Throughput Maximization for a Primary User with Cognitive Radio and Energy Harvesting Functions

Thanh-Tung Nguyen[1] and Insoo Koo[2]
[1] School of Electrical Engineering, University of Ulsan
680-749 San 29, Mugeo 2-dong, Ulsan, Republic of Korea
[e-mail: nguyenthanhtung@ctu.edu.vn]
[2] School of Electrical Engineering, University of Ulsan
680-749 San 29, Mugeo 2-dong, Ulsan, Republic of Korea
[e-mail: iskoo@ulsan.ac.kr]
*Corresponding author: Insoo Koo

Received January 2, 2014; revised March 29, 2014; revised June 20, 2014; accepted August 19, 2014; published September 30, 2014

Abstract

In this paper, we consider an advanced wireless user, called *primary-secondary user (PSU)* who is capable of harvesting renewable energy and connecting to both the primary network and cognitive radio networks simultaneously. Recently, energy harvesting has received a great deal of attention from the research community and is a promising approach for maintaining long lifetime of users. On the other hand, the cognitive radio function allows the wireless user to access other primary networks in an opportunistic manner as secondary users in order to receive more throughput in the current time slot. Subsequently, in the paper we propose the channel access policy for a PSU with consideration of the energy harvesting, based on a Partially Observable Markov decision process (POMDP) in which the optimal action from the action set will be selected to maximize expected long-term throughput. The simulation results show that the proposed POMDP-based channel access scheme improves the throughput of PSU, but it requires more computations to make an action decision regarding channel access.

Keywords: Primary-secondary user, energy harvesting, cognitive radio, channel access policy, Partially Observable Markov decision process (POMDP)

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (No. NRF-2013R1A2A2A05004535)

http://dx.doi.org/10.3837/tiis.2014.09.007

1. Introduction

The frequency spectrum is a limited resource and is being wasted by the fixed spectrum assignment policies of governmental agencies [1]. Cognitive radio (CR) technology allows cognitive radios to exploit both the unused frequency bands and the idle time slots in a primary channel in an opportunistic manner. Secondary users (SUs) are equipped with CR-enabled devices and can access the spectrum dynamically. When supported with the spectrum sensing capability, SUs can discover and select frequency bands that are suitable for use, and can change to other frequency bands when they detect a primary user presence.

From the primary user perspective, primary users can also benefit from the cognitive radio function if they are well integrated with cognitive-enabled radio devices for the purpose of achieving more throughput, which is defined as the number of bits transmitted over the network in a time unit, from the under-used channels of other primary networks. For example, a primary user who is implementing several online applications needs more bandwidth to satisfy the large amount of data being transferred over the network to improve or guarantee the QoS of these applications. It is obvious that, due to the functions of the primary user and secondary user, whenever primary users access their primary network, they have highest priority to use the channels assigned by the network operator, whereas when they utilize the other licensed channels, they have the same priority as other secondary users in accessing these primary channels. Note that when the primary users act as the secondary user, they not only avoid interference with the primary users of these primary networks, but also contend with other secondary users to occupy the channels. Collisions with other secondary users also may occur. Therefore, a novel access method for primary users with the cognitive radio function is required.

Subsequently, in the paper we propose a channel access policy for a primary user equipped with cognitive-enabled transceivers, called a *primary-secondary user* (PSU) or hybrid user, with consideration of the energy harvesting in order that the PSU can access both its primary network and secondary networks, and at the same time the PSU has a rechargeable battery powered by an energy harvester. The proposed access scheme is based on a Partially Observable Markov decision process (POMDP) by which we can select the optimal action from the action set to maximize expected long-term throughput. To the best of our knowledge, the throughput analysis and access policy for the primary and secondary transceivers with consideration of energy harvesting are investigated in this paper for the first time.

The remainder of this paper is organized as follows. In Section 2, we describe the related works. In Section 3, the system and the energy model are presented. In Section 4, we investigate the channel access policy for the PSU based on POMDP to find the optimal action which aims at maximizing the expected long-term throughput. In Section 5, the simulation results are presented to illustrate our problem investigation. Finally, Section 6 concludes our works.

2. Related works

Energy consumption should be addressed within all operations of any electronic devices. Energy efficiency is an interesting issue that attracts a great deal of attention from the research community, especially in the area of wireless communication. Many works in the literature have investigated energy issues. In the context of the fight against climate change, the target is to reduce the carbon emissions by 20% by 2020 [1]. Accordingly, the objective of wireless device design is to minimize energy consumption and to allow the device to maintain a connection to the network for a longer time. This issue is also related to the development of energy-efficient spectrum assignment algorithms. In most practical situations, the wireless users are battery-powered devices, and the spectrum sensing and data transmission procedures usually accompany energy consumption. In [2], the author investigated the optimal cognitive sensing and access policies for a secondary user with an energy harvester. The problem was formulated as a Markov decision process (MDP) to find an action to maximize the throughput. In [3], the authors investigated the problem of developing an energy efficient opportunistic spectrum access strategy for a secondary user with energy harvesting capability. They formulated the problem to determine the optimal sensing and access policy as a partially observable Markov decision process (POMDP). A cognitive radio network with an energy-harvesting secondary transmitter to improve both energy efficiency and spectral efficiency was presented in [4] to determine an optimal spectrum sensing policy that maximizes the expected total throughput subject to an energy causality constraint and a collision constraint. Optimal energy management has been addressed in several works such as [5], where the authors present throughput optimal and mean delay optimal energy management policies for an energy harvesting sensor node. The authors of [6] investigate spectrum-sensing policies for an energy-constrained cognitive terminal in considering the dynamics of the primary networks to determine the spectrum sensing duration for performance optimization. In [7], the authors proposed cognitive MAC protocols in cognitive radio decentralize network to allow secondary users to access the channel. The POMDP was used for the access decision. However, all mentioned works are only considered for either the secondary user or the primary user. Further joint investigations of a primary user equipped with cognitive-enabled transceivers and energy harvesting are not studied.

3. System and Energy Model

In this paper, we consider a PSU with two transceivers in which one transceiver has a cognitive capability. Consequently, the PSU can exploit two channels. The first channel is a primary channel of its licensed network and the second channel is also a primary channel which is managed by another primary network and is hereafter called the harvested channel. Normally, the operation of the two networks is completely independent but we assume that these two channels operate in time-slotted fashion with equal durations. The PSU will access its primary channel using the primary transceiver in time slots assigned by its network operator whereas utilizing free time slots of the harvested channel using its cognitive-enable transceiver for data transmission.

In this section, we first detail model of the primary channel via a slot assignment policy and model of the harvested channel. We then present spectrum sensing technique before investigating throughput of the PSU over the harvested channel and the primary channel in a time slot. Next, we describe model of energy harvesting which is related to energy arrival and battery state.

