Adaptive Sensing Policies for Cognitive Wireless Networks using Learning Automata

Panagiotis Sarigiannidis, *Member, IEEE,*
Malamati Louta, *Member, IEEE,*
Eleni Balasa,
Department of Informatics and telecommunications Engineering
University of Western Macedonia
Kozani, Greece
{psarigiannidis, louta}@uowm.gr,
elenibalasa@gmail.com

Thomas Lagkas, *Member, IEEE*
Department of Informatics and Telecommunications Engineering
University of Western Macedonia
Kozani, Greece
The University of Sheffield International Faculty, CITY College
Thessaloniki, Greece
tlagkas@ieee.org

***Abstract*—This paper introduces an adaptive spectrum sensing method for cognitive radio wireless networks. The proposed method enhances previously proposed random-based sensing policies, effectively selecting the licensed channels to be sensed by accurately estimating channels' availability, resulting, thus, to high system's resources utilization. The core mechanism of the adaptive method is an enhanced learning automaton, which efficiently interacts with the environment and provides accurate decisions on selecting the channel to be sensed on behalf of the secondary users. A thorough description of the introduced method is provided, while the performance of the enhanced sensing policies is verified through extensive simulation experiment.**

Keywords-cognitive radio; learning automata; multi-channel MAC; wireless networks

I. INTRODUCTION

One of the most challenging research issues to be addressed by both academia and industry is the accommodation of the ongoing developed wireless standards and products in the quite overcrowded, existing spectrum. The unlicensed frequency bands have been almost exhausted; however, substantial allocations of the spectrum remain idle or underutilized [1]. The concept of dynamic spectrum allocation comes in the light as one of the most flexible solutions to efficiently address this deficiency [2].

In essence, dynamic spectrum allocation involves two groups of users that may access and use the spectrum resources. Primary users, also known as licensed users, have the right to utilize spectrum channels; however, access opportunities are allowed to secondary users, also known as unlicensed users, when the related resources remain unoccupied by the primary users. In order to achieve this, secondary users should be able to dynamically change channels, adapting their transmission and reception configuration on demand. The cognitive radio concept satisfies this requirement, supporting dynamic alteration of the transmission frequency, modulation, data rate, and transmission power, mainly using software-defined radio (SDR) technology [1].

The cognitive radio concept involves several access issues. Thus, it entails a Media Access Control (MAC) protocol in order to practically operate. Generally speaking, a rigorous MAC protocol undertakes the following responsibilities: a) allows the secondary users to be aware of the cognitive radio configuration, i.e., number and capabilities of the licensed channels, synchronization details, and physical layer features, b) defines the way of exchanging messages (among the secondary users), c) defines the access framework of delivering data, e.g., CSMA, and d) incorporates an effective sensing policy.

In this paper our focus is laid on the sensing policy, since it constitutes one of the key players for the optimization of the network performance. In particular, the performance of the whole MAC protocol dramatically depends on the sensing outcome, as successful sensing processes lead to effective spectrum utilization. Different from the existing approaches, we design an adaptive sensing policy, capable of identifying the individual availability of each licensed channel. To this end, a powerful as well as simple learning tool is employed: the Learning Automata (LA) [3]. Each secondary user is enhanced with a learning automaton, which assists her accurately estimating an image of the channels' state. The overall aim is to take effective decisions regarding the selection of the licensed channels to be sensed. Moreover, the adaptive strategy is extended so as to beneficially replace the random channel selection in environments where the neighboring users are aware of the sender's channel selection. The performance of the proposed policy is evaluated and compared against similar sensing schemes operating in random fashion. Extensive assessment results are presented, providing evidence of the superiority of the proposed adaptive policy, improving channel availability sensing by approximately 30%.

The remainder of this paper is organized as follows. An overview of related research works is provided in Section II. In Section III, we describe the considered system architecture, while in Section IV the detailed adaptive method are presented in a detailed manner. Section V is dedicated to performance evaluation. We conclude the paper in Section VI.

II. RELATED WORK

A multitude of MAC strategies and sensing policies can be found in the literature. In [4] the Random Sensing Policy (RSP) and the Negotiation-based Sensing Policy (NSP) are

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: ARCHIMEDES III. Investing in knowledge society through the European Social Fund.

