TELKOMNIKA Indonesian Journal of Electrical Engineering
Vol.12, No.1, January 2014, pp. 149 ~ 156
DOI: http://dx.doi.org/10.11591/telkomnika.v12i1.3900

Cognitive Radio Channel Selection Strategy Based on Experience-Weighted Attraction Learning

Sun Yong*, Qian Jiansheng
School of Information and Electrical Engineering (SIEE), China University of Mining and Technology (CUMT), Xuzhou 221008
*Corresponding author, e-mail: SnowandSunshine@msn.com

Abstract
In this paper, an innovative proposed channel selection algorithm based on Experience-Weighted Attraction (EWA) learning allows Cognitive Radio (CR) to learn radio environment communication channel characteristics online. By accumulating the history channel experience, it can predict, select and change the current optimal communication channel, dynamic ensure the quality of communication links and finally reduce system communication outage probability. Validation and reliability have been strictly verified by Matlab simulations.

Keywords: *Wireless communications, Cognitive Radio, Experience-Weighted Attraction (EWA)*

1. Introduction

Traditionally, the licensed radio spectrum allocations are regulated by official authorities. The public and government use of radio spectrum is managed by the National Telecommunications and Information Administration (NTIA) and the Federal Communications Commission (FCC) is in charge of commercial radio resources respectively in the USA. However, as more and more applications of wireless devices, the rapid increasing requisition for radio spectrum licensing has led to current shortage of radio spectrum allocations and put their governing bodies into trouble. In fact, FCC's recent research has shown that these fixed static frequency channels are always idle or not occupied in most of time. Spectrum bands are not efficiently used and underutilization either at a temporal or on a geographical level. By seeking "spectrum holes" (unused frequency channels), Cognitive Radio(CR) can highly improve spectrum resources efficiency and solve these problems presented above in "secondary utilization" (with lower priority than legacy users) way. First introduced by Mitola [1], Cognitive Radio is often considered as an extension and expansion of Soft Radio(SR), which is equipped by general hardware and capable of programming to transmit and receive various radio waves. There have already been lots of research in many aspects of Cognitive Radio.In sensing, Hossain, M.S., et al. [2] evaluated the performance of cooperative spectrum sensing with the hard combination OR, AND and MAJORITY rules. Zhang, D., et al. [3] proposed a novel detection algorithm that the Fractal box dimension is used when the signal to noise (SNR) is high, while the improved TCC algorithm is used when the SNR is low and Khalaf, G. [4] formulated the detection problem based on the eigen-decomposition technique. In security,Shan-Shan, W., et al. [5] proposed a Four Dimensional Continuous Time Markov Chain model to analyze the communication performance of normal secondary users under PUEAs, typically affected by SMUs, and compared several PUEA detection schemes.

As revolutionary development of Intelligent Radio(IR), Cognitive Radio implements Soft Radio by adding Knowledge Base, Reasoning Engine and Learning Engine to be an independent Cognitive Engine (CE), which makes the radio capable of learning and adapting to the surrounding radio environment [6]. Knowledge Base is very common in Artificial Intelligence (AI) logic planning and can be seen as memory in human's brain and stores variety of cases, relations and rules. Reasoning Engine is just like expert system in Artificial Intelligence and executes all kinds of state information for reference of knowledge base by logic thinking and finally generates processed results or actions, driving soft radio changing setting parameters to adapt to changing environment. As the core component and key feature for Cognitive Radio implementation, Learning Engine is in charge of keeping Knowledge Base updated by

Received June 16, 2013; Revised August 24, 2013; Accepted September, 2013

accumulating new environmental experience into new knowledge extension and is what differentiates Cognitive Radio from traditional pre-programmed ones.