3.1 Slot assignment policy for PSU on the primary channel

Due to a primary user, the PSU accesses a primary channel of its licensed network using the primary transceiver. The primary channel is assumed to synchronize with the time structure of the harvested channel. Obviously, there may be many primary users sharing the primary channel. Network operator will assign each primary user time slots for before accessing the channel. It is assumed that the slot assignment for the PSU follows the Bernoulli process and that the probability mass function (PMF) of the slot assignment is then expressed as

$$S_k = \begin{cases} \Pr\{X_k = 1\} = p_a \\ \Pr\{X_k = 0\} = 1 - p_a \end{cases} \tag{1}$$

where $X_k = 1$ means that the primary channel is assigned to the considered PSU. It is noteworthy that S_k depends on the number of users sharing the same primary channel with the PSU.

3.2 Harvested channel model

As mentioned above, the PSU uses the cognitive-enable transceiver to utilize a harvested channel of another primary network for opportunistic data transmission. To access the harvested channel efficiently, statistical information regarding the harvested channel usage is gathered by observing user activities in the harvested channel. In cognitive radio, activities of primary users that are the presence or absence of the users on the channel can be sensed by spectrum sensing technique and is presented in the next subsection. Based on the statistics of the user activities in the harvested channel, the states of the harvested channel are known to the PSU and are modeled as a stationary Markov chain with two states: Active (A) and Silent (S), as shown in **Fig. 1**.

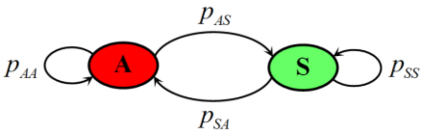

Fig. 1. The Markov chain of the harvested channel states

From the above figure, State A indicates that the harvested channel is being occupied by a user and that the PSU will not have an opportunity to utilize this channel. Conversely, State S indicates that the harvested channel is available to others and can be accessed by the PSU. Specifically, the transition probabilities from State A to State S and from State S to itself are p_{AS} and p_{SS}, respectively. These two parameters directly determine the possibilities that the PSU will occupy the harvested channel. From the perspective of the cognitive radio network, to exploit a primary channel effectively, a sensor network can be deployed to sense the primary channels in the long time and to make statistics of primary channels in the network. By this observation method, the value of p_{SS} and p_{AS} of a primary channel can be estimated and shared with the other secondary users to utilize the primary channels.

3.3 Spectrum sensing

Spectrum sensing is a mandatory step before using the harvested channel to avoid interference with primary user. In the case where the PSU wants to transmit data on the harvested channel, it first senses the channel using the spectrum sensing techniques to decide whether the channel is in state A or state S. If the harvested channel is in state S, then the PSU can send data over the channel; otherwise, it will wait for another chance in the next time slots. In the paper, the energy detection is used for spectrum sensing and the PSU spends the duration τ $(\tau \le T)$ at the beginning of the time slot for energy detection. Given a sensing power P_s, which is the required power to perform spectrum sensing, we now assume that the sensing time τ is proportional to the sensing energy e_s and can be expressed as follows

$$\tau = \frac{e_s}{P_s} \tag{2}$$

We further assume that given a sampling frequency of the energy on the harvested channel f_s, the number of energy samples taken in τ will be τf_s, and this value affects two important parameters: the probability of detection, P_D, and the probability of false alarm, P_F, which determine the quality of the energy detector. P_F denotes the probability that the PSU decides that the harvested channel is in state A while it is actually in state S, and P_D denotes the probability that the PSU correctly decides that the harvested channel is in state A. The relation between the probability of detection and the probability of false alarm is mentioned in [8] and given as follows:

$$P_F(e_s) = Q\left[\sqrt{2\beta+1}.Q^{-1}\left(P_D^*\right) + \beta\sqrt{\tau.f_s}\right] \tag{3}$$

$$P_D(e_s) = Q\left\{\frac{1}{\sqrt{2\beta+1}}.\left[Q^{-1}\left(P_F^*\right) - \beta\sqrt{\tau.f_s}\right]\right\} \tag{4}$$

where β is the received signal-to-noise ratio (SNR) on the harvested channel measured at the cognitive receiver. P_D^* is the designed probability of detection of the energy detector. f_s is the frequency of sampling energy on the harvested channel, and $Q(.)$ is the complementary distribution function of a standard Gaussian random variable, which is given by

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{\infty} e^{\frac{-t^2}{2}}\,dt \tag{5}$$

Obviously, from (3) and (4), with a given sampling frequency f_s and a received signal-to-noise ratio β at the energy detector, the sensing time τ will determine the detection performance of the energy detector of the PSU.

3.4 Throughput over harvested channel in a time slot

After performing the spectrum sensing, if the state of the harvested channel is determined as state S then the PSU can transmit data to its respective receiver. In the current time slot k, if the harvested channel is actually being in state S then the probability that the PSU uses the harvested channel is $1-P_F(e_s)$. In such case, the expected throughput obtained in this time slot can be calculated by

$$r\left(e_s\right) = \left[1-P_F\left(e_s\right)\right]W_{hc}\left(T-\tau\right)\log_2\left(1+SNR_{hc}\right) \tag{6}$$

where SNR_{hc} is the signal-to-noise ratio on the harvested channel when the transmission power of the PSU on the harvested channel is P_{hc}, e_{hc} is the energy expended in transmission on the harvested channel, and W_{hc} is the bandwidth of the harvested channel.

From (6) and (3), it can be observed that, given a value of the probability of detection P_D^*, the throughput is a function of the sensing energy e_s. It is inferred that the more energy the PSU consumes, the more the sensing time increases and the smaller the probability of false alarm would be, thus leading to more reliability in the detection of state S. Therefore, with the policy of throughput maximization, it is necessary to find the optimal amount of energy for spectrum sensing in order to maximize the expected throughput in the current time slot k. The optimal sensing energy can be found by solving (6) as follows:

$$e_s^* = \max_{0<e_s\le e_{s,\max}} r\left(e_s\right) \tag{7}$$

where $e_{s,\max} = \min(e, T.P_s)$ corresponds to the maximum energy spent for sensing in the duration $\tau = T$.

Fig. 2 shows an example of maximizing the throughput to find a value of the sensing energy in a time slot when the probability of detection P_D^* is 0.8 and other parameters are given in **Table 2** as follows: $W_{hc} = 3.0Mhz$, $f_s = 800Khz$, $SNR_{ss} = -15dB$, $SNR_{hc} = 10dB$ and $P_s = 40mW$. In this case, the optimal sensing energy is 0.1712 (mJ) and the optimal sensing time is 0.0043s.

As a result, the maximum throughput in the time slot k is updated as

$$r_{hc}\left(e_{sa}\right) = r\left(e_s^*\right) \tag{8}$$

where $e_{sa} = e_s^* + (\mathrm{T}-\tau)\mathrm{P}_{hc}$ is the energy which is required for the spectrum sensing and transmission on the harvested channel in the current time slot k.