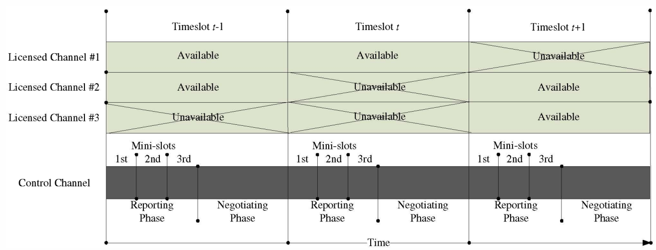

Figure 1. Cognitive system architecture with three licensed channels.

introduced. Both mechanisms operate in such a way that the information obtained by the sensing process is broadcasted in the most efficient way. However, the core procedure behind those mechanisms involves a random selection process. Specifically, having organized time into timeslots and mini-slots, in the context of the RSP, each secondary user randomly selects a single licensed channel to sense at the beginning of each timeslot. Afterwards, each secondary user reports the channel state by broadcasting a beacon in the corresponding mini-slot. This way, the RSP scheme allows the secondary users to be aware of the sensed channels in a timeslot basis. The deficiency of this approach lies in the common channel sensing case, where two or more secondary users select the same channels to perform sensing. Given that the target of a sensing policy is to maximize the channel set being sensed, the phenomenon of sensing common channels degrades the policy's performance, especially when the number of secondary users is equal or less than the number of licensed channels. To this end, the NSP extends the RSP mechanism adopting a corrective action in order to maximize the number of sensed channels per timeslot exploiting the Request To Send and Clear To Send (RTS/CTS) handshake scheme. In particular, a special byte denoting the sensed channel information is injected into the RTC/CTS packets, therefore the rest secondary users are aware of the channel sensed by the sender during the current timeslot. As a result, secondary users that decided to sense the same channel change their decision and chooses a different channel in the following timeslot, while all others remain at their selections.

In [5] the authors focus on the realization of cognitive networking using typical radio transceiver technologies. The core mechanism involves random sensing with probabilistic access. Nevertheless, the sensing policy implies random channel selection.

[6] refers to a MAC protocol that does not necessitate the usage of a dedicated common control channel for the exchange of control packets between secondary users. The proposed protocol requires synchronization among all nodes. At the beginning of each slot, every node tunes to the respective channel the slot represents and listens to that channel. The performance of the proposed model is compared against a common control channel based protocol considering throughput and network connectivity performance metrics. However, the specific protocol demands synchronization of each user in order to adequately operate.

In [7] two important issues associated with MAC-layer sensing are addressed in cognitive radio networks: a) when scanning of the availability of licensed channels should be performed and b) in which order these channels should be sensed in order firstly to maximize the discovery of spectrum holes in licensed channels to be exploited by secondary users. Concerning the sensing period optimization scheme proposed, the authors consider proactive sensing taking into account both the number of discovered opportunities and the sensing overhead cost. Similar sensing approaches are adopted in [8-10].

A common shortcoming of the sensing schemes found in the literature is the random manner in which channels to be sensed are selected. This means that the main decision of the sensing policy neglects specific access characteristics of each channel. Hence, in most cases, a static or a predefined way of examining the licensed channels is adopted, resulting in network performance degradation. Aiming to cover this gap, we propose an adaptive sensing policy, exploiting learning automata.

III. System Architecture

The cognitive wireless system under study considers a multi-channel licensed spectrum, in which the set of primary users has access to w different (licensed) channels. Fig. 1 illustrates the system architecture with three licensed channels. All network entities are synchronized to a common clock; time is organized into timeslots, where the beginning and ending point of each timeslot is predefined, identical for each channel, and known to any participating network entity. As time passes, each channel becomes available or unavailable instantly upon the beginning of each timeslot. Furthermore, we consider a simple, dedicated, and flexible control channel for message exchange between the secondary users. Control and licensed channels are considered synchronized with a common clock. In addition, each user is equipped with a couple of transceivers, one for accessing the control channel and the other for sending and receiving in data channels. The first transceiver is devoted to operating over the control channel, while the second consists of a SDR module. Thus, each user is capable of tuning to any of the w licensed data channels.

