

# An Agent-based Model of the Risk-based Spectrum Auction in the Cognitive Radio Networks

Ján Pastirčák\*, Lukáš Sendrei†, Stanislav Marchevský‡ and Juraj Gazda§

Department of Electronics and Multimedia Communications

Technical University of Košice

Košice, Slovak Republic

Email: \*jan.pastircak@tuke.sk, †lukas.sendrei@tuke.sk, ‡stanislav.marchevsky@tuke.sk, §juraj.gazda@tuke.sk

**Abstract**—In this paper, we propose an agent-based model for spectrum trading in the shared use model of dynamic spectrum access. Spectrum trading is employed using the single-unit sealed-bid first-price auction, which takes into the account risk due to the imperfect spectrum sensing. Bidding strategies of the bidder are controlled by the reinforcement learning algorithm. We consider cooperative energy-based spectrum sensing as a spectrum sensing mechanism. Two different decision fusion strategies, which provide different levels of risk are discussed. The results demonstrate that in risky environment, total revenue and total payoff of the auctioneer and bidder respectively is higher, than in the case of system with lower level of risk. On the other hand, normalized revenue and payoff per a single auction round is higher in the case with lower level of risk. Moreover, the results have shown that the optimum sensing time for maximizing revenue and payoff is different.

**Keywords**—Spectrum trading, imperfect sensing, agent-based modeling

## I. INTRODUCTION

A lot of attention in the area of telecommunications has been recently dedicated to the cognitive radio (CR), considering this technology as the promising solution for further increase in spectrum utilization efficiency. CR is an underlying platform for the dynamic spectrum access (DSA) networks, in which secondary access to the radio spectrum is allowed if the spectrum is not utilized by the primary user (PU). Unused spectrum resources of the PU can be utilized by the secondary users (SU). The major restriction posed on the secondary spectrum access in DSA consists in the fact that the transmission of SUs does not cause any measurable interference to the PU [1]. Therefore, SUs must identify possible free and not occupied space-time-frequency slots where the transmission of the PU is not present [2]. The most prevailing method for the identification of the spectrum holes is spectrum sensing, which can be used in the *shared use* model of DSA [3]. In spectrum sensing, SU analyzes its radio environment, creates a test statistics from the received signal and based on this statistics, decides if the transmission of the PU is present in the received signal. One can find several types of sensing methods in the literature [4], [5], e.g. energy detection or cyclostationary features detection. Sensing results can be significantly improved by exploiting the spatial diversity of several SUs and making the final sensing decision based on the local sensing results. This is referred to as cooperative spectrum sensing [6].

Another main challenge in the area of DSA is the design

of the spectrum trading mechanisms. Spectrum trading can be referred to as a process of leasing or selling unoccupied spectrum resources of the PU to the SU [7]. Several auction-based mechanisms for spectrum trading were proposed in [8]–[11]. Spectrum trading in the *dynamic exclusive use* model of DSA [3] is discussed [8], [9]. In this model, the SUs have information about the channel occupancy directly from the PUs, therefore spectrum sensing is not necessary and spectrum trading algorithms do not address the problem of risk caused by incorrect channel occupancy identification. On the other hand, shared model of DSA is exposed to the performance degradation due to the imperfect spectrum sensing, resulting in the occurrence of risk in the spectrum auction process. Here, the term "risk" reflects the probability of the missed detection of the PU. Thus, with the increased level of risk, the SU decision regarding the valuation of the channel is reduced. There are only several papers dealing with this critical issue, for the reference please see e.g. [10], [11].