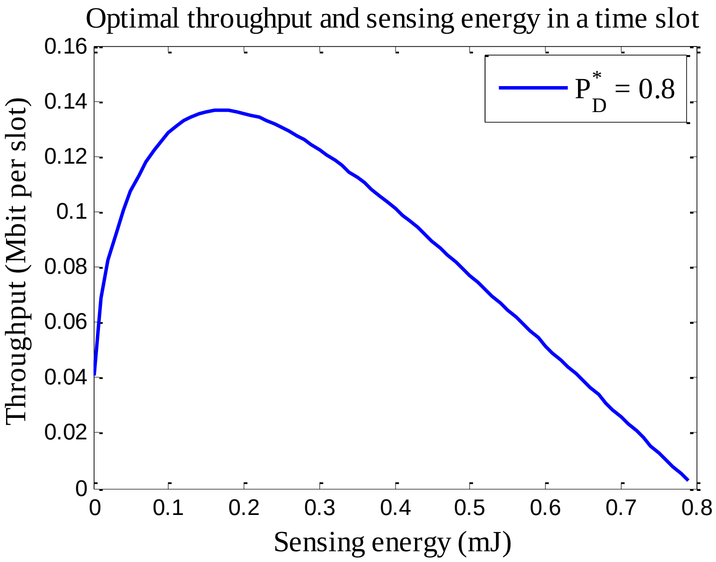

Fig. 2. An example of optimal sensing energy.

3.5 Throughput over primary channel in a time slot

Let W_{pc} and SNR_{pc} denote the bandwidth of the primary channel and the signal-to-noise ratio on the primary channel when the transmission power is P_{pc}, respectively. Accordingly, the throughput that the PSU can attain in a time slot if the PSU wishes to transmit data on the primary channel is expressed as

$$r_{pc}\left(e_{pc}\right) = W_{pc}\, T \log_2\left(1+SNR_{pc}\right) \tag{9}$$

3.6 Action set for PSU

Due to the capabilities to access the harvested channel and the primary channel, at the beginning of each time slot, the PSU will determine an action for throughput maximization. Let a_{hc} and a_{pc} denote the actions for access to the harvested channel and the primary channel, respectively. The action set that includes four possible actions can be described as

$$a_k = \left\{a_{pc}, a_{hc}\right\} \in \mathrm{A} = \left\{a^1, a^2, a^3, a^4\right\} \tag{10}$$

where $a^1 = \{0,0\}$, $a^2 = \{0,1\}$, $a^3 = \{1,0\}$, and $a^4 = \{1,1\}$, in which “zero” indicates that the PSU does not access the channel, and “one” indicates that the PSU accesses the channel. Channel access policy for PSU is to determine the optimal action to maximize the reward, which will be discussed in the next section.

3.7 Energy harvesting model

Powering mobile networks with renewable energy sources has been considered in order to reduce energy costs or the CO_2 emissions that harm our environment [4]. The energy generated from many kinds of renewable energy sources, such as solar, wind, vibration, and even RF signals, is harvested and stored in the power supply.

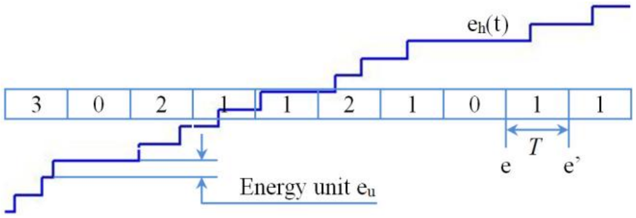

Fig. 3. Energy harvesting

It can be roughly assumed that the harvested arrival energy is packed into energy units, and each energy unit is denoted by e_u. Under this assumption, the energy harvesting can be modeled as a homogeneous Poisson process with a rate parameter λ, the expected number of harvested arrival energy units per unit of time, as shown in **Fig. 3**. Let N denote the number of arrival energy units in a particular time slot t with duration T, and the probability of u units of arrival energy, where $u = 0,1,\ldots$, which is denoted by P_u, is given as

$$P_u = \Pr\{N = u\} = e^{-\lambda T}\frac{(\lambda T)^u}{u!} \tag{11}$$

Assume that a rechargeable battery can contain at most B energy units ($1 \le B$) and can supply the energy for all operations of the PSU. Note that the goal is energy management in terms of spectrum sensing and data transmission so as to maximize PSU throughput. The energy consumed by the unmentioned operations of the PSU is negligible. In particular, the energy stored in the battery can be quantized into different states where each energy state is the amount of energy required to support a particular operation of the PSU. Let e_s, e_{pc} and e_{hc} represent the energy spent for spectrum sensing, transmissions on the primary channel, and transmissions on the harvested channel in a time slot, respectively. Thus, the amount of energy required for both spectrum sensing and harvested channel access is $e_{sa} = e_s + e_{hc}$.

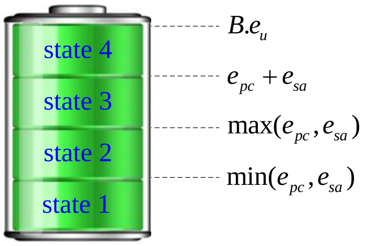

Fig. 4. Energy states of a battery.

Fig. 4 shows a rechargeable battery with 4 states, where battery state 1 is the lowest state and battery state 4 is the highest battery state. **Table 1** details energy ranges of the battery states, boundary among battery states, and possible actions of the PSU in a time slot.

Table 1. Battery state configuration

Battery State	**Energy at the beginning of time slot, e**	**Possible Actions of the PSU**
State 4	$e_{pc} + e_{sa} \le e \le B.e_u$	The PSU may transmit its data on one of two channels or transmit on both channels simultaneously.
State 3	$max(e_{pc}, e_{sa}) \le e < e_{pc} + e_{sa}$	The PSU may transmit its data on one of two channels.
State 2	$min(e_{pc}, e_{sa}) \le e < max(e_{pc}, e_{sa})$	The PSU may transmit its data on the harveted channel or the primary channel, depending on e_{pc} and e_{sa}. Note that e_{sa} includes energy spent for spectrum sensing and data transmission. ▪ If $e_{pc} > e_{sa}$, the PSU may perform spectrum sensing and transmit its data on its harvested channel but does not transmit the data on the primary channel. ▪ If $e_{pc} < e_{sa}$, the PSU may transmit its data on the primary channel but does not perform spectrum sensing and transmit the data on the harvested channel.
State 1	$e < min(e_{pc}, e_{sa})$	The PSU does not transmit its data on any channel due to inefficient energy.

4. Channel access policy based on POMDP

In this section, we will investigate the channel access policy for a PSU with consideration of the energy harvesting. At the beginning of time slot k, based on all of the information including available energy e_k, the belief, the probability of the arrival energy and the probabilities of the energy state from e_k to e_{k+1}, the PSU should select optimal action from action set to maximize its reward function. In the paper, we set reward function as expected long-term throughput, and utilize partially observable Markov decision process (POMDP) to formulate the channel access policy for the PSU. Before formulating the reward function, we first describe the belief function, and observation.

