Scientific Journal Impact Factor (SJIF): 1.711
International Journal of Modern Trends in Engineering and Research
ISSN (ONLINE) : 2349 - 9745 ISSN (PRINT) : 2393 - 8161

Performance analysis of POMDP for tcp good put improvement in cognitive radio network

Pallavi K. Jadhav[1], Prof. Dr. S.V.Sankpal[2]
[1]ME (E & TC) , D. Y. Patil college of Engg. & Tech. Kolhapur, Maharashtra, India.
[2]ME (E & TC) , D. Y. Patil college of Engg. & Tech. Kolhapur, Maharashtra, India.

Abstract: - In cognitive radio network tcp good put is one of the measure issues to improve the performance of the CR network. However most research work concentrated on performance improvement of tcp has two weaknesses as follows:-The underlying parameters are only considered to increase the TCP goodput, keeping the transport layer parameter unchanged. The second is formulated as the markov decision process in which the complete knowledge of the network is to be known. Hence to solve the above problem a POMDP base algorithm is proposed in this paper. In this proposed work each CRN users autonomously decides modulation type and power to be transmitted in PHY layer, channels which is to be selected in MAC layer to get best TCP Good put. As the channel is free space and the environment has perception error, this issue is formulated as Partial Observable Markov Decision Process (POMDP). Simulation result shows that the network can learn the optimal policy to improve the tcp good put in cognitive radio network.

Keywords: Cognitive radio network, Pomdp, Optimal parameter configuration.

I. INTRODUCTION

Wireless communication has enhance a lot now a days, and with this continues development of wireless application, the spectrum resources is becoming increasingly tense. Hence for the purpose of making effective use of radio resources, cognitive radio (CR) technology [1] has made lot of research in communication industry. Multi-user CRN can be modeled as a dynamic network which consists of interconnected CR users; by interacting with environment to improve network performance.TCP goodput is one of the important issues to measure network performance with following related research efforts.

UCLA Laboratory has done numerous relevant works of optimal parameters configuration in distributed CRNs. [2-3] proposed a priority queues based and a decomposition principle based optimal routing algorithm, respectively; By combining layered MDP with dynamic programming, [4][5] proposed a on-line routing scheme and a joint design of optimal routing and power, separately. However, the above studies have concentrated on knowing the complete knowledge. To put this similar technology in CRN, a new model is needed due to incomplete perception or the due to the existence of perception error.

To overcome the above problem, a POMDP based optimal parameters configuration is proposed in this paper for TCP goodput improvement, with Q-value iteration to find the optimal strategy.
By introducing Partial Observable Markov Decision Process (POMDP) of sequential decision making, solves the problem of spectrum access and spectrum sensing joint design, the optimal spectrum sensing problem, which both presented cognitive MAC protocols under interference constraints of registered users to maximize network performance. Such methods successfully employed POMDP theory into CRNs, Providing some reference significance for CR optimization design of perception error.

International Journal of Modern Trends in Engineering and Research (IJMTER)
Volume 02, Issue 08, [August– 2015] ISSN (Online):2349–9745 ; ISSN (Print):2393-8161

II. PROPOSED WORK

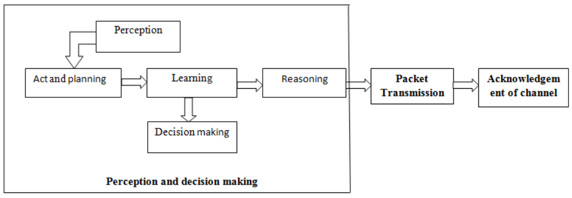

Fig 1:- Block schematic of proposed system.

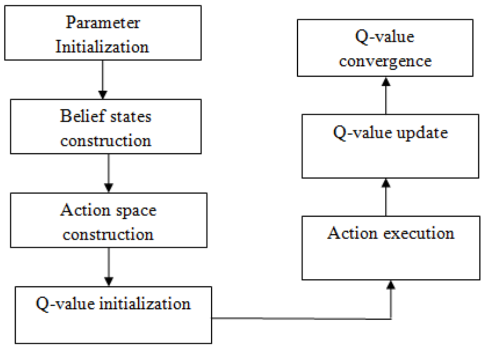

Fig 2. Work flow of proposed work.