There are varieties of learning algorithms available for Cognitive Radio, including neural networks, genetic models, and hidden Markov algorithms. Tsagkaris, K., A. Katidiotis, et al. [7] used neural network-based learning to predict data bit rate of Cognitive Radio. Galindo-Serrano, A. and L. Giupponi [8] proposed a form of real-time decentralized Q-learning to manage the aggregated interference generated by multiple WRAN systems. Li, H. S. [9] applied Multi-agent reinforcement leaning (MARL) for the secondary users to learn good strategies of channel selection. Chen, X. F., Z. F. Zhao, et al. [10] presented an intelligent policy based on reinforcement learning to acquire the stochastic behavior of Primary Users (PUs). Zhang, W. Z. and X. C. Liu [11] obtained the capability of iteratively on-line learning environment performance by using Reinforcement Learning (RL) algorithm after observing the variability and uncertainty of the heterogeneous wireless networks. Gallego, J. R., M. Canales, et al. [12] provided no-regret learning algorithms to perform the joint channel and power allocation and overcome the convergence limitations of the local game. Zhu, J., J. Wang, et al. [13] employed Reinforcement learning (RL) approach to finding a near-optimal policy under undiscovered environment. Torkestani, J. A. and M. R. Meybodi [14] proposed the learning automata-based cognitive radio to address the spectrum scarcity challenges in wireless ad hoc networks. Yang, M. F. and D. Grace [15] improved channel assignment in multicast terrestrial communication systems with distributed channel occupancy detection by using intelligence based on reinforcement learning and transmitter power adjustment. Zhou, P., Y. S. Chang, et al. [16] designed a robust distributed power control algorithm with low implementation complexity for Cognitive Radio networks through reinforcement learning, which does not require the interference channel and power strategy information among Secondary Users (SUs) and from SUs users to PUs.

However, as known till now with our best effort, little focus has been placed on implementing learning engine of cognitive radio with Experience-Weighted Attraction (EWA) algorithms. The innovative proposed channel selection algorithm based on EWA learning in this paper allows cognitive to learn radio environment communication channel characteristics online. By accumulating the history channel experience, it can predict, select and change the current optimal communication channel, dynamic ensure the quality of communication links and finally reduce system communication outage probability.

The rest of this paper is presented as follows. In section 2, the EWA algorithms will be introduced in full details. Intelligent channel selection algorithm of cognitive based on EWA learning is following in section 3. Then the simulation results comparison and analysis are presented in section 4. In the end the conclusion comes in the final section 5.

2. EWA Learning

Experience-Weighted Attraction(EWA) is derived from normal form multi-game theory [17]. Setting n players in the game, and denote $i(i = 1,2,3,\ldots,n-1,n)$ for each one. The strategy space for player i is S_i. There are m_i discrete choices in total for each S_i and can be expressed as $S_i = \{s_i^1, s_i^2, s_i^3, \ldots, s_i^{m_i-1}, s_i^{m_i}\}$. One strategy of player i, denoted by s_i, is the element of strategy space S_i, or $s_i \in S_i$. The entire strategy space of the game S is the n-Cartesian product of individual strategy space, that is, $S = S_1 \times S_2 \times S_3 \times \cdots \times S_{n-1} \times S_n$. Let s be the combination of all players' strategies in the game, then $s = \{s_1, s_2, s_3, \cdots, s_{n-1}, s_n\} \in S$. The combination of all other n-1 players' strategies except player i can be expressed as $s_{-i} = \{s_1, s_2, s_3, \cdots, s_{i-1}, s_{i+1} \cdots, s_{n-1}, s_n\}$. Denote m_{-i} be the combination number of s_{-i}, then $m_{-i} = \prod_{j=1,j\neq i}^{n} m_j$. The reward function for player i with scalar-value is $\pi_i(s_i, s_{-i})$. Take time dimension for consideration, then the strategy of player i in time period t can be expressed as $s_i(t)$, other players' strategy set(vector) $s_{-i}(t)$, and reward function for player i $\pi_i[s_i(t), s_{-i}(t)]$ respectively.

EWA learning algorithm assumes that any strategy has an attraction value. The model defines the initial values of attractions, how the attraction values are updated based on experience and determines the selection probabilities. The core algorithm is two variables updated each round. One is attraction value $A_i^j(t)$, it defines the attraction value of player i after selecting strategy j. The attraction update rule is that the current attraction value is the last attraction value $A_i^j(t-1)$ multiplied by attenuation coefficient φ plus (virtual) reward $\pi_i[s_i(t), s_{-i}(t)]$, then normalized by the updated experience weight $N(t)$. In mathematical form:

$$A_i^j(t) = \frac{\phi \cdot N(t-1) \cdot A_i^j(t-1) + \{\delta + (1-\delta) \cdot I[s_i^j, s_i(t)]\} \cdot \pi_i[s_i^j, s_{-i}(t)]}{N(t)}$$

Where I [·] is the indicator function, which is defined as follows

$$I(x,y) = \begin{cases} 1, x = y \\ 0, x \neq y \end{cases}$$

The algorithm updates attraction value by reward after selecting corresponding strategy and weight coefficient δ multiplying virtual reward it would have yielded if select other strategy in hypothetical scenes. Parameter δ is used to measure the weight of virtual reward in relative to the real weight, and virtual reward can be understood as estimate and expectation of payoffs after selecting other alternative strategy.

