POMDP Based Throughput Maximization Scheme For Wireless Powered Cognitive Radio Network

Hiep Vu-Van and Insoo Koo
School of Electrical Engineering
University of Ulsan, Ulsan, South Korea
Email: vvhiep@gmail.com and iskoo@ulsan.ac.kr

***Abstract*—In wireless powered Cognitive Radio Network (CRN), a Cognitive radio User (CU) can receive the wireless power transferred from an energy source for its operation. In this paper, we consider the operations of a CU in the CRN with wireless powered transfer. The operations of the CU include *three* modes in terms of *Data Transmission* (DT), *Receive Wireless Power* (WP) and *Sleep* (SL). In order to consider the effect of the future operation on the current reward of the CU, we formulate the problem to choose an optimal operation mode of the CU as a Partially Observable Markov Decision Process (POMDP), in which the *two* states in terms of belief about the channel state and remaining energy are utilized as main factors to determine an optimal operation of the CU. Simulation results show the efficiency of the proposed scheme.**

***Index Terms*—cognitive radio; wireless powered transfer; operation mode switching; maximizing throughput; POMDP.**

I. INTRODUCTION

Cognitive Radio (CR) technology can improve spectrum utilization by allowing Cognitive radio Users (CU) to share the frequency assigned to a licensed user, called the Primary User (PU). In order to avoid interference with the operation of the licensed user, the CU is allowed to be active only when the frequency is free. Otherwise, when the presence of the PU is detected, the CU has to vacate their occupied frequency.

In the CRN, the CU often has a small battery that can maintain operations of the CU in a short time. The wireless power source [1][2] can charge the battery to extend its lifetime. However, the wireless power is still limited, therefore, the performance of the CRN strongly depends on how effectively the CU uses its power. The problem of optimal energy management has been considered previously in [3][4] where an optimal energy management scheme for a sensor node with an energy harvester to maximize throughput is proposed. In [5][6], a scheme to find an optimal action policy including sleeping to save energy or being active to take an opportunity of transmitting data is proposed. In [5], the CU has an independent energy harvester that harvests non-radio frequency (RF) energy from surround environment (e.g., wind, solar etc.), hence, the CU can perform data transmission and energy harvesting concurrently. On the other hand, [5] adopts an RF energy harvester that harvests energy from a PU's signal in the closed environment. Subsequently, the energy harvesting performance in [5][6] significantly fluctuates due to the uncontrollable wireless energy source.

In this paper, we consider the operation of wireless powered cognitive radio network that includes a control center, a wireless power source and cognitive radio user. The control center has a strong power supply (i.e., it always has enough energy to take any operation) and controls the operation of the whole network. Because the wireless power source can be controlled by the control center, the energy harvester of the CU is more efficient than conventional where the harvester receives energy from an unknown source. In the CRN, the operations of the CU includes *three* modes in terms of *Data Transmission* (DT), *Receive Wireless Power* (WP) and *Sleep* (SL). Due to the limitation of the hardware, the CU has a limited capacity battery that can be recharged by a wireless power source. On the other hands, the CU cannot perform multiples actions at the same time (i.e., it cannot perform data transmission and receiving energy at the same time).

This paper will propose a scheme to determine an optimal operation for the CU that can maximize the average achieved throughput of the CRN. In order to take a long-term optimization for the operation selection, we formulate the problem to choose an optimal operation as a partially observable Markov decision process (POMDP) in which the CU's *two* states in terms of the belief about the absence probability of the primary user (PU) and the remaining energy are utilized as main factors to decide an optimal operation of the CU in current time slot. By applying POMDP, the effects of future operations on the current achieved throughput are also considered. It is expected that the proposed scheme based on POMDP theory will provide the CR system an improved performance.

The remaining part of this paper is organized as follows. Section 2 shows the system model of the considered wireless powered cognitive radio network. Section 3 formulates an optimization problem based on POMDP in order to find the optimal operation mode for maximizing throughput of the CRN. Section 4 introduces simulation model and simulation results of the proposed and reference schemes. Finally, Section 5 concludes this paper.