In this paper, we propose an agent-based model of spectrum trading for shared use model of DSA. We consider primary network consisting of a PU operating in TV band and secondary network consisting of secondary base transceiver station (BTS) and multiple SUs. SUs and secondary BTS perform cooperative spectrum sensing, in which secondary BTS makes a global sensing decision based on the local results from SUs. We consider two strategies for the decision fusion, where each creates different level of risk.. As a spectrum trading mechanism, we propose a single-unit sealed-bid first-price auction, where the bidding strategies of the SUs are controlled by the reinforcement learning algorithm [12]. Analysis of the total and normalized revenue and payoff is provided for both decision fusion strategies. Our results demonstrate that higher level of risk can provide higher total revenue of the auctioneer caused by more auction opportunities, but on the other hand, normalized (per-auction round) revenue is lower. The existing models do not take into the account the risk caused by imperfect spectrum sensing and commonly they operate with the global knowledge about the environment when optimizing the system. In our case, the presented agent based model aims to analyze and investigate the impact of the imperfect spectrum sensing determined by the duration of the spectrum sensing itself. Moreover, the strategies of the SUs are governed by the reinforcement learning and the knowledge of SUs is limited to the instantaneous payoff obtained in each auction round, thus no global environment knowledge is required. Last, the main advantage of the agent-based approach is that it can efficiently simulate the real-world scenario, where the agents (e.g., radio

nodes) have only limited knowledge about their environment.

The rest of the paper is organized as follows. Section II describes the system model considered in the paper. The agent-based model and the description of each network entity is described in Section III. Numerical analysis of the proposed model is pointed out in the Section IV. Finally, conclusion is given in the last Section V.

## II. SYSTEM MODEL

We assume a Cognitive Radio Network (CRN) operating in TV White Spaces. The CRN is represented by a secondary BTS and  $N$  SUs. The network scenario is depicted in Fig. 1. Data transmission is preceded by the spectrum sensing action in order to determine the availability of the desired channel. The local sensing result is sent to the secondary BTS. Based on this information, the secondary BTS decides whether the channel is occupied by the PU or not. This decision is done by applying decision fusion. If the channel is identified as idle the spectrum trading process is triggered. On the other hand, if the channel is identified as busy, no action will be taken, and thus no transmission will take place. Spectrum sensing and the spectrum trading algorithm will be revealed in detail in the following subsections.

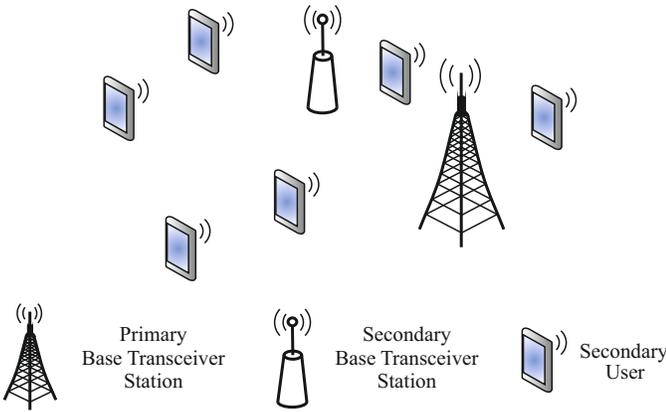


Fig. 1. Cognitive radio network considered in the paper

### A. Spectrum sensing

1) *General model:* In general, spectrum sensing is used to determine if a particular frequency channel is occupied by the PU transmission. Hence, we can establish two hypotheses: hypothesis  $\mathcal{H}_0$ , when no signal is transmitted and hypothesis  $\mathcal{H}_1$ , when a signal occupies the channel. It can be expressed mathematically as:

$$\mathcal{H}_0 : y(n) = w(n), \quad (1)$$

$$\mathcal{H}_1 : y(n) = s(n) + w(n), \quad (2)$$

where  $y(n)$  is the received signal,  $s(n)$  and  $w(n) \sim (0, \sigma_n^2)$  is the transmitted signal and noise, respectively.