The other is experience weight $N(t)$, which can be interpreted as equivalent observation of past experience with respect to present experience. The bigger the value of $N(t)$ is, the greater the influence of past experience to current attraction. The update rule of $N(t)$ is that the last experience weight $N(t-1)$ multiplied by attenuation coefficient ρ, then plus Incremental value of 1. That is

$$N(t) = \rho \cdot N(t-1) + 1$$

Denote $N(0)$ be the initial value of $N(t)$, while the initial value of $A_i^j(t)$ is $A_i^j(0)$, and $N(0)$ can be seen as equivalent assess of pregame thinking.

The value of attraction determines the probability of strategy selection. In other words, probability function of $P_i^j(t)$ should monotonously increase with $A_i^j(t)$. The mathematical expression of probability in exponential form is

$$P_i^j(t+1) = \frac{e^{\lambda \cdot A_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_i^k(t)}}$$

Parameter λ above is used to measure the player's sensitivity to attraction value. The more sensitive the player to the attraction, the bigger its value is.

3. Channel Selection Model

In the problem of radio communication channel selection, different wireless channels should have different channel availabilities, that is, the channel idle probabilities should be in difference for cognitive radio. Assuming radio propagation environment can be divided into n channels, then the idle probability of channel i can be expressed as α_i, or $A = \{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_{n-1}, \alpha_n\}$ in vector form. Let β_i be the successful transmission probability of channel i, then $B = \{\beta_1, \beta_2, \beta_3, \ldots, \beta_{n-1}, \beta_n\}$. Think of the radio channel characteristics change over time, the channel idle probability and successful transmission probability of channel i should not be the same at different time t, then the form of probabilities after introducing time parameter t are $A(t) = \{\alpha_1(t), \alpha_2(t), \alpha_3(t), \ldots, \alpha_{n-1}(t), \alpha_n(t)\}$ and $B(t) = \{\beta_1(t), \beta_2(t), \beta_3(t), \ldots, \beta_{n-1}(t), \beta_n(t)\}$ respectively.

According to EWA, the update equation of experience weight $N(t)$ is

$$N(t) = \rho \cdot N(t-1) + 1$$

However, in order to accommodate the cognitive radio channel selection and transmission characteristics, the update mathematical expression of attraction $A_i^j(t)$ should be modified or improved as follows

$$A_i^j(t) = \frac{\phi \cdot N(t-1) \cdot A_i^j(t-1)}{N(t)} + \frac{\{\delta + (1-\delta) \cdot I[s_i^j, s_i(t)]\} \cdot \langle \eta + (1-\eta) \cdot I[1, x(j)] \cdot I[s_i^j, s_i(t)] + (1-\eta) \cdot \{1 - I[s_i^j, s_i(t)]\} \rangle \cdot \pi_i[s_i^j, s_{-i}(t)]}{N(t)}$$

Where

$$x(j) = \begin{cases} 0, \text{Transmission failure on channel j} \\ 1, \text{Successful transmission on channel j} \end{cases}$$

$$\pi_i[s_i^j, s_{-i}(t)] = \begin{cases} 0, \text{channel j is sensing busy} \\ 1, \text{channel j is sensing idle} \end{cases}$$

And the probability of channel selection $P_i^j(t)$ meets equation of

$$P_i^j(t+1) = \frac{e^{\lambda \cdot A_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_i^k(t)}}$$

From the equation above, in the period of radio environment sensing of Cognitive Radio, when perceiving the current state of the channel j busy(strong Electromagnetic noise over interference threshold for transmission), the state flag status is set to 0 (unavailable), and the strategy of selecting channel j for transmission channel will get no payoff, or the award function value of $\pi_i[s_i^j, s_{-i}(t)]$ is 0; while perceiving the current state of the channel j idle(electromagnetic noise below interference threshold for transmission), the state flag status is set to 1(available), and the strategy of selecting channel j for transmission channel will get the payoff of $\pi_i[s_i^j, s_{-i}(t)]$ respectively. In addition, the value of $\pi_i[s_i^j, s_{-i}(t)]$ is assumed to equal 1.