II. SYSTEM MODEL

A wireless powered Cognitive Radio Network (CRN) shown in figure 1, that includes a control center, a wireless energy source and a CU, works in a cognitive manner, where the CRN is allowed to opportunely utilize the channel that is assigned to a primary user (PU). The operation of the PU changes between *two* states of the Markov chain model, that is, Presence (P) and Absence (A) as shown in figure 2. The transition probability of the PU from state P to state A, state A to state P and from

 ISBN: 978-1-61208-633-0

state A or state P to itself are defined as P_{PA}, P_{AP}, P_{AA} and P_{PP}, respectively.

The control center of the CRN controls the operations of CU that include *Sleep* (SL), *Receive Wireless Power* (WP) and *Data Transmission* (DT) modes. In *Sleep* mode, the control center does not request the CU to do any action, so the CU is in the sleep state. *Receive Wireless Power* mode firstly requests the wireless energy source to transfer energy to the CU, after that the energy source will wirelessly transfer energy to the CU where the energy will be received and stored in a limited capacity battery of the CU to use for future operation. In the *Data Transmission* mode, the control center firstly performs spectrum sensing to determine the state of the PU. If the sensing result is *Presence*, the control center does not ask the CU to do any action. Otherwise, if the sensing result is *Absence*, the control center asks the CU to transmit data. If the transmission successes, the CRN can achieve a throughput as:

$$R = \frac{T - t_s}{T} C_0 \tag{1}$$

where T is the duration of a time slot, t_s is the sensing time and C_0 is the throughput of the communication link, which is defined as $C_0 = \log_2(1 + SNR_{CR})$, where SNR_{CR} is the signal to noise ratio (SNR) received in the control center.

In the *Receive Wireless Power* mode, the CU receives the energy from the energy source. Let define $E_h(t)$ as the number of units of energy that the CU receives at the time slot t. Due to the fluctuation of wireless power transfer, the amount of energy that the CU received each time slot is not stable. Hence, we assume that $E_h(t)$ takes value from a finite number ξ_h of energy units.

$$E_h(t) \in \Upsilon_h = \left\{ e_1^h, e_2^h, ..., e_{\xi_h}^h \right\}, \tag{2}$$

where $0 \leq e_1^h < e_2^h < ... < e_{\xi_h}^h \leq E_{ca}$, and E_{ca} (units of energy) is the capacity of the battery of the CU.

The probability mass function (PMF) of the harvested energy is given as follows:

$$P_{E_h}(k) = \Pr\left[E_h(t) = e_k^h \right], \ k = 1, 2, ..., \xi_h. \tag{3}$$

We assume that the harvested energy follows the stochastic process that are marked by Poisson process. Subsequently,

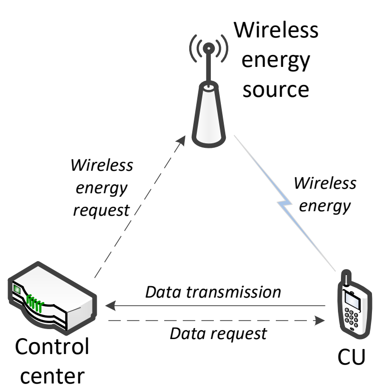

Figure 1. System model of the network.

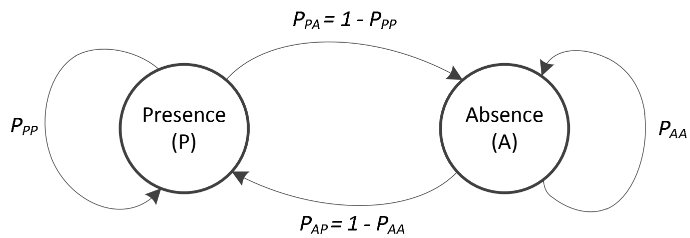

Figure 2. Markov chain states of the PU.