The control channel is primarily used for coordination; it is further divided into two phases, the reporting phase and the negotiation phase. During the reporting phase, the secondary users broadcast critical information about the availability of the licensed channels in a cooperative way. At the beginning of each timeslot, each secondary user senses one of each existing data channels in order to infer about its availability. The information obtained by this action is broadcasted to all secondary users, using the control channel as follows. The reporting phase is further divided into w periodical mini-slots. Hence, there is an opportunity for each secondary user to inform the rest users about the availability of the licensed channel sensed. To this end, secondary users that sensed channel's j availability broadcast a beacon during the jth mini-slot over the control channel. In this way, the secondary users cooperate in order to obtain as much as possible information regarding the current status of the licensed channels. Afterwards, the secondary users that wish to send data packets contend each other during the negotiation phase, if at least one data channel is available during the current timeslot. The negotiation takes place employing conventional hand-shake methods such as RTS and CTS message exchange between the sender and the receiver, using the control channel, without violating the ongoing timeslot duration.

IV. Adaptive Sensing Policies

A. Motivation

The sensing policy entails a set of aspects that should be carefully considered and addressed so as to operate in an efficient way. First, in order to gain complete information about their environment, the secondary users have to sense the whole spectrum, i.e., all licensed channel. Second, the secondary users ensure that all licensed channels were monitored, since no beacon message is sent if the data channel is sensed occupied. It becomes evident that if the number of secondary users is less than the number of channels, secondary users receive an incomplete image of the channel state. On the other hand, the set of secondary users may choose common channels to sense for the same timeslots, leaving other channels unmonitored. In the light of the aforementioned issues, a rigorous, efficient, and simple sensing policy is required in order to ensure an effective exploitation of the licensed channels' availability.

B. Learning Automata

A learning automaton constitutes a finite state machine that interacts with a stochastic environment and aims to perceive the optimal action offered by the environment via a learning process. Inheriting its basic aspects from the reinforcement learning field, LA act based on specific features that affect the environment. Specifically, the automaton chooses an action from a finite set of possible actions, while its decisions are updated based on feedback received from the environment concerning the impact of each selected action to the environment. A considerable number of research fields adopt LA as the main adaptive tool, such as pattern recognition, data networking, and scheduling [11]. LA could enhance the decision module in the current problem. The transceiver of each secondary user is enhanced with an automaton, which is responsible for choosing a licensed channel to be sensed in the beginning of each timeslot. The action pool is defined by the set of licensed channels. For each secondary user, the automaton decides on a channel for the current timeslot. The environment reacts by sending back feedback, which is defined as the result of the sensing process, i.e., the availability of the chosen channel. The automaton receives the feedback and updates its history, which is realized by a probability vector, expressing the availability probability of each channel. In this way, the automaton creates a whole image of the spectrum, allowing the SDR receiver to sense the channels that are more likely to be available.

C. Optimality Determination

Ideally, a sensing policy should yield the optimal number of spectrum opportunities for the secondary users, i.e., the optimal number of licensed channels sensed available. The determination of this optimal number involves the number of licensed channels as well as the number of the participating secondary users. Assuming that the number of channels is denoted by w and the number of secondary users is denoted by n, *Algorithm1* provides the optimal number of available licensed channels sensed during a timeslot. The logic behind the algorithm is quite simple. If the number of secondary users is equal or larger than the available channels during each timeslot, denoted by s ($s \leq w$), then the optimal number of sensed channels being available is s, since users could monitor all possible channels. On the contrary, in case the number of users is not large enough to cover the possible pool of channels, the optimal number of channels sensed available is equal to n.

Algorithm1: Optimality Determination

n: Number of secondary users.
o: Optimal average number of available licensed channels.
For each timeslot $t = 1,2,\dots,q$
  $s \leftarrow$ number of available channels
    If $n \geq s$
      $o = o + s$
    Else
      $o = o + n$
    End_If
End_For
$o = \frac{o}{q}$

