Transmit Power Control in Energy Harvesting Sensors: A Decision-Theoretic Approach

Anup Aprem
aaprem@ece.iisc.ernet.in

Advisor Dr Chandra R Murthy

SPC Lab
ECE Department
Indian Institute of Science, Bangalore

March 31, 2012

Table of Contents

Table of Contents

System Model I

- Measurement data to be periodically sent once in every measurement of interval of duration T_m.
- Time discretized into slots of duration T_p sec. $T_m/T_p = K$ denote the maximum number of retransmissions in a measurement interval
- Transmission Protocol
  - ARQ protocol data transmission
  - A NACK received triggers retransmission of packet
- Harvesting and Storage Model
  - Energy E_s harvested at each slot with probability ρ
  - Battery of (large) finite capacity
- Sensor Model
  - E - Minimum Transmission power possible. Usually limited by hardware specifications of EH node
  - Define $\frac{E_s}{E} = L \quad L \in \mathbb{N}$. Hereafter all power normalized w.r.t E

System Model II

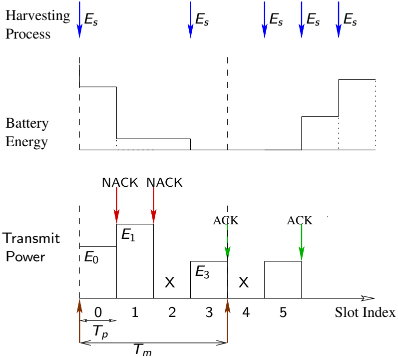

Figure: Transmission timeline of the EH node for $K = 4$, showing the random energy harvesting process ($\downarrow$) and periodic data arrival ($\uparrow$).The marker "X" denotes slots where the EHS does not transmit data

Channel Model

1. Time Varying Channel modelled using FSMC.

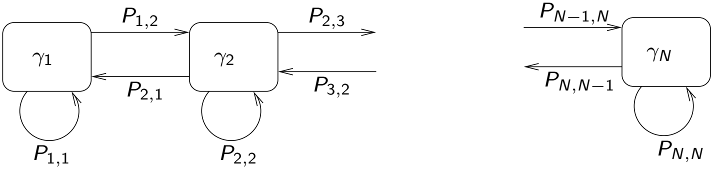

Figure: Finite State Markov Model for Rayleigh Fading Channel

2. Block-Fading Channel across measurement intervals.

Objective

Aim

Maximize the probability of packet reception at the destination / Minimize the probability of outage.

Outage occurs in a given measurement interval

- Receiver fails to decode packet even after successive retransmissions.
- EHS transmitter doesn't have enough energy to retransmit.

Objective

Aim

Maximize the probability of packet reception at the destination / Minimize the probability of outage.

Outage occurs in a given measurement interval

- Receiver fails to decode packet even after successive retransmissions.
- EHS transmitter doesn't have enough energy to retransmit.

Power Algorithm

Power management across successive transmissions to minimize the outage probability.

Table of Contents

Literature Survey I

- [Kansal et al., 2007]
  - System Design Aspects (Energy Harvesting Profile)
  - Use of storage buffer to reduce the randomness in the energy harvesting process.
- [Lei et al., 2009] Introduced Probabilistic Bernoulli injection model
- [Medepally et al., 2009] Fixed Power Retransmissions with Bernoulli injection model.
- More general models for data and energy arrival considered in [Sharma et al., 2010, Yang and Ulukus, 2010, Tutuncuoglu and Yener, 2011]. Link Design to optimize various metrics.

Motivation

Motivation

- Retransmission of the packet is triggered by channel errors
  - Knowledge of channel time correlation can be used to increase or decrease the transmission power.
  - Tradeoff between power of retransmitted packet and power of future packets.
- The number of NACK's received give information about channel. Can be exploited to optimize outage probability.

Optimum Power Algorithm

Can be cast as a Markov Decision Problem

Table of Contents

Markov Decision Process (MDP)

- Planning Under Uncertainty.

