Communication over a Gilbert-Elliot Channel with an Energy Harvesting Transmitter

Mehdi Salehi Heydar Abad[1] , Deniz Gündüz[2] , Ozgur Ercetin[1]

[1] Sabanci University, Istanbul, Turkey.
[2]Imperial College London, UK.

August, 2017

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

Motivation

- By 2020, **over 50 billion intelligent devices** connected to 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 Work

- 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!

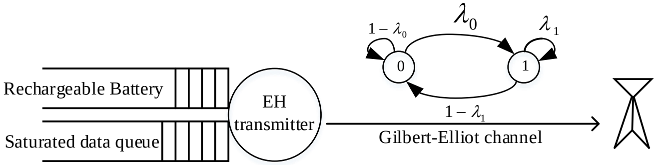
Main Results

How to balance CSI acquisition cost with performance loss from stale CSI?

Main Results

- 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).

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.

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* $b \geq 1$

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

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

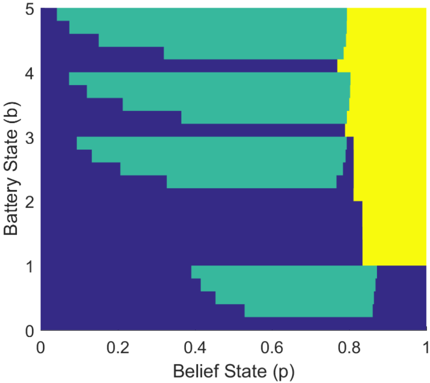
$$\pi^*(b,p) = \begin{cases} D, & \text{if } 0 \leq p \leq \rho_1(b) \text{ or } \rho_2(b) \leq p \leq 1, \\ O, & \text{if } \rho_1(b) \leq p \leq \rho_2(b). \end{cases} \quad (2)$$

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

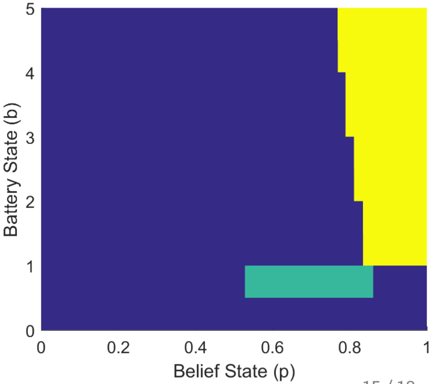
Numerical Results: Effect of Sensing Duration

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 $q = 0.1$.

- $\tau = 0.2$.

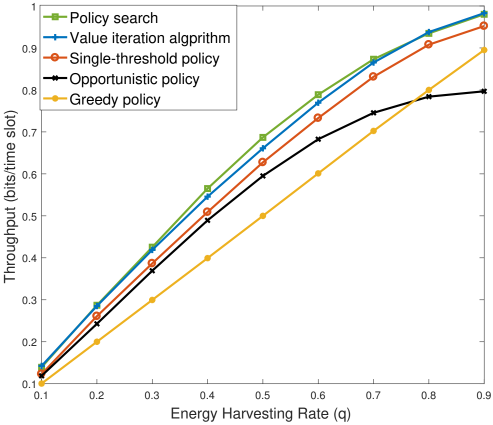


- $\tau = 0.5$.



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.

Conclusions and Future Work

- 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 unknown statistics using machine learning techniques.

Thank You!