$E_h(t)$ is a Possion random variable with mean e_m^h and the PMF in (3) can be rewritten as follows:

$$P_{E_h}(k) \approx \frac{e^{-e_m^h} \left(e_m^h\right)^k}{k!}, \ k = 1, 2, ..., \xi_h. \tag{4}$$

III. Throughput maximization scheme based on POMDP for Wireless Powered Cognitive Radio Network

In this section, in order to maximize its achieved throughput of the CRN, the problem of switching operation will be formulated as POMDP framework. As mentioned in section 2, we consider three operation modes of the CRN including *Sleep (SL)*, *Receive Wireless Power (WP)* and *Data Transmission(DT)* modes .

Hence, in order to formulate a POMDP, we define the operation space as: $\Psi = \{SL,\ WP,\ DT\}$. In addition, the state space of the system is defined as $S(t) = \{e_r(t), p_0(t)\}$, where $e_r(t)$ is the remaining energy of the CU and $p_0(t)$ is the belief about absence probability of the PU signal.

The reward of CRN is defines as its achieved throughput R according to the chosen operation. For POMDP framework, we define the value function $V(e_r(t), p_0(t))$ as the maximum total discounted throughput from the current time slot when the current state of the CU is $S(k)$. Subsequently, we have,

$$V(S(k)) = \max_{a(k)} \left\{ \sum_{t=k}^{\infty} \alpha^{t-k} R(S(t), a(t)) \left| S(k) \right. \right\} \tag{5}$$

where $0 \leq \alpha < 1$ is the discount factor, $S(k) = \{e_r(k), p_0(k)\}$. $R(S(t), a(t))$ is the throughput of the CU achieved at the t^{th} time slot, which is mainly dependent on state $S(t)$ and action decision $a(t)$.

A. Sleeping Mode (ϕ_1)

If the CU decides to remain sleeping, no throughput is achieved, then $R(S(t), SL\,|\phi_1) = 0$.

The belief p_0 for the next time slot will be updated as

$$p_0(t+1) = p_0(t)P_{AA} + (1 - p_0(t))\, P_{PA}. \tag{6}$$

In practice, the CU may consume energy even when it is in *Sleep* mode. However, the consumed energy is very small. Therefore, we assume that the CU does not consume energy in the *Sleep* mode. Hence, the remaining energy of the CU keeps the same as previous time slot:

$$e_r(t+1) = e_r(t), \tag{7}$$

 ISBN: 978-1-61208-633-0

with transition probability

$$\Pr\left(e_r(t) \to e_r(t+1)\left|\phi_1\right.\right) = 1. \tag{8}$$

B. Receive Wireless Power mode (ϕ_2)

In this operation mode, the CU does not transmit data. Therefore, no throughput is achieved, $R\left(S(t), WP\left|\phi_2\right.\right) = 0$

This operation mode has no more information to update the belief, p_0, so that the belief p_0 is updated as the *Sleep mode* in (6).

In this operation mode, the CU will receive energy from the wireless power source. So that, the remaining energy of the CU will be increased as,

$$e_r(t+1) = \min\{e_r(t) + e_k^h\left(t\right), E_{ca}\},\ k = 1, 2, ..., \xi_h, \tag{9}$$

with transition probability,

$$\Pr\left(e_r(t) \to e_r(t+1)\left|\phi_1\right.\right) = \Pr\left[E_h\left(t\right) = e_k^h\right]. \tag{10}$$

C. Data Transmission mode

This mode gives the CRN an opportunistic to achieve throughput by performing data transmission. So that, in order to transmit data the remaining energy of the CU must be higher than the energy spending for transmission, that is, $e_r(t) > e_t$.