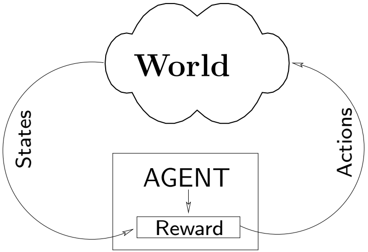

Figure: MDP model showing the interaction between agent and real world

MDP Formal Definition

- Formally, A Markov Decision Process is $\langle \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R} \rangle$, where
  - $\mathcal{S}$ is the finite set of states s
  - $\mathcal{A}$ is the finite set of actions a
  - $\mathcal{T}(s, a, s') = Pr(s'/s, a)$
  - $\mathcal{R}(s, a)$
- Objective: Obtain an optimal policy, $\pi : \mathcal{S} \rightarrow \mathcal{A}$
- Solution to MDP
  - Finite Horizon: Dynamic Programming
  - Infinite Horizon: Policy Iteration, Value Iteration

Partially Observable Markov Decision Process (POMDP)

Planning under uncertainty in partially observable environments

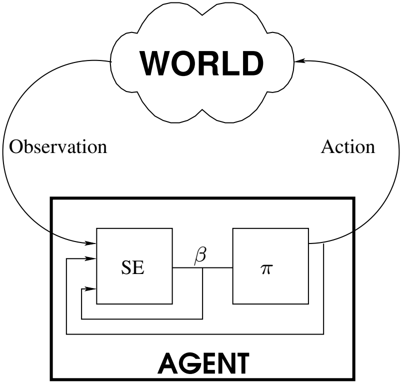

Figure: A POMDP agent can be decomposed into a state estimator (SE) and a policy (π).

Formal Definition of POMDP

Formally, POMDP is $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, \mathcal{T}, \mathcal{Z}, \mathcal{R} \rangle$

- $\mathcal{S}$, $\mathcal{A}$, $\mathcal{T}$ and $\mathcal{R}$ describe a Markov Decision Process.
- $\mathcal{O}$ Finite set of observations.
- Observation function $\mathcal{Z} : \mathcal{S} \times \mathcal{A} \rightarrow \Pi(\mathcal{O})$ which gives for each action a and resulting state s, a probability distribution over the observation states.

POMDP Solution

- Belief State $\beta(s)$: Distribution over the state space
- POMDP Control can be split to
  - State Estimator: Updating the belief state based on the last action, the current observation, and the previous belief state.
  - π: Taking actions as a function of the agent's belief state.
- Solution
  - Witness Algorithm [Littman, 1994]
  - Incremental Pruning [Cassandra et al., 1997]
  - PERSUS [Spaan and Vlassis, 2005]

Table of Contents

EHS POMDP: State, Observation & Action Space

- State Space $\mathcal{S} \triangleq \mathcal{B} \times \mathcal{G} \times \mathcal{U} \times \mathcal{K}$, where
  - $\mathcal{B}$ is the set of battery states, $\mathcal{B} \triangleq \{0, 1, \ldots, B_{\text{max}}\}$
  - $\mathcal{G}$ is the set of channel states.
  - $\mathcal{U} \triangleq \{0, 1\}$ is set of packet reception states. The packet reception state takes the value "1" when an ACK is received by the EHS, and "0" otherwise.
  - $\mathcal{K}$ is the set of packet transmission attempt indices within a frame, and hence, $\mathcal{K} \triangleq \{0, 1, \ldots, K-1\}$.
- Observation Space $\mathcal{O} \triangleq \{\text{ACK}, \text{NACK}\}$.
- Action Space $\mathcal{A} \triangleq \{0, 1, \ldots, B\}$, with $B \in \mathcal{B}$

EHS POMDP: Transition Function I

Let $s \triangleq (b, \gamma, u, k)$, $s' \triangleq (b', \gamma', u', k')$

- Correlated Case

$$\mathcal{T}(s, a, s') = \delta(k', k_+)\, P_{\gamma,\gamma'}\, \psi\left((b, u), a, (b', u')\right)$$

where $k_+ \triangleq (k+1) \bmod K$

- Block Fading Case

$$\mathcal{T}(s, a, s') = \delta(k', k_+)\, \zeta(\gamma, \gamma'; k)\, \psi\left((b, u), a, (b', u')\right)$$

where

$$\zeta(\gamma, \gamma'; k) = \begin{cases} \delta(\gamma, \gamma') & k \neq K-1 \\ \pi_{\gamma'} & k = K-1 \end{cases}$$