Typically, to specify the channel vacancy, a test statistic  $T(y)$  from the received signal is formed. Then the final decision is made by comparing  $T(y)$  with a predefined threshold  $\epsilon$

under certain hypothesis which is expressed as:

$$T(y) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \epsilon \quad (3)$$

In accordance with (3), we can define four probabilities related to the spectrum sensing. Under hypothesis  $\mathcal{H}_0$  the probability of false alarm  $p_f = Pr(T(y) > \epsilon | \mathcal{H}_0)$  and the probability of correct detection that the channel is idle  $P_i = Pr(T(y) < \epsilon | \mathcal{H}_0)$ . Similarly, under hypothesis  $\mathcal{H}_1$ , the probability of detection  $p_d = Pr(T(y) > \epsilon | \mathcal{H}_1)$  and the probability of missed detection  $P_m = Pr(T(y) < \epsilon | \mathcal{H}_1)$ . Obviously, the goal of a good detection algorithm is to achieve high probability of detection  $p_d$  and low probability of false alarm  $p_f$ . The first probability is important for the protection of the primary system, whereas the second yields the missed opportunity for the secondary user to use the spectrum hole.

2) *Energy detection:* As it was already mentioned, various spectrum sensing techniques are available. We have chosen energy detection as the sensing method in this paper, as it is the simplest one to implement.

Now, let us define  $\tau$  as the sensing time,  $f_s$  as the sampling frequency and  $L$  as number of samples for the sensing (Let us assume  $L = \tau f_s$ ). The test statistic in this case will be written as:

$$T(y) = \frac{1}{L} \sum_{n=1}^L |y(n)|^2. \quad (4)$$

According to [13], assuming the hypothesis  $\mathcal{H}_0$  and sufficiently large  $L$ , the probability density function (PDF) of  $T(y)$  can be approximated by a Gaussian distribution with mean  $\mu_0 = \sigma_n^2$  and variance  $\sigma_0^2 = \frac{1}{L} [E|w(n)|^4 - \sigma_n^4]$ , where  $E|\cdot|$  is the expected value operator. Providing that  $w(n)$  is circularly symmetric complex Gaussian (CSCG) variable, then  $E|w(n)|^4 = 2\sigma_n^4$  and consequently  $\sigma_0^2 = \frac{1}{L}\sigma_n^4$ . Based on this presumptions the probability of false alarm can be defined as:

$$p_f = \mathcal{Q}\left(\left(\frac{\epsilon}{\sigma_n^2} - 1\right)\sqrt{\tau f_s}\right), \quad (5)$$

where  $\mathcal{Q}(\cdot)$  is the Q-function defined as:

$$\mathcal{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_1^{\infty} \exp\left(-\frac{t^2}{2}\right) dx. \quad (6)$$

Further, under the hypothesis  $\mathcal{H}_1$  and for sufficiently large  $L$ , the PDF of  $T(y)$  can be approximated by a Gaussian distribution with mean  $\mu_1 = (\gamma + 1)\sigma_n^2$ , where  $\gamma = \frac{\sigma_s^2}{\sigma_n^2}$  is the Signal-to-Noise (SNR) ratio, and variance

$$\sigma_1^2 = \frac{1}{L} [E|s(n)|^4 + E|w(n)|^4 - (\sigma_s^4 - \sigma_n^4)^2], \quad (7)$$

if  $s(n)$  and  $w(n)$  are both circularly symmetric and complex valued. If  $s(n)$  is complex PSK modulated and  $w(n)$  is CSCG noise, then  $\sigma_1^2 = \frac{1}{L}(2\gamma + 1)\sigma_n^4$ . For the case of PSK modulated complex-valued signal and CSCG noise, the probability of detection can be approximated as:

$$p_d = \mathcal{Q}\left(\left(\frac{\epsilon}{\sigma_n^2} - \gamma - 1\right)\sqrt{\frac{\tau f_s}{2\gamma + 1}}\right). \quad (8)$$

Sensing threshold  $\epsilon$  can be calculated by evaluating (5) or (8) for given target probabilities  $\bar{p}_f$  or  $\bar{p}_d$  respectively. Note, the

proof of the above mentioned presumptions and relations is based on the central limit theorem and detailed information can be found in [13].