The achieved throughput of the system depends on their observations. In this paper, we define *three* observations for the *Data Transmission* mode as follows:

Observation 1 *(ϕ_3):* The control center detects that the PU is *present* (the channel is used by the PU - state *P*). Then, the CU is not asked to transmit data and there is no achieved throughput $R\left(S(t), DT\left|\phi_3\right.\right) = 0$. The probability that this observation (ϕ_3) happens is:

$$\Pr\left(\phi_2\right) = p_0(t)P_f + (1 - p_0(t))P_d, \tag{11}$$

where P_f and P_d are spectrum sensing performance of the control center. P_f is the probability of false alarm that event happens when the PU signal is actually *absence*, however the sensing result is given as *presence*. P_d is the probability of detection that event happens when the sensing result is correct as the PU signal is *presence*.

The sensing result can be used to correct the belief p_0 in the current time slot as,

$$p_0^u(t) = \frac{p_0(t)P_f}{p_0(t)P_f + (1 - p_0(t))P_d}. \tag{12}$$

As a results, the updated belief for the next time slot is given by:

$$p_0(t+1) = p_0^u(t)P_{AA} + (1 - p_0^u(t))P_{PA}. \tag{13}$$

Since, this observation of the *Data Transmission* mode does not require the CU to do any action, the energy of the CU is not changed. So that, the remaining energy and transition probability are updated as the *Sleep* mode in (7) and (8).

TABLE I. SIMULATION PARAMETERS

Symbol	Description	Value
SNR_{CR}	SNR of the sensing channel	-10 dB
Pr(H_0)	The absence probability of the PU	0.5
P_{AA}	Transition probability from state A to itself	0.8
P_{PA}	Transition probability from state P to state A	0.2
E_{ca}	Total capacity of battery	15 units of energy
e_m^h	mean of harvested energy	2 units of energy
e_t	Transmission energy	2 units of energy

Observation 2 *(ϕ_4):* The control center does not detect any signal from the PU (i.e., the state *A* of the PU). The CU is requested to transmit its data to the control center through the channel assigned to the PU and the transmission is success (i.e., the CU can receive an ACK message). This means that the sensing results is correct (the PU signal is really absent). The throughput is achieved as:

$$R\left(S(t), DT\left|\phi_4\right.\right) = \frac{T - t_s}{T}C_0. \tag{14}$$

The probability that the observation ϕ_4 happens is:

$$\Pr\left(\phi_3\right) = p_0\left(t\right)\left(1 - P_f\right). \tag{15}$$

The belief and remaining energy for the next time slot can be updated, respectively, as:

$$p_0(t+1) = P_{AA} \tag{16}$$

and

$$e_r(t+1) = e_r(t) - e_t, \tag{17}$$

where e_t is the energy spent for data transmission.

The transition probability of the updated energy is given as,

$$\Pr\left(e_r(t) \to e_r(t+1)\left|\phi_4\right.\right) = 1. \tag{18}$$

Observation 3 *(ϕ_5):* This observation is similar to the observation ϕ_4, state A of the PU is detected and the CU transmits its data. However, the CU can not receive ACK message. This means that the sensing results is incorrect (the PU signal is *present*), and the transmission data fails, no throughput is achieved, $R\left(S(t), DT\left|\phi_5\right.\right) = 0$. The probability that ϕ_5 is obtained is:

$$\Pr\left(\phi_5\right) = \left(1 - p_0\left(t\right)\right)\left(1 - P_d\right). \tag{19}$$

The belief that the PU is in state A at the next time slot is:

$$p_0(t+1) = P_{PA} \tag{20}$$

The remaining energy and transition probability of the CU for the next time slot can be updated similar to the case of observation ϕ_4 in (17) and (18), respectively.

According to those observations, the expected value function in (5) can be expressed as (21). In order to find an optimal mode policy for maximizing throughput, the optimization problem in (21) will be solved by using the *value iterations* method [7].

$$V\left(S\left(k\right)\right)=\max_{a_k}\left\{\sum_{t=k}^{\infty}\alpha^{t-k}\sum_{\phi_i\in a(t)}\Pr\left(\phi_i\right)\sum_{e_r(t+1)}\Pr\left(e\left(t\right)\to e\left(t+1\right)|\phi_i\right)R\left(S\left(t\right),a\left(t\right)|\phi_i\right)|S\left(k\right)\right\} \tag{21}$$