D. Model Formulation

In this sub-section we provide the formulation of the proposed adaptive model including the definition of the action pool, the probability vector, and the feedback vector. Each learning automaton operates independently on behalf of each secondary user. Upon the beginning of each timeslot, each secondary user chooses one of the w possible actions, determining the licensed channel to be sensed at the current timeslot. Let $A = \{a_1, a_2, \dots, a_w\}, i = 1,2,\dots,n$ denote the pool of the w possible actions. The decision at each timeslot is supported by an availability probability vector, which represents the probability for selecting one of the possible actions from the A set. The probability vector at timeslot t is defined as follows: $p(t) = \{p_1(t), p_2(t), \dots, p_w(t)\}, 0 \leq p(t) \leq 1$. Each action taken by the automaton is followed by a feedback, originated from the environment, as a reaction to the channel selection. In essence, the feedback refers not only to the channel selected by the corresponding secondary user, but it reflects the behavior of all licensed channels of the spectrum. According to the sensing architecture, each secondary user broadcasts a beacon in case the channel sensed is available. Hence, the feedback encloses the status information of all licensed channels, informing the secondary user about the status of the spectrum during the current timeslot. Thus, the feedback received by each secondary user is defined as a notification vector at timeslot t as follows: $F(t) = \{F_1(t), F_2(t), \dots, F_w(t)\}$, where each element may take a logical value, i.e., true or false in case the licensed channel $1,2,\dots,w$ is available or either not available or not monitored, respectively. For instance, if the value of $F_2(4)$ is true, it means that the second channel is sensed available. The feedback is delivered by each user at the end of the reporting phase. The learning nature of the automaton lies in the probability vector update. *Algorithm2* gives the probability vector update process in detail. Upon the reception of the feedback provided by the environment at each timeslot t, the automaton updates the probability vector. Initially, the automaton of each secondary user sets the availability probability vector equal to $1/w$, which means that initially each channel is selected uniformly. Thereupon, and at each timeslot t, the set of licensed channels that are sensed available is updated ($V(t)$). For each licensed channel the corresponding feedback is examined. If the feedback implies availability, the probability of the examined licensed channel is increased, receiving a reward, otherwise it is decreased, receiving a penalty. The magnitude of the increase is governed by two factors (*L*,*b*), where the parameter b is employed in order to prevent reaching zero probabilities; parameter L controls the convergence speed of the learning process. The lower the L value, the more accurate the estimation held by the automaton—a fact, however, that comes at the expense of the convergence speed. Furthermore, the reward and the penalty define the level of the increment and the decrement respectively, of the channel's availability probability. The reward is expressed as the summation of the probability of each channel sensed unavailable (or not sensed at all), divided by the number of channels found available (line 9 of *Algorithm2*). Accordingly, a small penalty is imposed to each unavailable channel, expressed as a portion of its probability (line 11 of *Algorithm2*).

Algorithm2: Availability Probability Update

1: w: Number of licensed channels.
2: *L*,*b*: Speed of the automaton convergence, $L,b \in (0,1)$.
3: $p_i(1) = \frac{1}{w}, i = 1,2,\dots,w.$
4: **For** each timeslot $t = 1,2,\dots,q$
5: $V(t) = \{V_1(t), V_2(t), \dots, V_s(t)\}, s \leq w$. $V(t)$ ← The set of licensed channels sensed available.
6: **For** each channel $i = 1,2,\dots,w$
7: **If** $F_i(t) == true$
8: $p_i(t+1) = p_i(t) + \frac{L\cdot\sum_{y\notin V(t)}(p_y(t)-b)}{s}$.
9: **Else**
10: $p_i(t+1) = p_i(t) - L\cdot(p_i(t)-b)$.
11: **End_If**
12: **End_For**
13: **End_For**

E. Main Operation

The operation of the proposed adaptive method is given in this sub-section. *Algorithm3* describes the Adaptive Sensing Policy (ASP) and *Algorithm4* provides the steps of the Adaptive Negotiation-based Sensing Policy (ANSP). According to ASP, each secondary user selects a licensed channel to sense based on the corresponding availability probability vector (line 8). The probability vector is normalized, so as the summation of the availability probabilities is equal to 1. In the sequel, a generated random number in [0,1) produces the selected licensed channel. This step is introduced in order to avoid having all secondary users select the same licensed channel to sense. The update of the probability vector takes place at the end of the reporting phase, upon receiving the feedback vector in accordance with *Algorithm2*.