4.1 Belief function

As mentioned in [6], it is well known that for each POMDP, all information that is useful for making decisions can be encapsulated in the belief vector.

For our POMDP framework, based on the actions and access observations on the harvested channel in the past, at the beginning of each time slot, the PSU can know the probability that the harvested channel is in state S to serve the harvested channel access in the time slot. Let p denote this probability, which is called the "belief function". At the end of each time slot, after performing its operation for the action, the PSU updates its belief function according to following observations.

4.2 Observation

4.2.1 Case 1

In this case, the PSU does not have enough energy to transmit data on the harvested channels in the time slot k $(e_k < e_s + e_{hc})$, or it wants to conserve energy to get more throughput in next time slots even though the PSU has enough energy for spectrum sensing and transmission on the harvested channel $(e_k \geq e_s + e_{hc})$. Therefore, the PSU decides to come to an *idle state* such that very low energy $(e_I \approx 0)$ is consumed. Since no spectrum sensing is performed, the belief of the PSU is updated regardless of the state of the harvested channel by using the Markovian property as follows:

$$p_S^u = p\, p_{SS} + (1-p)\, p_{AS} \tag{12}$$

In this case, the PSU does not observe the harvested channel.

4.2.2 Case 2

In the case, the PSU has sufficient energy for data transmission on the harvested channel $(e_k \geq e_s + e_{hc})$ and it wants to access the harvested channel. Before transmitting data on the harvested channel, the PSU will perform spectrum sensing. Therefore, the PSU has three *observations*, based on the spectrum sensing and access to the harvested channel.

a. Observation 1 (Φ_1)

The PSU cannot transmit data on the harvested channel because the harvested channel is now detected as "active" by the energy detector. Since the energy detector has the designed probability of detection P_D^* and corresponding probability of false alarm P_F, the probability of observation 1 is given by

$$\Pr\left(\Phi_1\right) = pP_F + (1-p)P_D^* \tag{13}$$

In this observation, the belief of the PSU that the harvested channel is in state S in the next time slot is updated as follows

$$p_{SS}^u(\Phi_1) = \frac{pP_F\, p_{SS} + (1-p)P_D^* p_{AS}}{pP_F + (1-p)P_D^*} \tag{14}$$

b. Observation 2 (Φ_2)

In this case, the PSU detected the harvested channel as being “silent”, and it transmitted data and successfully received an ACK from its receiver side.

The probability of observation 2 is calculated as

$$\Pr(\Phi_2) = p\left[1 - P_F(e_s)\right] \tag{15}$$

In this case, the PSU confirms that the harvested channel was actually free during the time slot due to ACK from its receiver side. Therefore, the belief of the PSU that the harvested channel is in state S in the next time slot can be updated as

$$p_{SS}^u(\Phi_2) = p_{SS} \tag{16}$$

c. Observation 3 (Φ_3)

In this case, the PSU decided that the harvested channel was “silent” by spectrum sensing and transmitted data to its receiver, but no ACK was received in the rest of the time slot. So, the PSU presumes that the spectrum sensing decision is wrong, and the harvested channel was actually “active”.

The probability of observation 3 is given by

$$\Pr(\Phi_3) = \left(1-p\right)\left(1-P_D^*\right) \tag{17}$$

Consequently, in this case, the PSU updates its belief as follows:

$$p_{SS}^u(\Phi_3) = p_{AS} \tag{18}$$

In addition, a collision between the PSU and other secondary users may be happened during transmission time and the PSU will not receive an ACK from its receiver. In this case, the PSU presumes that a collision with the primary user occurred due to a wrong decision of the spectrum sensing and the PSU also updates its belief as (18).

4.3 Reward function

In the paper, we formulate the reward function, which is defined as the expected total throughput in long-term consideration from the current time slot $k\ (n \geq k)$, as follows: [2]

$$R_k(e,p) = \max_{\left\{a_n^m\right\}_{n=k}^{\infty}} E\left\{\sum_{n=k}^{\infty} \alpha^{n-k} r\left(e_n, p_n, a_n^m\right) \mid \mathrm{e_k} = \mathrm{e},\, p_k = p, m = 1,\ldots,4\right\} \tag{19}$$

where α is the discount factor $(0 \leq \alpha < 1)$. e_n and p_n are the available energy and the belief of the PSU at the beginning of the *n*-th time slot, respectively. $r(e_n, p_n, a_n^m)$ is the reward obtained on both channels at the *n*-th time slot as the PSU takes an action a_n^m, where $n \geq k$ and $m = 1,...,4$.

Consequently, the optimal action can be found and be expressed as follows

$$a_k^* = \underset{\{a_n^m\}_{n=k}^{\infty}}{\arg\max}\, E\left\{\sum_{n=k}^{\infty} \alpha^{n-k} r\left(e_n, p_n, a_n^m\right) | \, \mathrm{e_k} = \mathrm{e},\ p_k = p, m = 1,...,4\right\} \tag{20}$$

The reward function satisfies the following Bellman equation

$$R_k(e,p) = \max\left\{R_k^1(e,p), R_k^2(e,p), R_k^3(e,p), R_k^4(e,p)\right\} \tag{21}$$

where $R_k^1(e,p)$, $R_k^2(e,p)$, $R_k^3(e,p)$, and $R_k^4(e,p)$ are the expected rewards at time slot k when the PSU takes an action a_k^1, a_k^2, a_k^2, or a_k^4, respectively.

The above equation (21) indicates that the maximum expected reward equals the maximum of the expected rewards yielded by taking all actions from the action set. The expected reward comprises two parts: the immediate reward due to taking an action $a_{n=k}^m$ and the expected future reward due to the sequence of actions $a_{n>k}^m$, respectively.

4.4 Action reward

The PSU can obtain expected reward by taking one of four actions from the action set A as following:

4.4.1 Idling (action a_k^1)

As mentioned previously, the PSU does not access any channels so no immediate reward can be achieved in the current time slot k. The PSU is expecting to harvest more energy, and the expected reward will come in the next time slots $n > k$. When the PSU takes the action a_k^1, the expected reward, denoted by $R_k^1(e,p)$, is shown as follows:

$$R_k^1(e,p) = 0 + \alpha \sum_{e_{k+1}} \Pr\left\{b\left(e_k\right) \to b\left(e_{k+1}\right)\right\} R_{k+1}\left(e_{k+1}, p_s^u\right) \tag{22}$$

where $b(e_k)$ and $b(e_{k+1})$ are the battery states corresponding to the available energy e_k at time slot k and to the energy e_{k+1} at time slot $k+1$, respectively. $\Pr\left\{b\left(e_k\right) \to b\left(e_{k+1}\right)\right\}$ is the transition probability of the battery state during the time slot corresponding to e_k and e_{k+1}.