These available channels are candidate channels for channel selection of Cognitive Radio, and the candidate channel with the highest probability of channel selection will be chosen for transmission. However, other candidate channels unselected will get the virtual award of $\delta \cdot \pi_i[s_i^j, s_{-i}(t)]$. Next reaches the innovation of this paper: a feedback mechanism of channel transmission is introduced and the update of selected channel attraction is divided in two phases. Before the transmission starts, selected channel gets some actual award of $\eta \cdot \pi_i[s_i^j, s_{-i}(t)]$ if feedback coefficient is set to $1-\eta$. After successful transmission, the selected channel will get another $(1-\eta) \cdot \pi_i[s_i^j, s_{-i}(t)]$ award as payoff. But if the transmission is unfortunately failure, the selected channel will not get the second part of the actual award.

4. Results and Discussion

Assuming the number of channels in simulation environment is 5, or n =5. The initial value of attraction $A_i^j(0)$ and experience weight $N(0)$ are the same to 1, while the selection probability of each channel is the same in the initial state, that is, $P_i^j(0) = 1/n = 1/5 = 0.2$. For the coefficients, ϕ for attraction is 0.72, ρ for experience weight is 0.2, δ for virtual award is 0.4, η for feedback is 0.6 and λ for sensitivity is 0.9. The initial channel idle probability vector $A_0 = \{0.4,0.9,0.6,0.5,0.7\}$, while the initial channel successful transmission probability vector $B_0 = \{3/4,8/9,5/6,4/5,6/7\}$. Then the initial channel available probability vector $\Gamma_0 = A_0 \cdot B_0 = \{0.3,0.8,0.5,0.4,0.6\}$. In order to verify this intelligent algorithm is capable to decide and guide cognitive radio real-time switch to the newly transmission channel with the highest available probability online accurately, the channel idle probability vector will change to $A_1 = \{0.6,0.4,0.7,0.9,0.5\}$, and the channel successful transmission probability vector will change to $B_1 = \{5/6,3/4,6/7,8/9,4/5\}$ after 33 rounds during the simulation process. Therefore the channel available probability vector will be $\Gamma_1 = A_1 \cdot B_1 = \{0.5,0.3,0.6,0.8,0.4\}$ after the simulation environment change.

After parameters above are set, the simulation result of channel selection algorithm based on EWA learning is shown in Figure 1