Algorithm3: Adaptive Sensing Policy

1: n: Number of secondary users.
2: w: Number of licensed channels.
3: $p(t)$ ← Availability probability vector.
4: $F(t)$ ← Feedback vector.
5: **For** each cognitive cycle
6: **For** each secondary user
7: //Reporting Phase//
8: Choose a licensed channel based on $p(t)$.
9: $j = 1,2,\dots,w$ ← the selected channel to sense
10: Send a beacon at *j*-th mini-slot if the channel sensed is available.
11: Receive the feedback and update the availability probability vector.
12: //Negotiating Phase//
13: Employ RTS/CTS messages to enable communication.
14: **End_For**
15: **End_For**

The ANSP is more efficient than ASP, since, during the negotiation phase, secondary users are enabled to overhear the RTS sender and change their channel selection in case they selected for data transmission the same channel as the sender (lines 15-22). This is accomplished through the usage of the special byte in the RTS/CTS messages. Specifically, if a secondary user perceives that possesses the same channel selection with the RTS sender (line 17), instantly chooses a different channel from the set channels that returned a negative feedback during the reporting phase (lines 18-19).

Algorithm4: Adaptive Negotiation-based Sensing Policy

1: n: Number of secondary users.
2: w: Number of licensed channels.
3: $p(t)$ ← Availability probability vector.
4: $F(t)$ ← Feedback vector.
5: **For** each cognitive cycle
6: **For** each secondary user
7: //Reporting Phase//
8: $j = 1,2,\dots,w$ ← the selected channel to sense
9: Send a beacon at *j*-th mini-slot if the channel sensed available.
10: Receive the feedback and update the availability probability vector.
11: //Negotiating Phase//
12: Overhear the RTS/CTS message exchange and determine the channel selected by the sender.

$z = 1,2, \ldots, w$ ← the sender's channel selection.
If $j == z$
$R(t)$ ← the set of licensed channels having not true feedback.
Choose a different channel from the $R(t)$ set and name it j.
Else
Keep the same channel selection
End_If
End_For
End_For

V. PERFORMANCE EVALUATION

In order to evaluate the performance of the proposed schemes, a simulation model was developed using Matlab. As a first step, the accuracy of the proposed adaptive method was assessed. In this way, the automaton ability on accurately estimating the channel availability probabilities was examined. Second, the proposed ASP was compared against the random-based RSP [4] in order to provide evidence of the performance improvements introduced due to the adaptive nature of our proposed policy. Finally, assuming that the usage of the special byte in RTS/CTS messages is feasible, informing thus the secondary users about the channel selected for data transmission by the RTS sender, the proposed ANSP was compared against NSP [4], so as to infer the efficiency of the proposed method.

Three main scenarios were considered. Specifically, in the first scenario the assumed wireless system consisted of three licensed channels having availability probabilities 0.7, 0.4, and 0.1. The simulation took place for 10000 timeslots. In this scenario, it was assumed that the special byte into the RTS/CTS messages, indicating the licensed channel selected for data transmission by the RTS sender, is not used, so the secondary users are not able to know the sensing intention of each sender. Regarding the automaton configuration, the parameter L was set to 0.01, holding a typical value of convergence speed [3], while the parameter b was chosen equal to 10^{-5}. Figures 2 and 3 illustrate the learning accuracy of the enhanced automaton when the number of secondary users is 10 and 20 respectively. In essence, figures show the automaton performance on estimating the availability probabilities of each licensed channel.

Undoubtedly, as illustrated in the figures, the automaton creates a quite adequate image of the channel availability probabilities, since the error rate is fairly limited. In general the error rate is below 5%. Moreover, it is observed that the error level of the automaton tends to increase as the actual availability probability decreases. The rationale behind this lies in the way the reporting phase is operating. In particular, a channel that is often active, i.e., its availability probability is low, produces a small amount of beacons by the secondary users during its corresponding mini-slot. Hence, the automaton receives a questionable feedback regarding this channel, causing a false probability update. However, the operation of the automaton is deemed as accurate, since the noticed error rate is limited.

In the second scenario, the proposed method is evaluated

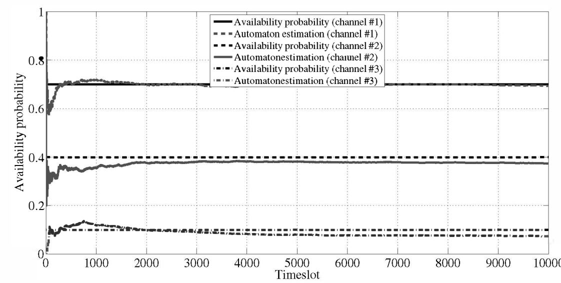

Figure 2. Learning accuracy of the automaton with 10 users.

