h by choosing:

$$n^* = \arg\min_n W_n(t) = \arg\min_n Q_n(t) R_n(t) \tag{3}$$

- However, MWS needs to know $R_n(t)$ $(c_n(t))$ for all n.
- The feedback channel can support feedback from $L < N$ users. $S(t)$: set of L users whose CSI is acquired.
- Then MWS solves the following:

$$n^* = \arg\min_{n \in S(t)} W_n(t) = \arg\min_{n \in S(t)} Q_n(t) R_n(t) \tag{4}$$

Research question

- $\pi(\eta)$: joint channel probing and scheduling policy that works together with a channel prediction algorithm η.
- Let $\Lambda^{\pi(\eta)}$ be the capacity region achieved by $\pi(\eta)$.

Research problem

Find a scheduling policy, $\pi(\eta)$, that determines $S(t)$ such that its $\Lambda^{\pi(\eta)}$ becomes as close as possible to Λ_h.

- $\hat{c}_n(t)$, $\hat{R}_n(t)$, predicted values.
- L users added to $S(t)$ with highest *estimated* $\hat{W}_n(t) = \hat{R}_n(t) \cdot Q_n(t)$.
- Let $\rho^{\pi(\eta)}(\mathbf{Q}(t)) = \mathbb{P}(W_n(t) = \hat{W}_n(t) | \mathbf{Q}(t))$.
- It can be shown that if $\rho^{\pi(\eta)}(\mathbf{Q}(t)) \geq \varepsilon$, then $\Lambda^{\pi(\eta)} \subseteq \varepsilon \cdot \Lambda_h$

Adding the element of surprise!

- Some users with smaller backlogs may not be probed for a long time!
- We need to modify $\pi(\eta)$.
- $I_n^{\pi(\eta)}(t)$: information of an unexplored channel.
- Information of a channel observed recently and many times before is less than the channel not been probed for a long time or one that varies rapidly.
- Hence, $S(t)$ is determined according to Multi-Objective Scheduling and Feedback (MOSF) algorithm as follows:
  - $\hat{W}_n(t) = Q_n(t) * \hat{R}_n(t) + \zeta I_n(t)$
  - add L users with highest values of $\hat{W}_n(t)$ to $S(t)$.

Prediction using Gaussian Process Regression (GPR)

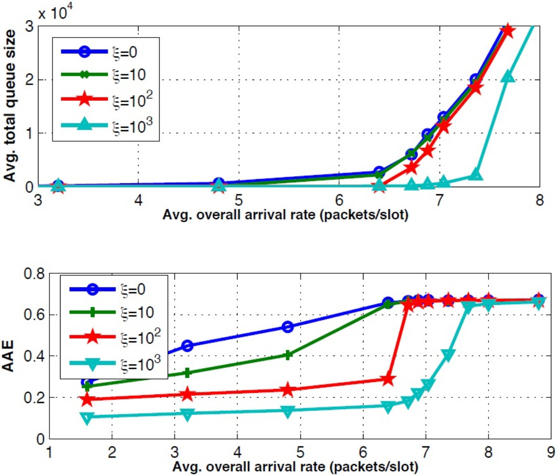
- Predict the value of $\hat{R}_n(t)$ using GPR:
  - $\mathbf{c}_n = (c_n^1, c_n^2, \ldots, c_n^w)$ latest w CSI values.
  - taken at times $\tau = (\tau_n^1, \tau_n^2, \ldots, \tau_n^w)$ for user n before t.
  - Gaussian kernel function, $k_n(\tau_n^i, \tau_n^j)$ describes the correlation of channel n between two of its measurements taken at times τ_n^i, and τ_n^j.
- Posterior distribution of $c_n(t)$ given $\mathbf{c}_n$ and τ is Gaussian with mean $\hat{c}_n(t)$ and variance $v_n(t)$ calculated as follows.

$$\hat{c}_n(t) = \mathbf{k}_n^T(t)\mathbf{K}_n\mathbf{c}_n, \qquad v_n(t) = k_n(t,t) - \mathbf{k}_n^T(t)\mathbf{K}_n\mathbf{k}_n(t) \tag{5}$$

  where $\mathbf{k}_n(t) = [k_n(\tau_n^i, t)]_i$.
- In GPR method it turns out that: $\arg\max_n I_n(t) = \arg\max_n v_n(t)$

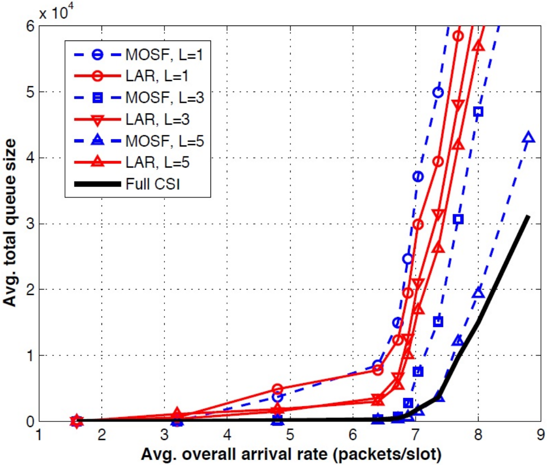
Numerical Results

- AAE: average channel estimation error.
- $L = 4$, $N = 16$ by MOSF algorithm.



Numerical Results

- LAR: by p previous values of CSI uses Auto-regression (AR) model to predict the CSI.
- LAR probes L users with the highest backlog-estimated rate product.



Summary

- Cost and practicality of obtaining CSI cannot be neglected.
- In general, wireless channel statistics is unknown and possibly non-stationary.
- Developed a learning framework to probe/learn/schedule wireless users with provable performance.
- Proving performance guarantees difficult with learning algorithms in wireless networks demanding further research.

Thank You!

Collaborators

Mehdi Abad (Sabanci University), Deniz Gunduz (Imperial College London), Mehmet Karaca (Ericsson), Tansu Alpcan (University of Melbourne).