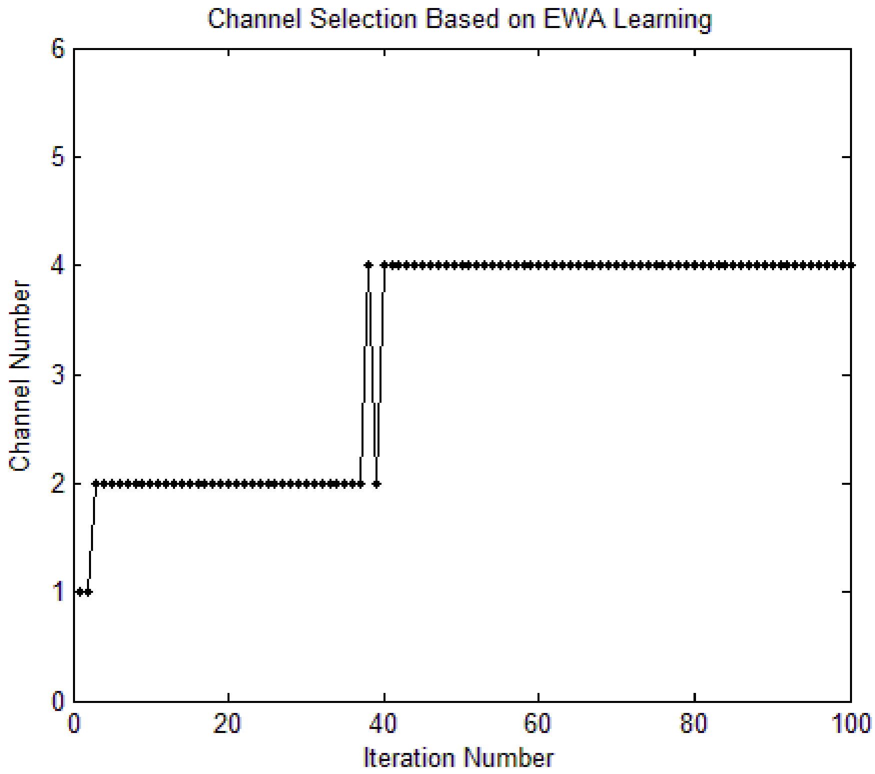

Figure 1. The Result of Channel Selection Based on EWA Learning

In the figure above, the intelligent algorithm of Cognitive Radio select channel 2 with the highest availability of 0.8 as its transmission channel at 3rd round(only 2 rounds delay), then this state keeps stable until environment changes. It shows that channel selection algorithm based on EWA learning can quickly track and lock the highest availability channel from the initial state with high stability.

When radio environment changes after 33 rounds, the intelligent algorithm of Cognitive Radio first selects channel 4 with the highest availability of 0.8 as its transmission channel at 38th round(only 4 rounds delay). After experiencing one round of unstable state, it selects channel 4 for the second time at 40th round and keep stable afterwards. It shows that channel selection algorithm based on EWA learning can also quickly track and lock new highest availability channel as radio environment changes with high stability.

The explanation about turning back event at 39th round is that as selection probability of channel 4 is first higher than selection probability of channel 2 with small margin after 37 rounds, then channel 4 is selected as transmission channel at 38th round for the first time instead of channel 2. Unfortunately, there is accidental transmission failure in channel 4 at 38th round as channel 4 has little but does exist transmission failure probability of $1/9$, and this failure directly results of sharp decreasing in reward, attraction and selection probability of channel 4 to lower than selection probability of channel 2 again, then it is no exception to return to select channel 2 at 39th round. This interpretation can also be verified in the following analysis.

Channel selection probability based on EWA learning is illustrated in Figure 2. From the figure above, it can be perceived that the selection probability of channel 2 ascends rapidly after the initial 2 rounds, then fluctuated slightly when reaching around the probability of 0.5. For the reason of channel availability probability changes after 34th round, the selection probability of channel 2 falls dramatically, while the selection probability of channel 4 increases respectively and steadily overtake the selection probability of channel 2 after 39 rounds.

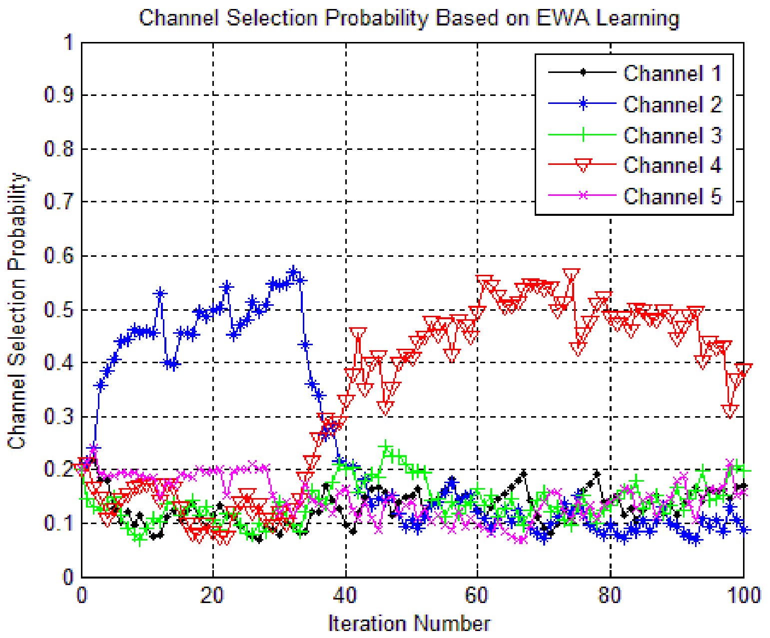

Figure 2. Channel Selection Probability Based on EWA Learning

Channel experience weighted attraction based on EWA learning is depicted in Figure 3. It can be clearly seen from the figure above that the attraction of channel 4 drops at 38th round, and that confirms the interpretation of turning back event at 39th round. Then the attraction of channel 4 back to rise and eventually reaches a relatively stable state.