III. POMDP FRAMEWORK

Sequential decision making is provided by the natural model of POMDP under uncertainty. A framework of MDP is augmented by this model to the situation where the secondary user cannot reliably identify the underlying environment of spectrum occupancy state. The important characteristics which keep the POMDP apart from different models are that the state is not directly observable. Instead the agent can only perceive observations which convey incomplete information about the world's state [6]. It is very important tool which increases the application of MDP to many realistic problems. POMDP is characterized by seven distinct quantities namely states (S), actions (A), observations (Θ), reward (R) and the three probability distributions namely transition probabilities (P), initial belief ($b0$), and observation probabilities (θ) [6].All of this items together describes the probabilistic system model that underlies each POMDP. In this work we have not studied deep about the development and analysis of the POMDP solution, instead we will make use of available POMDP solution in our paper to achieve optimal parameters.

3.1. Pomdp Formulation:-

Let s denotes the instantaneous state of the system, so that the finite set is denoted by $S = \{s_1, s_2 \ldots\ldots s_n\}$ and the nth channel state is denoted by $S_n(t)$. As Under the POMDP frame work the state of the system is not directly observable by the CR users so the CR user can calculate only the belief state over the state space. CR nodes take the sensor measurement result regarding the information of the belief state. Φ denotes the sensor measurement result such that $\Phi = \{\Phi 1, \Phi_2, \Phi_3 \ldots.\}$ and the nth channel observation is denoted by $\Phi n(t)$. Hence due to the spectrum sensing error $\Phi n(t)$is a incomplete projection of n^{th} channel at time t. The POMDP framework can be defined precisely only by specifying the state transition and observations by probabilistic law. This law includes the initial probability distribution (b_0) which gives the probability of the system to be in state s at time t=0, provided that this distribution is defined over all states in S.

The topology used is distributed topology hence CR users can get only the part state information of the whole network. Therefore, the dynamic parameters configuration is formulated as a POMDP which can be defined as a tuple (S, A, P, R, Z, O).In this model, S defines the all possible state space, A defines the array of action to be performed in state s: R:$S\times A\rightarrow$R represents the reward function that gives reward for the action perform in state s; P : $S\times A\times S\rightarrow S$ is the state transition probability for the state transition from s to s' ; Z stands for the set of observable history information which will give history base on the observations; O : $S\times A\times Z\rightarrow O$ depicts the observation function, which can calculate the potential observation of next state after an action.The brief information of each element in POMDP is given bellow.

3.2. System State:-

System state defines the set of possible states. It gives channel gain as $s^n=(G^n)$ in POMDP as system state, where G^n is a matrix of channel gain and $gcl^n=g^n(c,l)$ is the channel gain of c.

3.3. System Action:-

Let $A^n=(Pow^n, mod^n, X^n)$ depict the action space of n^{th} slot ,where $A^n=A_1{}^n\times\ldots\times A_L{}^n$, $A_l{}^n$ is the action space of l. Pow^n and $mod^n\in R^L$ represent transmission power vector and modulation type vector, respectively. $X^n=(X_1{}^n,\ldots,X_L{}^n)$stands for the channel allocation vector, which meets the condition $X_l{}^n=\{x\in\{0,1\}C$, $x\cdot Y^n=0\}$, in other words, the channels employed by CRN cannot conflict with registered network's, and the channels selected by each CR user should be less than or equal to m_l.

3.4. Reward Function:-

For the packet transmission, action $a_l{}^n$ is performed by the CRs users. Corresponding to the action perform TCP goodput is taken as the reward in acknowledgement stage of each slot.

$$r(s^n, a^n) = \frac{\sum_{Xl}\, n(c)=1 Th^n(c,l)}{\sum_{Xl}\, n(c)=1\, Band(c)} \ldots\ldots(1)$$

Where band(c) is the bandwidth of channel c. The product of TCP good put for each CR user is expressed as the average network utility, which is presented by equation.

$$R(S^n,a^n) = \prod_{i=1}^{L} r(S^n, a_l^n) \ldots\ldots(2)$$

3.5. Observation History and Observation Function:-

Let z^n represent the history information collection of past n slots, where $z^n=\{s^0, a^0, r^0, \ldots, s^n, a^n, r^n\}$ .This includes three elements such as state, action and reward function. O is the confidence probability which represents the distribution function of system states from s^n to s^{n+1} after action a^n. This transition takes place base on the history observation information which is express by $o(s^{n+1},a^n,z^n)=Pr(z^n|s^{n+1}, a^n)$.