3) *Decision fusion*: Let us define  $p_d^{(k)}$  as the probability of detection of the  $k$ -th secondary user. This local information, and also the local binary decision of the channel state  $D_k \in \{0, 1\}$ , is produced in each SU individually and sent to the secondary BTS in each sensing period. The final decision of the channel availability is made by the secondary BTS by using decision fusion. We assume two different decision fusion strategies.

- The first is the "Logic-OR" Rule. In this case the final decision says that the channel is occupied by a primary signal when one of the local decision declares the channel as busy. Mathematically can be this relation written as  $\Lambda = \sum_{k=1}^N D_k$ . If  $\Lambda \geq 1$  then the channel is busy. The final probability of detection, assuming that all local decisions are independent, is given by:

$$Q_d = 1 - \prod_{k=1}^N (1 - p_d^{(k)}). \quad (9)$$

- The second is the "Logic-AND" Rule. For this rule, if the all the local decisions says that the channel is occupied by a primary signal then the final decision declares the channel as busy. Mathematically can be this relation written as  $\Lambda = \prod_{k=1}^N D_k$ . If  $\Lambda = 1$  then the channel is busy. Once again, assuming independent local decisions, the final probability of detection can be calculated as:

$$Q_d = \prod_{k=1}^N p_d^{(k)}. \quad (10)$$

### B. Proposed auction algorithm for spectrum trading

We propose a spectrum trading algorithm based on the sealed-bid first-price auction. In this type of auction, the highest bidder wins and pays the amount he/she bids. In general, this auction type can be carried out either in sequential or concurrent version, but in our specific case, where only single unit of the spectrum is traded in each auction round, both sequential and concurrent version of the auction provides the same results<sup>1</sup>. In our model, BTS has a role of an auctioneer and SUs send their bids to BTS in order to purchase spectrum resources. The auction is being held if the channel is identified as idle. We define and instantaneous payoff of the  $i$ -th SU in the  $t$ -th time frame  $P(i, t)$  as:

$$R(i, t) = \begin{cases} V(i, t) - b(i, t) : \text{SU wins the auction} \\ 0 : \text{SU loses the auction} \end{cases} \quad (11)$$

where  $V(i, t)$  is a spectrum valuation of the  $i$ -th bidder in the  $t$ -th time frame and  $b_{i,t}$  is a bid of the  $i$ -th bidder in the  $t$ -th time frame. We define  $V(i, t)$  as:

$$V(i, t) = \omega[1 - r(t)](T - \tau) \frac{C(i, t)}{C_{ref}}, \quad (12)$$

<sup>1</sup>In the scenario, where bidder is constrained to obtain only single unit of spectrum from the multi-unit spectrum pool in each auction round, sequential auction provides better revenue, than the concurrent auction [8].

where  $r(t)$  is the risk connected with purchasing the spectrum in the  $t$ -th time frame,  $T$  is the duration of the time frame,  $\tau$  is the sensing time,  $C(i, t)$  is the channel capacity between the  $i$ -th SU and BTS in the  $t$ -th time frame,  $C_{ref}$  is the reference channel capacity and  $\omega$  is a scaling parameter. The probability that the  $i$ -th bidder wins the auction can be expressed as:

$$P(i\text{-th bidder winning}) = \left( \frac{b(i) - b_{min}}{V_{max} - b_{min}} \right)^{(N-1)}, \quad (13)$$

where  $b_{min}$  is the minimum bid,  $V_{max}$  is the maximum spectrum valuation and  $N$  is the number of bidders. The expected payoff of the  $i$ -th bidder can be then expressed as:

$$E(i) = (V(i) - b(i)) \times P(i\text{-th bidder winning}) \quad (14)$$

Optimal bid can be determined by substituting  $P(i\text{-th bidder winning})$  from (13) into (14) and maximizing the resulting  $E(i)$  by equating the first derivative of  $E(i)$  to 0. Thus, the optimal bid  $b(i)^*$  for the  $i$ -th bidder can be expressed as:

$$b(i)^* = \frac{(n-1)V(i) + b_{min}}{n} \quad (15)$$

One can easily observe that the bidder needs an information about the minimum bid in order to determine its optimal bid and this is usually not possible in the real scenario. Therefore, we propose an algorithm based on the reinforcement learning in order to determine the optimal bid locally for each bidder. The auction process works as follows:

- 1) BTS (auctioneer) notifies SUs that the auction is being held in the current time frame
- 2) SUs send their bids to the BTS
- 3) BTS determines the winner (SU with the highest bid) and sends auction results to the SUs
- 4) SUs calculate their instantaneous payoff and adapt their bids for the next auction round

1) *Learning process of the SU*: In order to determine the SU decision about their bids in each auction round, we have employed the reinforcement learning algorithm. This algorithm dictates the SUs to choose the bid, which would maximize their longterm payoff. Let us define a vector of the available bids for the  $i$ -th bidder in the  $t$ -th time frame  $\mathbf{b}(i, t)$  with elements  $b_j(i, t)$ , which can be expressed as:

$$b_j(i, t) = V(i, t) \frac{j}{M}, \quad (16)$$

where  $j = 1, 2, \dots, M$  and  $M$  is the number of the possible bid options. Each bid has its corresponding selection probability  $p_{i,t,j}$ . For the sake of the notation simplicity, we ignore the index  $i$  in the following expressions, but it is important to mention that these quantities are unique for each bidder. Bid valuation adaptation for the  $j$ -th bid choice can be expressed as:

$$\pi_j(t+1) = \pi_j(t) + \alpha(\rho(t) - \pi(t)), \quad (17)$$

where  $\pi_j(t)$  is a valuation of the  $j$ -th bid choice in the current time frame,  $\pi_j(t+1)$  is a valuation of the  $j$ -th bid choice in the next time frame,  $\rho(t)$  is the reference payoff and  $\alpha$  is a adaptation parameter. Adaptation of the  $\rho(t)$  can be expressed as:

$$\rho(t+1) = \rho(t) + \beta(R(t) - \rho(t)), \quad (18)$$

where  $\rho(t)$  is the reference payoff in the current time frame,  $\rho(t+1)$  is the reference payoff in the next time frame and  $\beta$  is the adaptation parameter. Probability that the  $j$ -th bid will be selected in the  $t+1$ -th time frame is calculated as follows:

$$p_j(t+1) = \frac{\exp(\frac{\pi_j(t+1)}{\gamma})}{\sum_{l=1}^M \exp(\frac{\pi_l(t+1)}{\gamma})}, \quad (19)$$

where  $\gamma$  is a parameter. The bid  $b(i, t+1)$  is selected randomly from  $\mathbf{b}(i, t)$  according to the vector of bid probabilities  $\mathbf{p}(i, t)$  with elements  $p_j(i, t)$ .

### III. AGENT-BASED IMPLEMENTATION OF THE MODEL

In order to analyze the impact of the risk caused by imperfect spectrum sensing on the spectrum trading, we propose an agent-based model consisting of three types of agents: single PU, single secondary BTS and multiple SUs. The main advantage of the agent-based implementation of the model is the possibility of studying very complex and dynamic scenarios, which is not possible to achieve using conventional simulation methods. Our goal is to analyze the revenue of the secondary BTS and the payoff of the SUs in the relation with the risk, caused by the imperfect spectrum sensing. This section provides the description of the agent types and their behaviour.

#### A. Primary user

PU is represented by a simple agent entity in our model. It can inhabit one of two states, which are directly related to its data traffic: *ACTIVE* or *IDLE*. Let us define the probability of the PU transmission  $p_{tx}$ , which is related to the transition probabilities between the PU states as shown in Fig. 2.