4.4.2 Sensing and accessing the harvested channel and no data transmission on the primary channel (action a_k^2)

In this case, the expected reward of the action a_k^2, denoted by $R_k^2(e,p)$, is shown as follows

$$R_k^2(e,p) = p.r_{hc}\left(e_{sa}\right) + \alpha\sum_{e_{k+1}}\left\{\Pr\left\{b\left(e_k\right)\rightarrow b\left(e_{k+1}\right)\right\}\sum_{\Phi\in\left\{\Phi_1,\Phi_2,\Phi_3\right\}}\Pr\left(\Phi\right)R_{k+1}\left(e_{k+1}, p_{SS}^u(\Phi)\right)\right\} \tag{23}$$

where $r_{hc}(e_{sa})$ is the expected throughput due to the transmission on the harvested channel and can be obtained using (8).

4.4.3 No sensing the harvested channel and transmitting data on the primary channel (action a_k^3)

In this case, the expected reward of the action a_k^3, denoted by $R_k^3(e,p)$, is shown as follows:

$$R_k^3(e,p) = X_k.r_{pc}\left(e_{pc}\right) + \alpha\sum_{e_{k+1}}\Pr\left\{b\left(e_k\right)\rightarrow b\left(e_{k+1}\right)\right\}R_{k+1}\left(e_{k+1}, p_s^u\right) \tag{24}$$

where $r_{pc}(e_{pc})$ is the immediate reward coming from the transmission on the primary channel, and can be obtained by using (9). X_k is the policy function of the PSU slot assignment from the primary network operator.

4.4.4 Sensing and accessing the harvested channel and transmitting data on the primary channel (action a_k^4)

In this case, as mentioned previously, since there is sufficient energy inside the battery, the PSU attains the maximum immediate reward because it is able to transmit data on both channels. The expected reward of the action a_k^4, denoted by $R_k^4(e,p)$, can be expressed as follows:

$$\begin{aligned} R_k^4(e,p) = {} & p.r_{hc}\left(e_{sa}\right) + X_k.r_{pc}(e_{pc}) + \\ & \alpha\sum_{e_{k+1}}\left\{\Pr\left\{b\left(e_k\right)\rightarrow b\left(e_{k+1}\right)\right\}\sum_{\Phi\in\left\{\Phi_1,\Phi_2,\Phi_3\right\}}\Pr\left(\Phi\right)R_{k+1}\left(\mathrm{e}_{\mathrm{k+1}}, p_{SS}^u(\Phi)\right)\right\} \end{aligned} \tag{25}$$

It is clear that due to the long-term throughput calculation, the POMDP scheme computation is more complex than the Myopic scheme computation. In this paper, the recursive method is used to solve POMDP. This method searches all space of a policy for the best course of action. Therefore, to determine an action for the PSU in a time slot, the PSU needs to calculate the values of $R_k^1(e,p)$, $R_k^2(e,p)$, $R_k^3(e,p)$ and $R_k^4(e,p)$ in equations (22)-(25). For complexity calculation, let's assume that there are M cases of arrival energy units in a time slot, and we consider V future time slots for the long-term throughput calculation. For the Myopic scheme, as above mentioned, only the immediate throughput in the current time slot is computed. In this scheme, four of the first terms in the equations (22)-(25) are calculated. For the POMDP scheme, the PSU needs to calculate the second terms in (22)-(25). For example, if only one future time slot is considered ($V = 1$), M cases of $\mathrm{R}_{\mathrm{k+1}}$ in the equations (22) and (24) and $3M$ cases of $\mathrm{R}_{\mathrm{k+1}}$ in the equations (23) and (25) are computed.

Consequently, *8M* cases of R_{k+1} need to be calculated. Similarly, if two future time slots are considered (V = 2), *8M* cases of R_{k+2} are calculated for each case of R_{k+1}. Therefore, the computational complexity grows with $(8M)^V$ in making the action decision for the PSU in a time slot.

5. Simulation Results

In this section, we present simulation results that illustrate the performance of the POMDP scheme and the Myopic scheme. Here, Myopic scheme is considered for performance comparison. Myopic scheme is a special case of POMDP scheme because it only seeks the action a_k^* based on the immediate reward at the current time slot $(n=k)$ without considering the future expected reward $(n>k)$. We can obtain Myopic scheme by setting the parameter α to zero $(\alpha=0)$ in (22) – (25). We observe throughput of the PSU in different settings. Simulation parameters are summarized in **Table 2**.

Table 2. Simulation parameters

Symbol	Description	Value/metric
T	Slot time	20 ms
f_s	Sampling frequency for sensing	800 kHz
p_{SS}	Transition probability from state S to state S	0.7
p_{AS}	Transition probability from state A to state S	0.2
α	Discount factor	0.9
β	SNR in energy detection	-15 dB
SNR_{hc}	SNR of the link on harvested channel	10 dB
SNR_{pc}	SNR of the link on primary channel	08 dB
W_{hc}	Harvested channel bandwidth	3.0 MHz
W_{pc}	Primary channel bandwidth	1.0 MHz
P_{hc}	Transmission power on harvested channel	100 mW
P_{pc}	Transmission power on primary channel	80 mW
P_{sl}	Power in Idle state	05 mW
P_s	Sensing power	40 mW
B	Maximum energy stored in battery	20 mJ
e_a	Available energy in the battery	10 mJ
e_u	Energy unit	1 mJ
p_a	Probability that the primary channel will be assigned to the considered PSU in a slot	1

In the first experiment, we first set the energy arrival rate as 30 (λ = 30 e_u/s) and then evaluate the throughput of the PSU as the designed probability of detection P_D^* varies from 0.8 to 1. **Fig. 5** shows the number of bits that PSU transmits in one second for two schemes. It can be observed that the throughput of POMDP scheme is higher than that of the Myopic scheme when $P_D^* < 1$. For the case of $P_D^* = 1$, however two schemes get the same throughput. It is mainly due to the fact that the energy detector spends almost whole duration τ $(\tau \approx T)$ for spectrum sensing in both schemes so as to achieve an absolutely exact detection of the presence of a primary user on the harvested channel. As a result, there is almost no chance for data transmission on the harvested channel, and data is only transmitted on the primary channel. Accordingly, two schemes obtain the same throughput.