3.6. Belief States:-

As in POMDP each CR user does not have the complete information, but only consist of the partial information, history action and immediate reward for policy decision, hence POMDP is a non-Markov problem. Hence this can be difficult to solve if large state space is consider. So we have transfer the POMDP into belief MDP by including belief states. The brief information about the belief state formation is given below. Let B represent the belief state such that B : $O\times A\times B\rightarrow B$ which gives the distribution of each state s^n. The system belief state is of sufficient statistics for obtaining the optimal action policy A*. Hence to get the optimal policy the solution of the POMDP can be converted into

Volume 02, Issue 08, [August– 2015] ISSN (Online):2349–9745 ; ISSN (Print):2393-8161

belief state and policy π i.e. solution of BMDP. The model is described by figure 2: 1) SE: $O \times A \times B(S) \rightarrow B(S)$, 2) π: $B(S) \rightarrow A$.

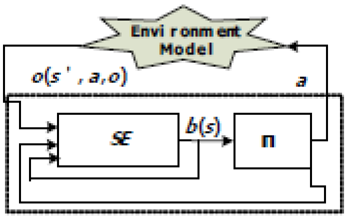

Fig.3 The BMDP model

According to belief state s^n, action a^n and observation function o^n, the belief probability of state s^{n+1} can be calculated by Bayesian formula and as equation (3).

$b(S^{n+1})=\eta o(S^{n+1},a^n,z^n) \quad \sum_{S} n \quad \in S$
$P(s^n,a^n,s^{n+1})b(S^n)$…………..……………………….……..(3)

Where η is regularization constant. In POMDP model, belief state based instantaneous reward can be weighted by all possible states as (4).

$r(\bar{b},a)=\sum_{s\in S} b(s)r(s,a)$………………………………………………...………..……..(4)

3.7. BMDP Based Q-Learning:-

Q-learning is machine learning. It does not have the prior knowledge about the environment but gets it by learning through the environment on trial and error bases. Hence action is performed by the CR users. Corresponding to each action reward is received. The rewards are of two type positive or negative reward. If the positive reward is obtained than the action trend is enhance, but if the reward received is negative in respect to the action perform then action trend might be reduced independent of a priori probability.

Consider mapping Q-learning to the BMDP four-tuple <B, A, R, SE>, the implementation process can be as follows: In this the agent gets the belief state b(s) and then performs an action that leads to the probability b(s') of transforming the state to new state that is s'. Here the immediate reward will be obtained based on the action perform. There are three main components of Q-learning: action policy, Q-value and reward function. The other two components are as follows.

3.8. Q-BMDP

Under BMDP model, $Q_\pi(\bar{b}, a)$ is defined as the expected discounted cumulative reward that can be written as (5), during which the agent performs an action *a* according to policy π and then the next action by the same policy.

$Q_\pi(\bar{b}, a) =\sum_{s\in S} b(s)E_\pi [\sum_{n=0}^{\infty} \Upsilon^n r(s^n,a^n)|s^0=s,a^0=a]$ ……………………………...….(5)

$0\leq\gamma\leq1$ is a discount factor which indicates the weighs of future rewards on current Q-value. The Q-value can be rewritten as (6) following (4), (5).