EHS POMDP: Transition Function II

$$
\begin{aligned}
\psi\left((b,u), a, (b',u')\right) = {} & \rho\delta(b', b+L)\delta(u')\delta(u) \\
& + (1-\rho)\delta(b', b)\delta(u')\delta(u) \\
& + (1-\rho)(1-P_e(aE;\gamma))\delta(b', b-a)\delta(u')\delta(1-u) \\
& + \rho(1-P_e(aE;\gamma))\delta(b', b-a+L)\delta(u')\delta(1-u) \\
& + (1-\rho)P_e(aE;\gamma)\delta(b', b-a)\delta(1-u')\delta(1-u) \\
& + \rho P_e(aE;\gamma)\delta(b', b-a+L)\delta(1-u')\delta(1-u)
\end{aligned}
$$

Alternate Form

EHS POMDP: Observation Function, Reward & Objective

- Observation Function

$$\begin{aligned} P\left(\text{NACK}/a,\gamma\right) &= P_e(aE;\gamma) \\ P\left(\text{ACK}/a,\gamma\right) &= 1 - P_e(aE;\gamma), \end{aligned} \tag{1}$$

- Reward

$$\mathcal{R}(s,a) = \begin{cases} 1 - P_e(aE;\gamma) & a \leq b, \quad u = 0 \\ -1 & a > b, \quad u = 0 \\ -1 & a \neq 0, \quad u = 1 \\ 0 & \text{else} \end{cases}$$

- Objective: Maximize over an infinite horizon the expected reward collected by the EHS node and is given by

$$J = \lim_{m\rightarrow\infty} \frac{1}{m}\mathbb{E}\left\{\sum_{n=1}^{m} \mathcal{R}(s_n, a_n)\right\}$$

where $n \in \{1, 2, \ldots, \}$ denotes the time step, and s_n (and a_n) is the state (and action) sequence

Table of Contents

Solution to EHS POMDP

- General POMDP solutions is PSPACE-complete.
- Explore two heuristic solution.
- Heuristic Solution depend on the underlying MDP.
- Optimality Equation for MDP.

$$\lambda^* + h^*(s) = \max_{a \in \mathcal{A}, a \leq B(s)} \left[\mathcal{R}(s,a) + \sum_{s' \in \mathcal{S}} \mathcal{T}(s,a,s') h^*(s') \right], \forall s \in \mathcal{S}$$

- Above equation solved by value iteration

$$J_{k+1}(s) = \max_{a \in \mathcal{A}, a \leq B(s)} \left[\mathcal{R}(s,a) + \sum_{s' \in \mathcal{S}} \mathcal{T}(s,a,s') J_k(s') \right], \forall s \in \mathcal{S}$$

J_k is the value function at the k'th iteration, $k = 0, 1, \ldots$.
It can be shown that [Bertsekas, 2005]

$$\lim_{k \to \infty} \frac{J_k(s)}{k} = \lambda^*, \quad \forall s \in \mathcal{S}.$$

Belief Update

- Battery, Packet reception and Slot state fully observable.
- Channel state only partially observable through the ACK/NACK messages.
- Belief over the channel state $\beta(\gamma)$.
- Belief update equation

$$\beta_n(\gamma_j) = \frac{P(o_n/a_n, \gamma_j) \sum_{i=1}^{N} P_{\gamma_i,\gamma_j} \beta_{n-1}(\gamma_i)}{\sum_{o' \in \mathcal{O}} P(o'/a_n, \gamma_j) \sum_{i=1}^{N} P_{\gamma_i,\gamma_j} \beta_{n-1}(\gamma_i)},$$

for $j = 1, 2, \ldots, N$

ML Heuristic [Cassandra, 1998]

- ML state of channel: $\gamma_{\text{ML}} \triangleq \arg\max_{\gamma \in \mathcal{G}} \beta(\gamma)$
- ML state of system: $s_{\text{ML}} \triangleq (b, \gamma_{\text{ML}}, u, k)$
- ML Heuristic Solution of POMDP.