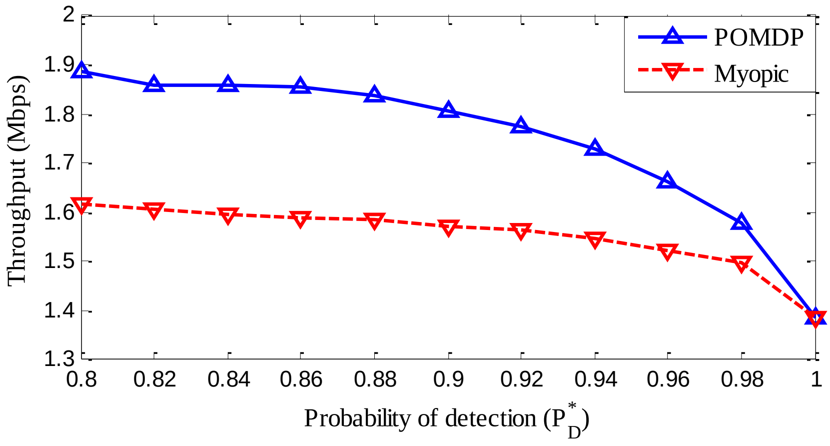

Fig. 5. Throughput of the primary-secondary user for different values of the designed probability of detection when λ = 30 e_u/s.

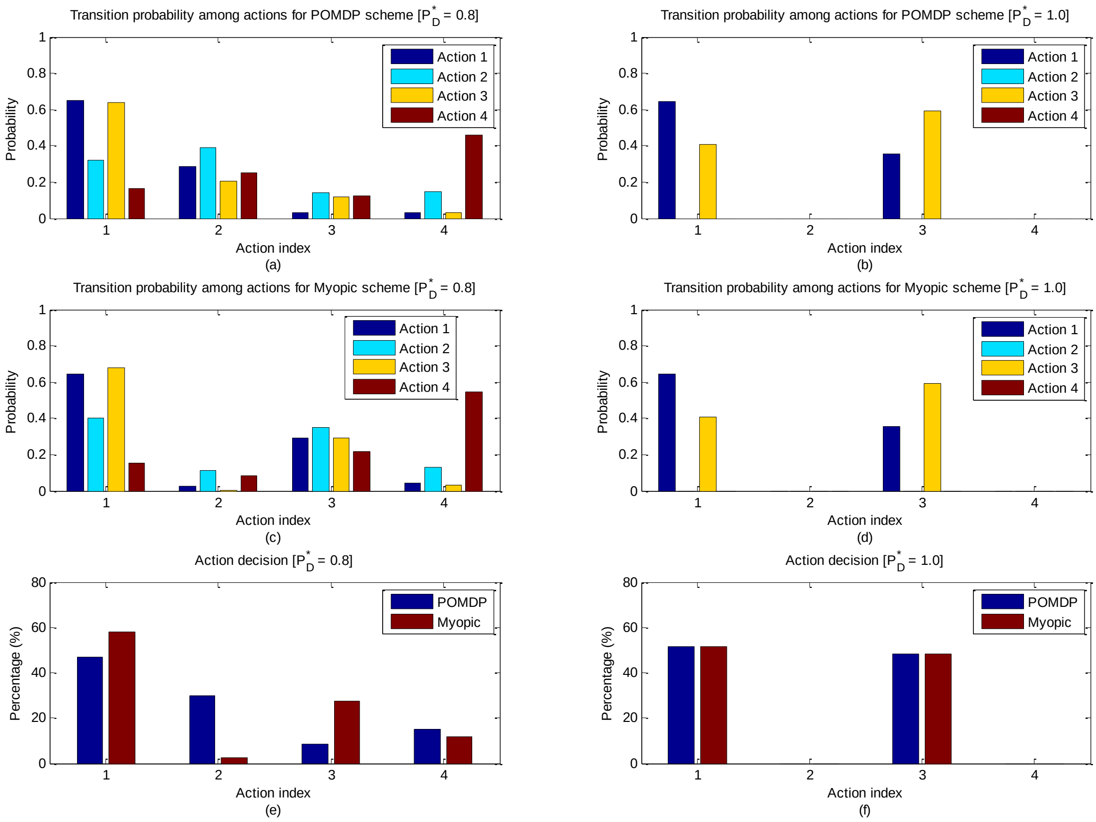

Fig. 6. Transition probability among actions and action decision of the PSU with $P_D^* = 0.8$ and $P_D^* = 1$ when λ = 30 e_u/s.

Fig. 6 shows the probability of transition among actions and the percentage of four actions taken by the considered PSU in the first simulation. For the case of $P_D^* = 0.8$, the transition probabilities among actions are absolutely positive in both of the schemes. This means that, after performing an action, and based on available energy, the PSU can take one of four actions.

However, the transition probability from action a^1 to action a^4 is small in both of the schemes because the amount of harvested energy during the implementation of action a^1 is less sufficient for taking action a^4. It can be further seen that, the transition probability from any action to action a^2 is higher than the one to action a^3 in POMDP scheme and vice versa in Myopic scheme. This explains why the action a^2 in POMDP scheme is more taken than in Myopic scheme as shown in **Fig. 6e**. **Fig. 6e** also shows that all types of actions have been taken by the PSU in both of the schemes and that, in POMDP scheme, the PSU took more two actions a^2 and a^4 and less action a^1, compared to the case of Myopic scheme, which means that harvested channel is more utilized in POMDP than in Myopic. These statistical results can also explain why the PSU attains higher throughput in the POMDP scheme than in the Myopic scheme. For the case of $P_D^* = 1$, only transition probabilities between action a^1 and action a^3 are positive and equal as shown in **Fig. 6b** and **Fig. 6d**. Therefore, the percentage of two actions a^1 and a^3 is equal in both of the schemes. By noting that the action a^3 corresponds to the case in which the PSU transmits its data on the primary channel, we know that the PSU did not take any action to access the harvested channel and did not use the cognitive radio function. The PSU acts like a traditional primary user when $P_D^* = 1$.

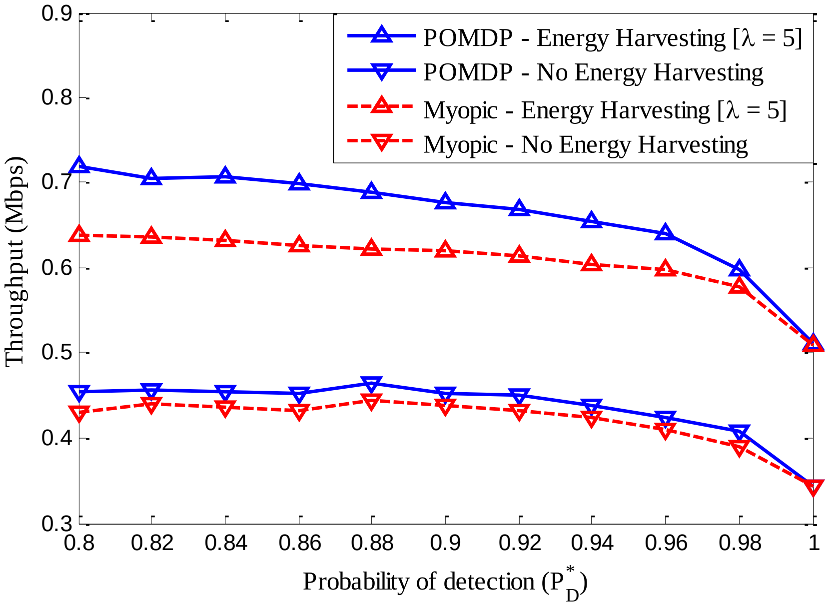

Fig. 7. Throughput of the PSU in two modes: energy harvesting and no energy harvesting, with different values of the designed probability of detection.