$Q_\pi (\vec{b}, a)= E [\sum_{n=0}^{\infty} \Upsilon^n r(b^n,a^n)|s^0=s,a^0=a]$ ……………………………….....(6)

After each action a^n, the environment state transforms to s^{n+1}, the corresponding $b(s^{n+1})$ can be calculated by (3) and Q-value by equation (7).

$$Q(\vec{b}, a) \leftarrow Q(\vec{b}, a) + \alpha(n)\delta(\bar{b}, a) \quad \ldots\ldots(7)$$

Where $\alpha(n)$ denotes the learning rate and $\delta(\vec{b}, a)$ is the step error from slot n to $n+1$, which can be expressed by (8).

$$\delta(\bar{b}, a) = r(\bar{b}, a) + \gamma \max_{a'} Q(\vec{b'}, a') - Q(\bar{b}, a) \quad \ldots\ldots (8)$$

For a particular application, the Q value of every state action pair can be derived from sequence $\{Q_n(\bar{b},a)\}$ generated by policy π. Sequence $\{Q_\pi(\vec{b},a)\}$ can converge to the optimal values with $\pi^*(B)$ depicting the corresponding optimal strategy. Thus the action policy of BMDP is derived, i.e.,

$$\pi^*(\vec{b}) = \arg\max_a Q^*(\vec{b}, a) \quad \ldots\ldots(9)$$

IV. SIMULATION RESULTS ANALYSIS

Here we have taken NS2 platform. Under this platform we have assume 75 CR users are randomly distributed in a square area of 600mx500m and they can access 5 wireless channels. Each channel occupancy is given by pu for the registered users. The TCP packet length *Ltcp* is set to be 1500 bytes and the maximum number of retransmission *Nre* is 5. The maximum congestion window *cwnd* is given by 6000 bytes and initialize timeout *T0* is 2*s*. The ARQ protocol is selected in MAC layer and the maximum frame retransmission *Nfr* is 10, of which header length *Lfrh* is set to be 20 bits. The ACK frame length *Lack* is 24 bits and the bandwidth is assumed to be 1MHz. This paper assumes that each CR user can either be a sender or a receiver in a certain slot, while all of them are working abidingly.
After the simulation of 30 slots, the fig 4 calculated the delay for the packet transmission for the different nodes. According to the delay calculated, the throughput is calculated as shown in the figure 6. From figure 7 and 8 the optimal goodput and the average goodput is been calculated for different nodes by the simulation using the POMDP algorithm as shown in the graph below.