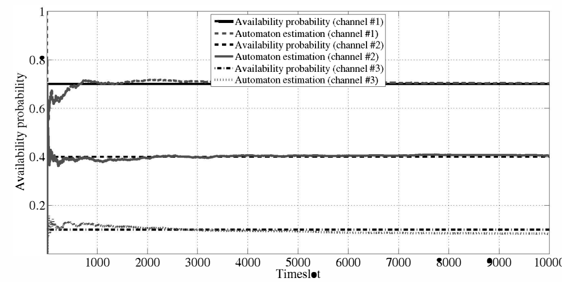

Figure 3. Learning accuracy of the automaton with 20 users.

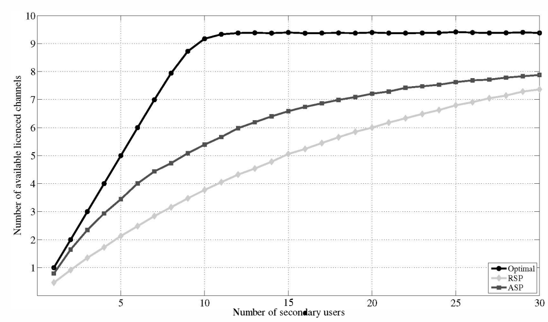

Figure 4. ASP performance with 20 licensed channels.

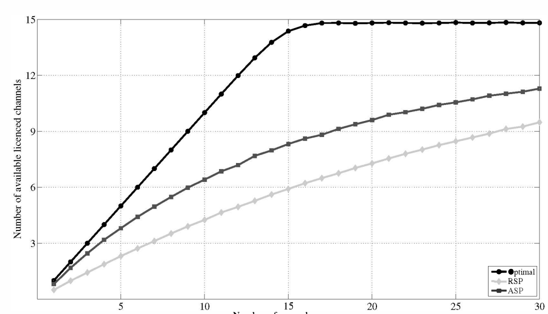

Figure 5. ASP performance with 30 licensed channels.

in comparison with the RSP scheme. The number of participating secondary users was varied from 1 to 30, while the number of licensed channels was set 20 and 30, as Fig. 4 and 5 depict. Again, the simulation lasted for 10000 timeslots. The available probabilities were chosen uniformly and independently. In this scenario the number of available

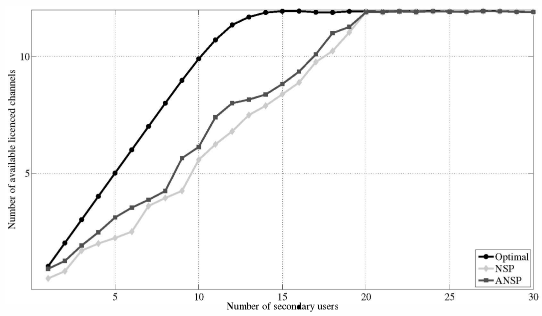

Figure 6. ANSP performance with 20 licensed channels.

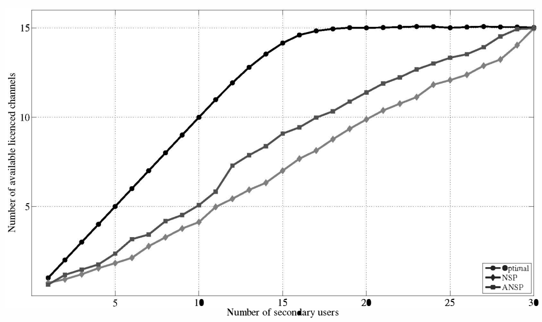

Figure 7. ANSP performance with 30 licensed channels.

channels sensed by the applied policy is examined, compared against the determined optimal channel number in accordance with *Algortihm1*. The performance of the proposed scheme is verified in both figures; the automaton acts beneficially, offering more channels sensed available than the RSP scheme, independently of the number of the connected

secondary users. The improvements reach almost 30% (i.e., 30% more available channels are identified), a critical advantage that potentially leads to 30% more efficient spectrum exploitation. As expected, the performance of both schemes is improved as the number of secondary users exceeds the number of licensed channels, a fact that reflects the wireless architecture: the more the users the more the sensed available channels.