$$\mu_{\text{ML}} = \mu^*_{\text{MDP}}(b, \gamma_{\text{ML}}, u, k).$$

Voting Heuristic [Simmons and Koenig, 1995]

- Each state votes for an action as determined by the optimal policy of underlying MDP.
- Vote for each action weighted by the belief of the state.
- Sum of all weighted votes for each action is determined.
- Action with largest sum selected as the optimal action.

$$\mu_{\text{voting}} = \arg\max_{a \in \mathcal{A}} \sum_{\substack{s=(b,\gamma,u,k) \\ \gamma \in \mathcal{G}}} \beta(s)\delta(\mu^*_{\text{MDP}}(s) - a).$$

Table of Contents

Simulation Results I

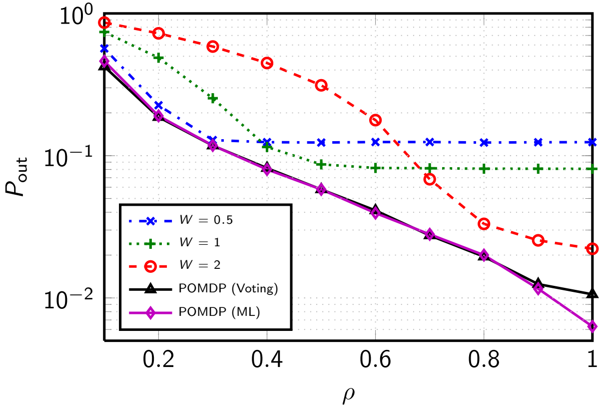

Figure: Correlated channel case: outage probability vs. ρ. System Parameters: $K = 3$, $L = 4$, $N = 7$, $\ell = 50$, $E_s = 12$ dB (relative to N_0), $N_0 = 1$ and $B_{max} = 10E_s$.

Simulation Results II

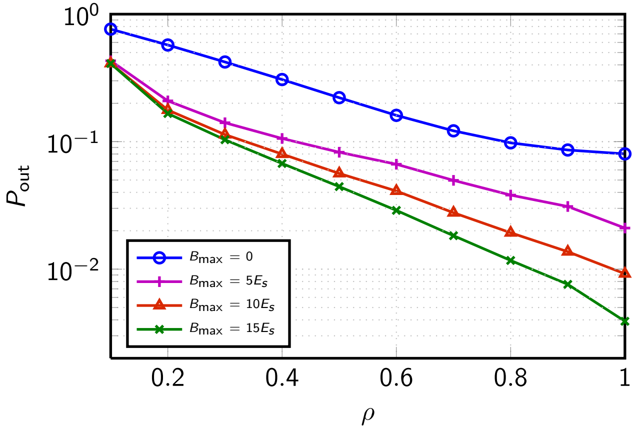

Figure: Correlated channel case: outage probability vs ρ for different values of the battery capacity

Simulation Results III

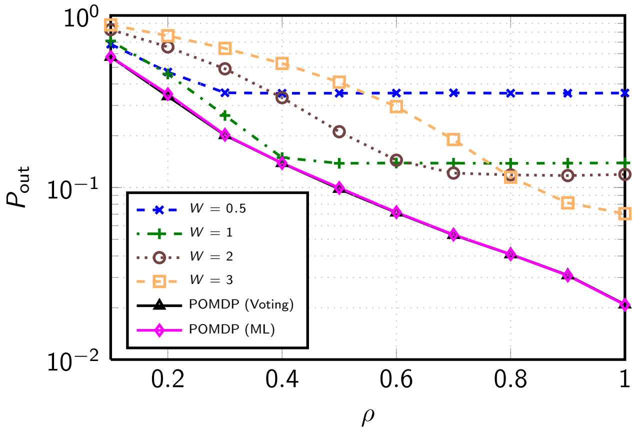

Figure: Block fading channel case: outage probability vs. ρ. System Parameters: $K = 4$, $L = 4$, $N = 7$, $\ell = 50$, $E_s = 12$ dB (relative to N_0), $N_0 = 1$ and $B_{\max} = 20E_s$.