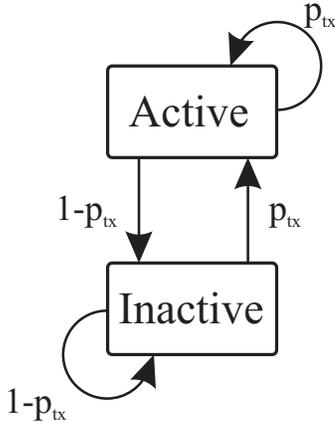


Fig. 2. Traffic model of the PU

#### B. Secondary BTS

Secondary BTS agent makes the global sensing decision based on the local decisions from the SUs, it calculates the risk  $r$ , which is connected with purchasing the spectrum band and it determines the winner of the auction. In each time frame, BTS waits for the local sensing results from each SU. The global

sensing decision is then made using either *OR* or *AND* rule as described in Section II-A3. The global probability of correct detection  $Q_d$  is calculated using (9) or (10) based on which decision fusion scheme is applied. If the channel is identified as idle, auction is being held, and risk  $r$  is calculated. Risk  $r$  represents in fact the global probability of missed detection and it can be expressed as:

$$r = 1 - Q_d \quad (20)$$

Risk  $r$  is sent to each SU over a dedicated channel and BTS waits for the bids from each SU. When all bids are collected, the winning SU is determined and all SUs are notified about the auction results. If the channel is identified as busy, auction is not being held in the current time frame.

#### C. Secondary user

SU agent performs four operations in each time frame:

- 1) Spectrum sensing
- 2) Participation in the auction (if the auction is being held)
- 3) Adaptation of its bidding choices using the learning algorithm
- 4) Data transmission (if SU wins the auction)

Spectrum sensing is performed as described in II. After sending its local sensing result to the BTS, SU waits for the global sensing decision from the BTS. If the channel is identified as idle, SU chooses bid  $b_j$  with the probability  $p_j$  and send it to the BTS. Based on the auction results, SU calculates its payoff from the current auction round according to (16) and (19). Then it adapts its bidding choices and their corresponding probabilities using the algorithm described in II-B1.

### IV. NUMERICAL RESULTS

In this section, we provide Monte-Carlo simulation results of the agent-based model described in the previous section. Based on these results, we analyze the impact of the sensing time on risk, revenue of the BTS and average payoff of the SU. We assume 1 PU, 1 BTS and 5 SUs located in the area with the dimensions of  $200 \times 200$ m and SUs moving randomly with the velocity  $v_{SU}$ . Moreover we assume that all channels (i.e., between the SUs and PU and between the SUs and BTS) to be the channels with just the path loss component affecting the attenuation. The simulation parameters are shown in Tab. I.

In Fig. 3 we illustrate risk  $r$  in terms of sensing time  $\tau$  for both *OR* and *AND* fusion rules. It can be observed that  $r$  is monotonically decreasing with the increasing  $\tau$  and  $r$  is in general higher, when *AND* decision fusion rule is applied. It can be explained using the fact that missed detection can occur more frequently when *AND* rule is applied, because each sensing node must identify the channel as busy in order to identify the channel as busy globally.

Fig. 4 illustrates the revenue of the secondary BTS in terms of sensing time  $\tau$ . It can be observed that the total revenue of the secondary BTS is around 20% higher when *AND* fusion rule is applied, than in the case with the *OR* fusion rule applied. However, by applying the *AND* fusion rule, the auction is being held in more time frames as by using the *OR* fusion rule

TABLE I. SIMULATION PARAMETERS

Parameter	Value
Number of iterations	10000
Number of time frames per iterations	100000
Time frame length	$T = 100\text{ms}$
Sampling frequency	$f_s = 200\text{kHz}$
Carrier frequency	$f_c = 800\text{MHz}$
Transmit power of the PU	$P_{tx,PU} = 60\text{dB}$
Transmit power of the BTS	$P_{tx,BTS} = 80\text{dB}$
Velocity of the SUs	$v_{SU} = 5\text{m/s}$
PUs probability of transmission	$p_{tx} = 0.5$
Variance of the noise on the SU receiver	$\sigma_n^2 = 1$
Target probability of false alarm	$\bar{p}_f = 0.1$
Parameter of the learning algorithm $\alpha$	$\alpha = 0.1$
Parameter of the learning algorithm $\beta$	$\beta = 0.1$
Parameter of the learning algorithm $\gamma$	$\gamma = 10$
Number of bid choices	$M = 10$
Reference channel capacity	$C_{ref} = 660\text{kbps}$