In the third scenario, we assess the proposed scheme in comparison with the NSP scheme. The operation of both schemes entails the usage of the special byte into the RTS/CTS messages. In this case, the secondary users are aware of the channel selected for data transmission by the sender of the RTS message, thus, the optimal state could be reached after a period of time, if the number of secondary users is equal to or larger than the number of licensed channels. During the negotiation phase it is assumed that a single secondary user, uniformly selected, gains access to send data packets. The available probabilities were chosen uniformly and independently, while the simulation lasted for 10000 timeslots. In Fig. 6 the cognitive system consists of 20 channels, while in Fig, 7 it includes 30 channels. As it may be observed, a) both schemes converge to the optimal value when the number of users exceeds the number of licensed channels, and b) the adaptive scheme offers the asset of identifying more available channels than the NSP scheme when the number of users is less than the number of channels. This merit, stemming from the adaptive nature of the LA, allows for a more effective exploitation of the available spectrum.

VI. CONCLUSIONS

We designed and evaluated a novel adaptive sensing method for cognitive radio wireless networks. The novelty of this work lies in the enhancement of sensing policies with learning from experience concepts. By incorporating a learning automaton, users that seek opportunities for accessing the existing licensed channels operate more efficiently compared to random-based strategies. Having conducted extensive experiments, we demonstrated the performance of the proposed schemes. Specifically, a) the accuracy of the proposed method was identified, and b) considerable improvements are provided in terms of the number of available licensed channels sensed.

REFERENCES

[1] Q. Zhao, B.M. Sadler, "A survey of dynamic spectrum access," IEEE Signal Processing Mag., vol. 24, no. 3, pp. 79-89, May, 2007.

[2] L. Cao, H. Zheng, "Balancing Reliability and Utilization in Dynamic Spectrum Access," IEEE/ACM Transactions on Networking, vol.20, no.3, pp.651-661, June 2012.

[3] P. Nicopolitidis, G. I. Papadimitriou, A. S. Pomportsis, P. Sarigiannidis, M. S. Obaidat, "Adaptive wireless networks using learning automata," IEEE Wireless Communications, vol.18, no.2, pp.75-81, April 2011.

[4] H. Su, X. Zhang, "Cross-Layer Based Opportunistic MAC Protocols for QoS Provisionings Over Cognitive Radio Wireless Networks," IEEE Journal on Selected Areas in Communications, vol.26, no.1, pp.118-129, Jan. 2008.

[5] A. Khattab, D. Perkins, M. A. Bayoumi, "Experimental evaluation of Opportunistic Spectrum Access in distributed cognitive radio networks", In proc. of 8th Intrnational Wireless Communications and Mobile Computing Conference (IWCMC), pp. 245-250, 2012.

[6] Y. R. Kondareddy, P. Agrawal, "Synchronized MAC Protocol For Multi-Hop Cognitive Radio Networks," In proc. of IEEE International Conference on Communications (ICC), pp.3198-3202, 19-23 May 2008.

[7] Le Thanh Tan, Long Bao Le, "Channel Assignment With Access Contention Resolution for Cognitive Radio Networks," IEEE Transactions on Vehicular Technology, vol.61, no.6, pp.2808-2823, July 2012.

[8] Jing Zhang, Lina Qi, Hongbo Zhu, "Optimization of MAC Frame Structure for Opportunistic Spectrum Access," IEEE Transactions on Wireless Communications, , vol.11, no.6, pp.2036-2045, June 2012.

[9] N. Shah, D. Habibi, I. Ahmad, "Multichannel Cognitive Medium Access Control Protocol for Vehicular Ad-Hoc Networks", In Proc. of IEEE Vehicular Technology Conference (VTC), pp. 1-5, 2012.

[10] Gang Uk Hwang, S. Roy, "Design and Analysis of Optimal Random Access Policies in Cognitive Radio Networks," IEEE Transactions on Communications, vol.60, no.1, pp.121-131, January 2012.

[11] B. J. Oommen, "Recent advances in Learning Automata systems," 2nd International Conference on Computer Engineering and Technology (ICCET 2010), pp.V1-724-V1-735, 16-18 April 2010.