In the second experiment, we observe the throughput of the PSU when the PSU operates in two modes: the energy harvesting and no energy harvesting. We first set the energy arrival rates as 5 and 0 (e_u/s) for energy harvesting mode and no energy harvesting mode, respectively, and then examine the throughput of the PSU as the designed probability of detection P_D^* varies from 0.8 to 1. **Fig. 7** illustrates the throughput of the PSU according to the variation of the probability of detection. It is obvious that the achievable throughput in the mode of no energy harvesting is less than the one in the mode of energy harvesting in both POMDP and Myopic schemes when $P_D^* < 1$.

In the third experiment, we observe the PSU throughput according to the variation of lambda, from 0 to 20 (e_u/s), in different designed values of probability of detection as 0.84, 0.9, and 0.96. From the **Fig. 8,** it can be observed that, for a fixed P_D^*, the throughput of the PSU linearly increases with the increase of the energy arrival rate in both of the schemes. It can be further seen that the throughput gap between POMDP scheme and Myopic scheme becomes larger as the energy arrival rate increases. Furthermore, the PSU throughput decreases with increase of detection probability value for a fixed λ. This result is consistent with the result of the second experiment.

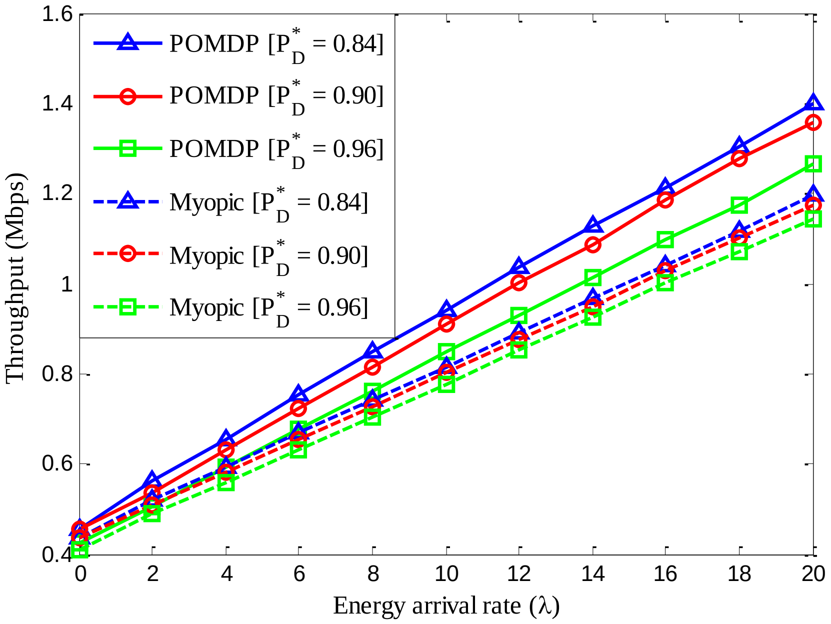

Fig. 8. The PSU throughput with the variation of the energy arrival rate and the probability of detection.

In the fourth simulation, we keep the transmission power of the PSU on both of the channels and vary the bandwidth of the harvested. From the **Fig. 9**, it can be easily seen that, in the case of $P_D^* = 1$, the PSU only takes the action to access the primary channel demonstrated by a similar amount of throughput in all cases of the bandwidth. It also shows that the more bandwidth the harvested channel has, the more throughput the PSU obtains in all case of $P_D^* < 1$ in two schemes and that the throughput obtained in POMDP scheme is always larger than the one obtained in Myopic scheme.

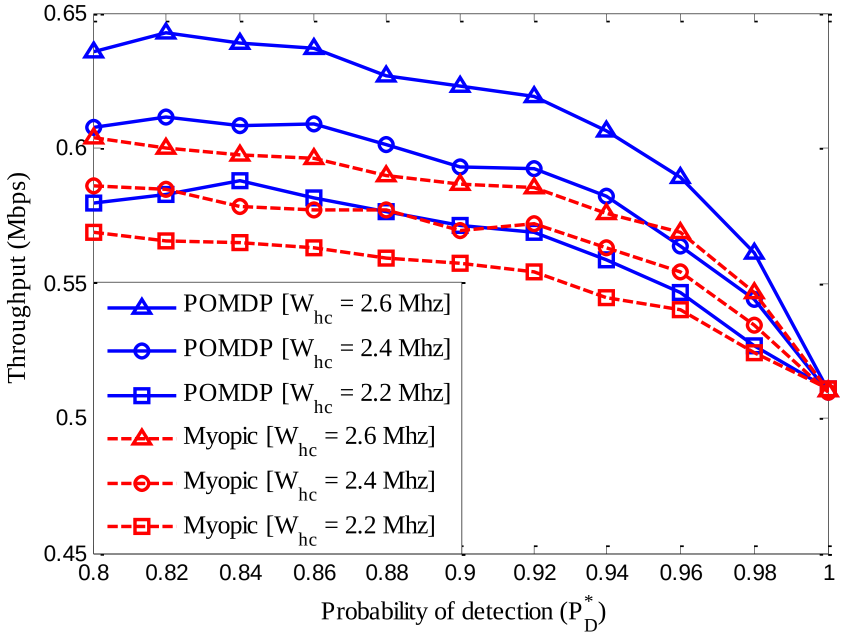

Fig. 9. The PSU throughput according to the variation of the bandwidth of the harvested channel (W_{hc}).

In the next simulation, we examine the throughput of the PSU in different cases of the SNR at receiver on the harvested channel with the variation of probability of detection when the energy arrival rate λ is fixed at 5 (e_u/s). **Fig. 10** shows that the SNR at the receiver also impacts on the PSU throughput. A harvested channel with higher SNR at the receiver is preferred.