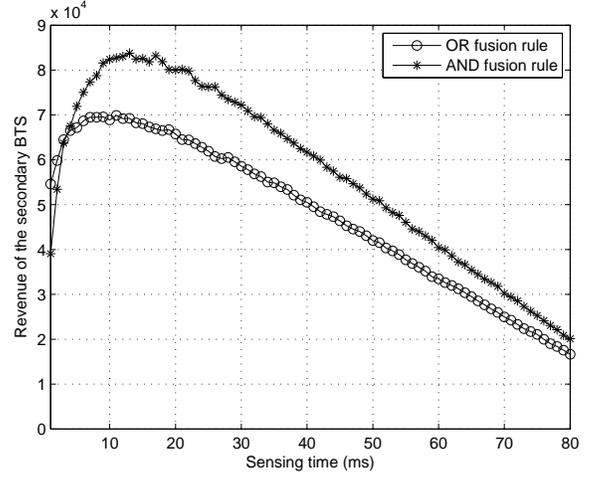


Fig. 4. Total revenue of the secondary BTS with respect to the sensing time

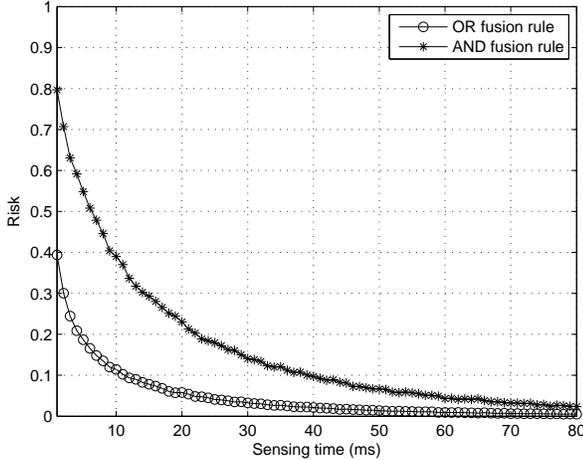


Fig. 3. Development of the risk vs. sensing time

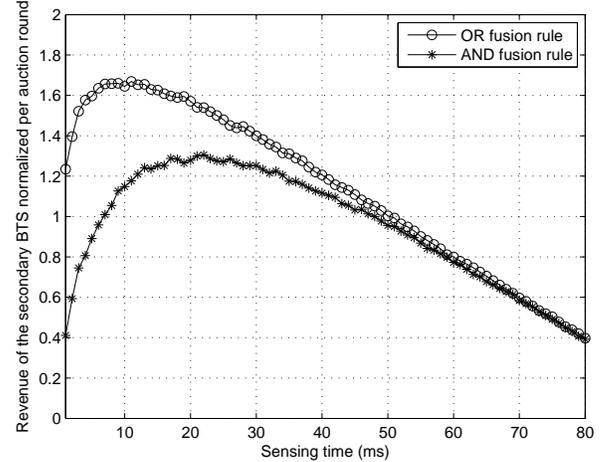


Fig. 5. Normalized revenue of the secondary BTS with respect to the sensing time

and this is the reason, why the AND fusion rule provides better total revenue for the auctioneer as the OR fusion rule. The optimum sensing time, which maximizes the total revenue is about 10 ms and 14 ms for OR rule and AND rule respectively. Fig. 5 illustrates the normalized revenue of the secondary BTS (i.e., total revenue divided by the number of time frames in which the auction was being held) and it can be observed that the OR fusion rule provides better normalized revenue, than the AND fusion rule. It can be explained in such way that the risk of purchasing the spectrum in the scenario with the OR fusion rule is lower, than in the case of the AND rule and therefore the SUs are willing to pay more for the same spectrum resource. The optimum sensing time, which maximizes the normalized revenue is about 10 ms and 20 ms for OR rule and AND rule respectively.