4.1 Simulation Result:

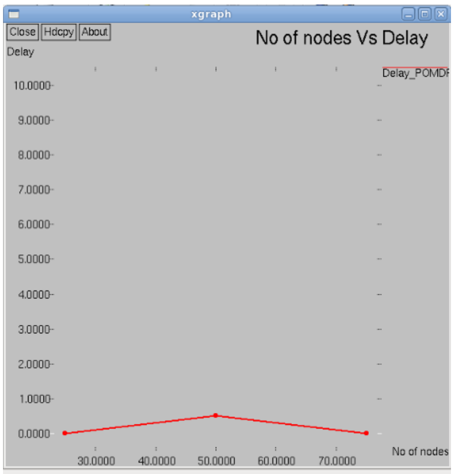

Fig 4:- Graph of no. of nodes vs delay

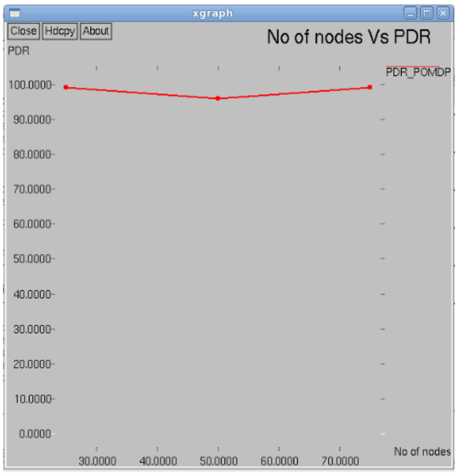

Fig 5:- Graph no. of nodes vs PDR

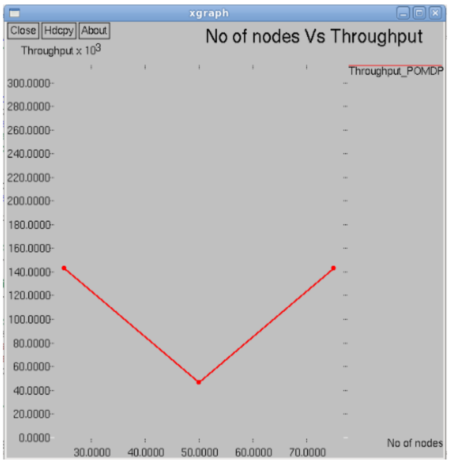

Fig 6:- Graph of no. of nodes vs Throughput

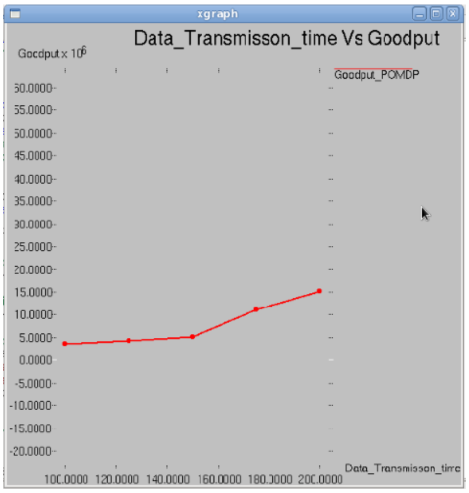

Fig 7:- Graph of Data transmission time vs Goodput

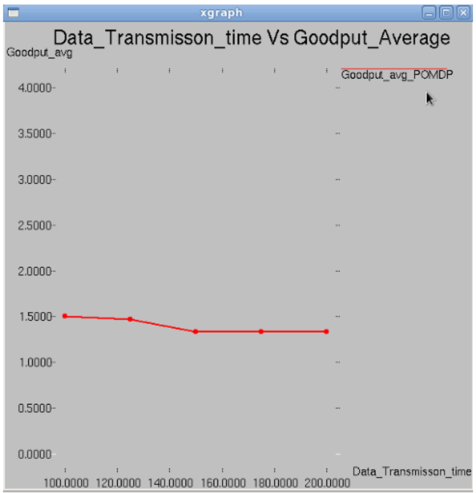

Fig 8:-Graph of data transmission time vs. good put average

I. CONCLUSION

In the distributed network, end to end TCP performance is one of the main criteria to measure the network performance. This performance in the previous studies is based on the complete knowledge of the network and perception error is not considered. This paper present a POMDP base optimal parameter configuration schemes in CRN where the complete knowledge of the network is not necessary and the POMDP is based on the partial information about the network. The optimal parameters are obtained by Q-value iteration for maximizing the TCP Goodput. In this paper the POMDP can find the optimal parameter in environment of perception error as compared with the traditional Q-MDP algorithm.

REFERENCES

[1] S. Hykin, “Cognitive Radio: brain-empowered wireless communications,” *IEEE J.Sel. Areas* Commun, Vol. 23, no. 2, pp. 201-220, 2005.

[2] Hsien-Po Shiang and Mihaela van der Schaar, “Multi-User Video Streaming Over Multi-Hop Wireless Networks: A Distributed, Cross-Layer Approach Based on Priority Queuing, ”IEEE Journal on Selected Areas in Communications, vol. 25(4), 2007.

[3] Fangwen Fu, Mihaela van der Schaar, “Decomposition Principles and Online Learning in Cross-Layer Optimization for Delay-Sensitive Applications, ”IEEE Transactions on Signal Processing, vol. 58(3),2010.

[4] Yu Zhang, Fangwen Fu and Mihaela van der Schaar, “On-Line Learning and Optimization for Wireless Video Transmission, ” IEEE Transactions on Signal Processing, vol. 58(6), 2010.

[5] Zhichu Lin, Mihaela van der Schaar, “Autonomic and Distributed Joint Routing and Power Control for Delay-Sensitive Applications in Multi-Hop Wireless Networks,” IEEE Transactions on Wireless Communications, vol(1), 2011.

[6] J. Pineau, Tractable planning under uncertainty: Exploiting structure, Ph.D. Dissertation, Rutgers University, 2004.

[7] Q. Zhang, S.A. Kassam, Finite-state Markov model for Rayleigh fading channels, IEEE Transactions on Communications47(11)(199)1688-1692.