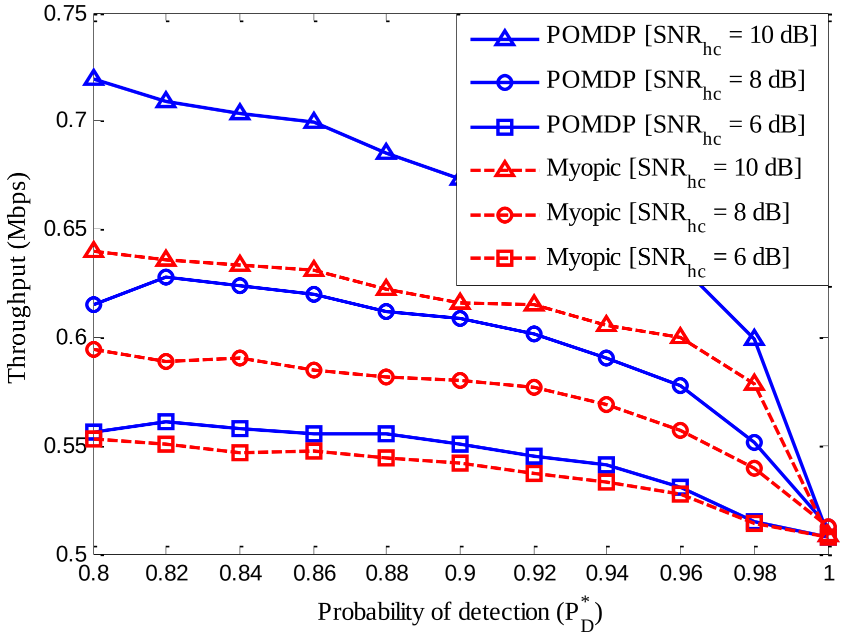

Fig. 10. The PSU throughput according to the variation of the SNR_{hc} when λ = 5 (e_u/s).

In the last experiment, the PSU will transmit data over different harvested channels with different transmission powers but other simulation parameters are preserved for all channels. The simulation result is shown in **Fig. 11**. It can be easily seen that, for each case of $P_D^* < 1$, POMDP scheme outperforms the Myopic scheme in terms of throughput. A similar result can be found in the case of $P_D^* = 1$, the PSU obtains a similar amount of throughput. The figure shows that if the transmission power of the PSU is high then the PSU will obtain less throughput because much energy is consumed for the transmission on the harvested channel.

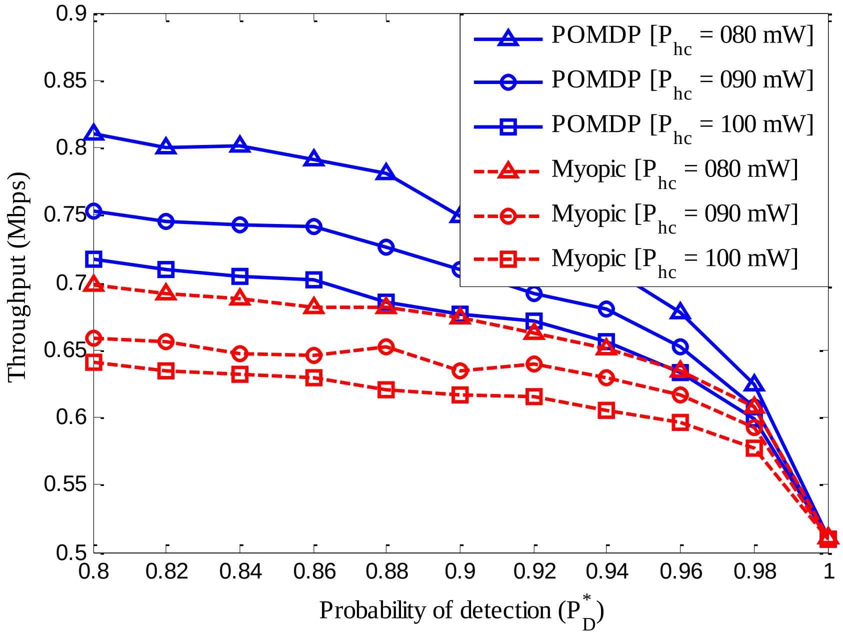

Fig. 11. The PSU throughput according to the variation of transmission power on the harvested channel when $\lambda = 5$ (e_u/s).

6. Conclusion

In this paper, a PSU equipped with an energy harvester has been investigated in terms of throughput and channel access policy. The PSU has the highest priority in using its primary channel and can access the harvested channel using the cognitive radio function when more throughput is needed. The channel access policy of PSU is formulated in form of POMDP framework in order to select optimal action in the current time slot to maximize the expected long-term throughput. Furthermore, the Myopic scheme, which uses only the immediate reward in making decisions without taking the future reward into account, was considered for performance comparison. The simulation results show that the POMDP scheme outperforms the Myopic scheme in terms of throughput, but it requires more computations to make an action decision regarding channel access. Through this study, we demonstrate that a primary user can obtain more throughput when it is equipped with cognitive-enable radio. It is also noted that due to the future reward consideration to find the optimal action a_k^* at the time slot k, the POMDP scheme requires much more complexity than the Myopic scheme, not only in terms of the computation to make the action decision, but also in updating all of the information of the PSU.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MEST) (No. NRF-2013R1A2A2A05004535)

References

[1] E. Z. Tragos, S. Zeadally, A. G. Fragkiadakis, and V. A. Siris, "Spectrum assignment in cognitive radio networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1108-1135, July, 2013. Article (CrossRef Link)

[2] A. Sultan, "Sensing and transmit energy optimization for an energy harvesting cognitive radio," *IEEE Wireless Communications Letters*, vol. 1, no. 5, pp. 500-503, October, 2012. Article (CrossRef Link)

[3] S. Park, S. Lee, B. Kim, D. Hong, and J. Lee, "Energy-efficient opportunistic spectrum access in cognitive radio networks with energy harvesting," in *Proc. of the 4th International Conference on Cognitive Radio and Advanced Spectrum Management*, pp. 62:1–62:5, 2011. Article (CrossRef Link)

[4] S. Park, H. Kim, and D. Hong, "Cognitive radio networks with energy harvesting," *IEEE Transactions on Wireless Communications*, vol. 12, no. 3, pp. 1386-1397, March, 2013. Article (CrossRef Link)

[5] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1326-1336, April, 2010. Article (CrossRef Link)

[6] A. T. Hoang, Y. C. Liang, D. T. C. Wong, Y. Zeng, and R. Zhang, "Opportunistic spectrum access for energy-constraint cognitive radios," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1206-1211, March, 2009. Article (CrossRef Link)

[7] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 589-600, April, 2007. Article (CrossRef Link)

[8] Y. C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326-1337, April, 2008. Article (CrossRef Link)

Thanh-Tung Nguyen received the B.E. degree in electronics & telecommunications engineering from Can Tho University, Vietnam in 2004. Since 2011 he has been involved in the combined degree programs (Master's and PhD.) in University of Ulsan, Korea. His current research interests include Resource Allocation for Cognitive Radio Networks.

Insoo Koo received the B.E. degree from the Kon-Kuk University, Seoul, Korea, in 1996, and received the M.S. and Ph.D. degrees from the Gwangju Institute of Science and Technology (GIST), Gwangju, Korea, in 1998 and 2002, respectively. From 2002 to 2004, he was with Ultrafast Fiber-Optic Networks (UFON) research center in GIST, as a research professor. For one year from September 2003, he was a visiting scholar at Royal Institute of Science and Technology, Sweden. In 2005, he joined University of Ulsan where he is now a full professor. His research interests include Next Generation Wireless Communication Systems and Wireless Sensor Networks.