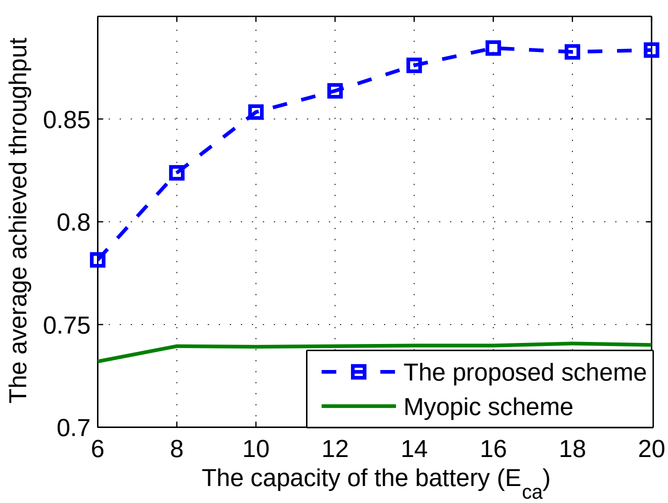

Figure 3. The average achieved throughput of the system versus battery capacity E_{ca} when $e_t = 2$ and $e_m^h = 1$.

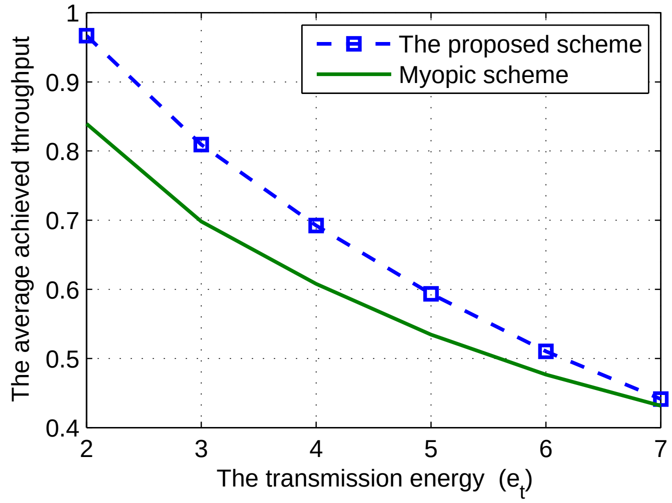

Figure 5. The average achieved throughput of the system versus transmission energy e_t when $e_m^h = 2$ and $E_{ca} = 15$.

IV. SIMULATION RESULTS

In this section, we present simulation results to show the efficiency of the proposed scheme. *Myopic* scheme only considers the current time slot for the *value function* (i.e. $\alpha = 0$) to determine the operation mode. This means that *Myopic* scheme always chooses the third operation (i.e., *Data Transmission*) when the CU has enough energy to perform transmission (i.e., $e_r > e_t$). Otherwise, they will choose the second operation (i.e., *Receive Wireless Power*. Simulation results of *Myopic* will be provided in this section for a reference. The parameters for simulations are shown in Table I.

In order to evaluate the performance of the proposed scheme, the average achieved throughput in each time slot of the system will be utilized.

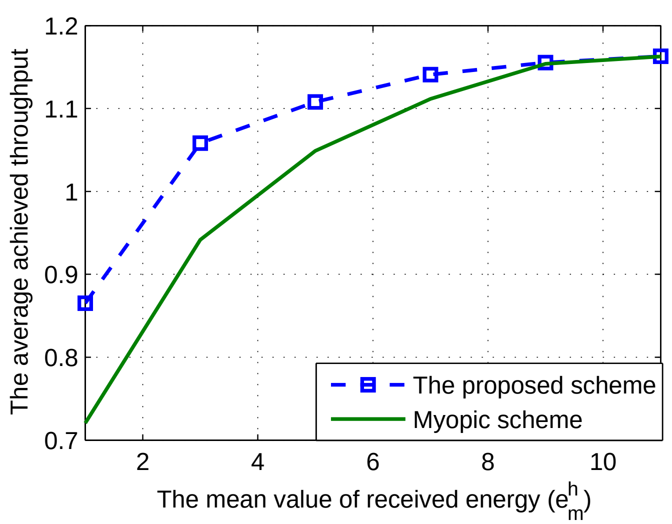

Figure 4. The average achieved throughput of the system versus the mean value of received energy e_m^h when $E_{ca} = 15$ and $e_t = 2$.