Table of Contents

Conclusion and Future Work

- Conclusion
  - Considered the problem of finding optimal power policy for EHS node with no CSI
  - Problem cast in POMDP framework and Heuristic solutions found.
  - Heuristic Solutions performs better than existing algorithms
- Future Work
  - Hybrid ARQ(HARQ) schemes, shown to be energy efficient could be implemented in EHS.
  - Maximizing the average rate by adapting the modulation and coding scheme.


THANK
YOU
THANK

Transition Function: Alternate Form of ψ

Back

$$\psi\left((b,u),a,(b',u')\right)=\begin{cases}\rho & \text{if } \begin{cases} b'=b+L, \\ u'=1, \quad u=1\end{cases} \\ 1-\rho & \text{if } \begin{cases} b'=b, \\ u'=1, \quad u=1\end{cases} \\ (1-\rho)(1-P_e(a;\gamma)) & \text{if } \begin{cases} b'=b-a, \\ u'=1, \quad u=0\end{cases} \\ \rho(1-P_e(a;\gamma)) & \text{if } \begin{cases} b'=b-a+L, \\ u'=1, \quad u=0\end{cases} \\ (1-\rho)P_e(a;\gamma) & \text{if } \begin{cases} b'=b-a, \\ u'=0, \quad u=0\end{cases} \\ \rho P_e(a;\gamma) & \text{if } \begin{cases} b'=b-a+L, \\ u'=0, \quad u=0\end{cases} \\ 0 & \text{else}\end{cases}$$

Value Function

- A policy π is a mapping from $\mathcal{S}$ to $\mathcal{A}$ which gives action to select in each state.
- Value of a state is expected long term return starting from that state.

$$V_\pi(s) = R(s, \pi(s)) + \gamma \sum_{s' \in \mathcal{S}} T(s, \pi(s), s') V_\pi(s')$$

Value Iteration

$$V_{k+1}(s) = \max_{a \in \mathcal{A}} \left[R(s,a) + \gamma \sum_{s' \in \mathcal{S}} T(s, \pi(s), s') V_k(s') \right]$$

Bibliography I

Bertsekas, D. (2005).
Dynamic Programming and Optimal Control 3rd Edition, Volume I.

Cassandra, A., Littman, M., Zhang, N., et al. (1997).
Incremental pruning: A simple, fast, exact method for partially observable Markov decision processes.
In *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence*, pages 54–61. Citeseer.

Cassandra, A. R. (1998).
Exact and Approximate Algorithms for Partially Observable Markov Decision Processes.
PhD thesis.

Kansal, A., Hsu, J., Zahedi, S., and Srivastava, M. (2007).
Power management in energy harvesting sensor networks.
ACM Trans. Embedded Computing Systems (TECS), 6(4):32.

Bibliography II

Lei, J., Yates, R., and Greenstein, L. (2009).
A generic model for optimizing single-hop transmission policy of replenishable sensors.
IEEE Trans. Wireless Commun., 8(2):547–551.

Littman, M. (1994).
The witness algorithm: Solving partially observable Markov decision processes.
Brown University, Providence, RI.

Medepally, B., Mehta, N., and Murthy, C. (2009).
Implications of energy profile and storage on energy harvesting sensor link performance.
In *Proc. Globecom*, pages 1–6. IEEE.

Sharma, V., Mukherji, U., Joseph, V., and Gupta, S. (2010).
Optimal energy management policies for energy harvesting sensor nodes.
IEEE Trans. Wireless Commun., 9(4):1326–1336.

Bibliography III

Simmons, R. and Koenig, S. (1995).
Probabilistic robot navigation in partially observable environments.
In *International Joint Conference on Artificial Intelligence*, volume 14, pages 1080–1087.

Spaan, M. and Vlassis, N. (2005).
Perseus: Randomized point-based value iteration for POMDPs.
Journal of Artificial Intelligence Research, 24(1):195–220.

Tutuncuoglu, K. and Yener, A. (2011).
Short-term throughput maximization for battery limited energy harvesting nodes.
In *Proc. ICC*, pages 1–5.

Yang, J. and Ulukus, S. (2010).
Transmission completion time minimization in an energy harvesting system.
In *Proc. Conf. on Inform. Sci. and Syst. (CISS)*, pages 1–6. IEEE.