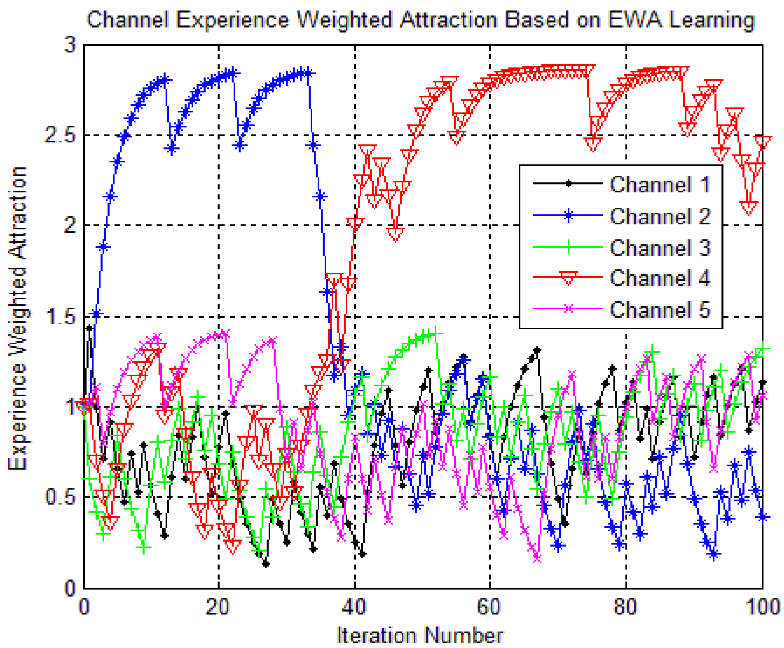

Figure 3. Channel Experience Weighted Attraction Based on EWA Learning

In order to highlight better performance of channel selection algorithm based on EWA learning than other traditional radio with fixed transmission channel, the times of available to access the channel and successful completion of transmissions in 100 rounds are collected in 3 scenes: fixed channel 2 as transmission channel, fixed channel 4 as transmission channel and channel selection algorithm based on EWA learning. The statistical data is compared in Figure 4

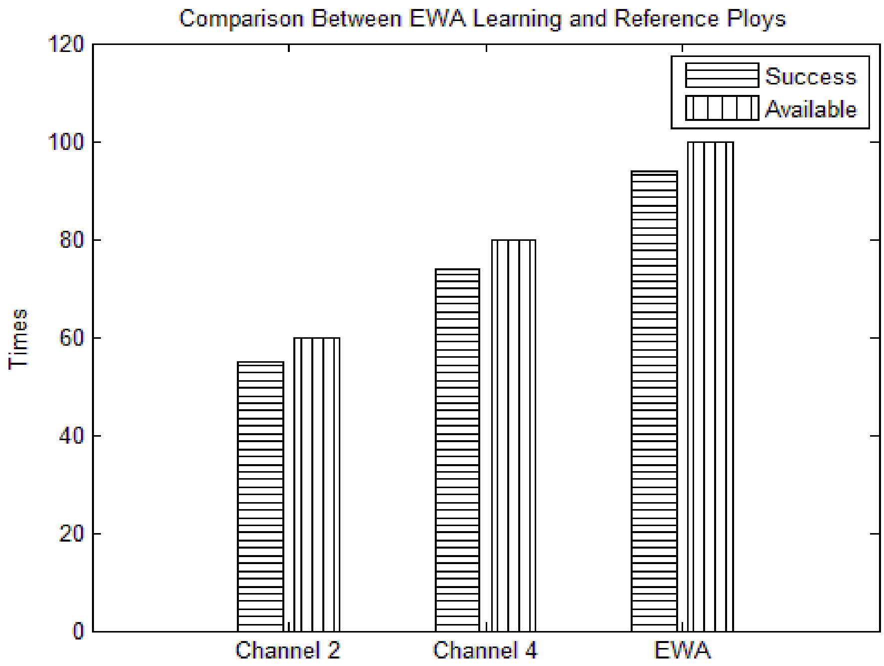

Figure 4. Comparison Between EWA Learning and Reference Ploys

The number of available to access the channel with fixed channel 2 as transmission channel is 60, and the number of successful completion of transmissions with fixed channel 2 as transmission channel is 55; the number of available to access the channel with fixed channel 4 as transmission channel is 80, and the number of successful completion of transmissions with fixed channel 4 as transmission channel is 74. While the number of available to access the channel with channel selection algorithm based on EWA learning is 100 with no block, and the number of successful completion of transmissions with channel selection algorithm based on EWA learning reaches the point of 94. Finally, the probability of successful completion of transmission with channel selection algorithm based on EWA learning is 94%, much higher than that of channel 2(55%) and channel 2(74%). By evident statistical comparison, channel selection algorithm based on EWA learning can greatly improve the probabilities of successful channel access and transmission completion.