Figure 3 shows the average achieved throughput (bits/s/Hz) of the considered schemes according to the capacity of the CU battery E_{ca} (units of energy). The proposed scheme always achieves better performance in comparison with *Myopic* scheme as shown in the figure. It can be seen that the increase of battery capacity may help the CU to achieve higher throughput. However, when the battery is big enough (i.e., it has enough space to store all received energy), the increase of battery will not affect the throughput. However, since *Myopic* scheme always choose *Data Transmission* mode when it has enough energy so that the received energy is often spent before the battery is fully charged. Subsequently, the battery capacity seems not to affect on *Myopic scheme*.

Figure 4 presents the relation between the average achieved throughput and the mean value of the received energy e_m^h (units of energy) of the CU. The higher amount of e_m^h provides more energy for *Data Transmission* mode of the CU so that the CU can get more throughput. However, when e_m^h is high enough for *Data Transmission* mode of the CU in all time, the increase of e_m^h has no more effect on the CU's throughput. In this case, energy constrained will be disappeared, so that the chosen operation of the proposed scheme and *Myopic* scheme will be the same and that is the reason why they have the same performance when the received energy is high enough.

The effect of the transmission energy e_t (units of energy) on the average achieved throughput of the proposed scheme is shown in Figure 5. The higher energy spent for transmission may reduce the throughput of the system. However, the proposed scheme always obtains better performance in comparison with *Myopic scheme*.

Simulations results shown in all figures show that the proposed scheme can offer the CRN a better performance than the conventional *Myopic* scheme. That benefit of the proposed scheme is achieved by considering the future operations on the current achieved throughput.

V. Conclusion

In this paper, we consider the operations of wireless powered cognitive radio network in which the CU is powered by a wireless power source. The proposed scheme based on POMDP will determine an optimal operation of the CU to maximize the average achieved throughput with limited power supply. In the other hand, the proposed scheme can take consideration the effect of the future operation on the current throughput of the system by applying POMDP theory. Simulation results show that the proposed scheme can obtain better performance than conventional *Myopic* scheme.

Acknowledgement

This work was supported by the National Research Foundation (NRF) of Korea funded by the MEST (2017R1D1A1B03029448).

References

[1] Y. Wang, Y. Wang, F. Zhou, Y. Wu, and H. Zhou, "Resource allocation in wireless powered cognitive radio networks based on a practical non-linear energy harvesting model," *IEEE Access*, vol. 5, pp. 17 618–17 626, 2017.

[2] S. A. Mousavifar, Y. Liu, C. Leung, M. Elkashlan, and T. Q. Duong, "Wireless energy harvesting and spectrum sharing in cognitive radio," in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, 2014, pp. 1–5.

[3] 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, 2010.

[4] S. Mao, M. H. Cheung, and V. Wong, "An optimal energy allocation algorithm for energy harvesting wireless sensor networks," in *Communications (ICC), 2012 IEEE International Conference on*, 2012, pp. 265–270.

[5] A. Sultan, "Sensing and transmit energy optimization for an energy harvesting cognitive radio," *Wireless Communications Letters, IEEE*, vol. 1, no. 5, pp. 500–503, 2012.

[6] S. Park et al., "Optimal mode selection for cognitive radio sensor networks with rf energy harvesting," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on*, 2012, pp. 2155–2159.

[7] D. P. Bertsekas, *Dynamic Programming and Optimal Control.* Athena Scientic, 2nd edition, 2001, vol. 1 and 2.