5. Conclusion

In this paper, an innovative channel selection algorithm based on EWA learning is proposed and a feedback mechanism of channel transmission is also introduced to facilitate Cognitive Radio to learn radio environment communication channel characteristics online. By accumulating the history channel experience, it can predict, select optimal communication channel, and capable to decide and guide cognitive radio real-time switch to the newly transmission channel with the highest available probability online accurately, dynamic ensure the quality of communication links and finally reduce system communication outage probability. Compared to traditional radio with fixed transmission channel, this intelligent algorithm can greatly improve the probabilities of successful channel access and transmission completion.

Acknowledgements

This work was supported by the National High Technology Research and Development Program of China (863 Program) (No. 2012AA062103)

References

[1] Mitola J. Cognitive radio for flexible mobile multimedia communications. *Mobile Networks & Applications*. 2001; 6(5): 435-441.
[2] Hossain MS, et al. Hard Combination Data Fusion for Cooperative Spectrum Sensing in Cognitive Radio. *International Journal of Electrical and Computer Engineering (IJECE)*. 2012; 2(6): 811-818.
[3] Zhang D, et al. An Improved Cognitive Radio Spectrum Sensing Algorithm. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(2): 583-590.
[4] Khalaf G. An Optimal Sinsing Algorithm for Multiband Cognitive Radio Network. *International Journal of Information and Network Security (IJINS)*. 2013; 2(1): 60-67.
[5] Shan-Shan W, et al. Primary User Emulation Attacks Analysis for Cognitive Radio Networks Communication. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(7): 3905-3914.
[6] Bantouna A, et al. An overview of learning mechanisms for cognitive systems. *Eurasip Journal on Wireless Communications and Networking*. 2012: 1-6.
[7] Tsagkaris K, A Katidiotis, and P Demestichas. Neural network-based learning schemes for cognitive radio systems. *Computer Communications*. 2008; 31(14): 3394-3404.
[8] Galindo-Serrano A and L Giupponi. Distributed Q-Learning for Aggregated Interference Control in Cognitive Radio Networks. *IEEE Transactions on Vehicular Technology*. 2010; 59(4): 1823-1834.
[9] Li HS. Multiagent Q-Learning for Aloha-Like Spectrum Access in Cognitive Radio Systems. *Eurasip Journal on Wireless Communications and Networking*. 2010: 1-15.
[10] Chen XF, et al. Reinforcement Learning Enhanced Iterative Power Allocation in Stochastic Cognitive Wireless Mesh Networks. *Wireless Personal Communications*. 2011; 57(1): 89-104.
[11] Zhang WZ and XC Liu. Centralized Dynamic Spectrum Allocation in Cognitive Radio Networks Based on Fuzzy Logic and Q-Learning. *China Communications*. 2011; 8(7): 46-54.
[12] Gallego JR, M Canales and J Ortin. Distributed resource allocation in cognitive radio networks with a game learning approach to improve aggregate system capacity. *Ad Hoc Networks*. 2012; 10(6): 1076-1089.
[13] Zhu J, et al. Adaptive transmission scheduling over fading channels for energy-efficient cognitive radio networks by reinforcement learning. *Telecommunication Systems*. 2009; 42(1-2): 123-138.
[14] Torkestani JA and MR Meybodi. A Learning Automata-Based Cognitive Radio for Clustered Wireless Ad-Hoc Networks. *Journal of Network and Systems Management*. 2011; 19(2): 278-297.
[15] Yang MF and D Grace. Cognitive Radio with Reinforcement Learning Applied to Multicast Downlink Transmission with Power Adjustment. *Wireless Personal Communications*. 2011; 57(1): 73-87.
[16] Zhou P, YS Chang and JA Copeland. Reinforcement Learning for Repeated Power Control Game in Cognitive Radio Networks. *IEEE Journal on Selected Areas in Communications*, 2012; 30(1): 54-69.
[17] FENG Wen-jiang ZC, JIANG Wei-heng. Channel Selection Based on EWA Game Abstraction in Cognitive Radio Network. *Journal of Applied Sciences*. 2010; 28(6): 580-584.