Fig. 6 illustrates the total average payoff of the SU (i.e., payoff of all SUs cumulated during the entire duration of one simulation run averaged by the number of SUs) in terms of

sensing time  $\tau$ . The results are similar to the results of the total revenue of the secondary BTS. By applying the AND fusion rule, the total average payoff of the SUs is higher (about 25% higher for  $\tau = 20\text{ms}$ ), than by applying the OR fusion rule, but this result doesn't take into account the number of time frames in which the auction was being held. As it is stated above, this number differs for AND and OR fusion rule. More detailed analysis can be provided by evaluating the normalized average payoff of the SU (i.e., total average payoff of the SU divided by the number of time frames, in which the auction was being held). These results are shown in Fig. 7. Similarly to the results of the normalized revenue, normalized average payoff of the SUs is higher in the case of OR rule applied. Optimum sensing time maximizing the total and normalized payoff is the same as the optimum time maximizing the total and normalized revenue respectively for both fusion rules.

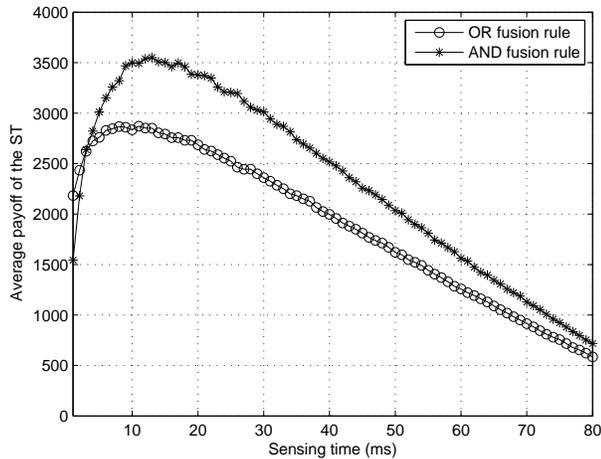


Fig. 6. Total average payoff of the SU with respect to the sensing time

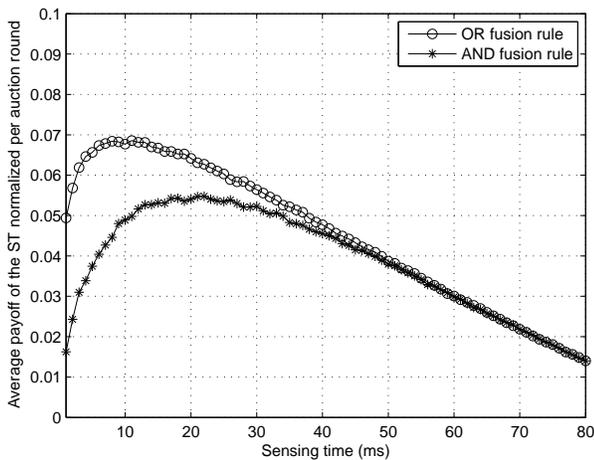


Fig. 7. Normalized average payoff of the SU with respect to the sensing time

## V. CONCLUSION

In this paper, we have considered spectrum trading mechanism for shared use model in cognitive radio networks. Spectrum trading is employed using the single-unit sealed-bid first-price auction, in which the bidding strategies of the individual SUs are controlled using the reinforcement learning algorithm. The proposed learning algorithm takes into account the risk of identifying the channel as idle, when it is in fact occupied by the PU. We have employed a simple energy-based spectrum sensing with the decision fusion of the local sensing result. The AND and OR fusion rules have been applied in the decision fusion process. As we have expected, the risk is higher when the AND fusion rule is applied. On the other hand, the employment of the OR rule provides less auction opportunities, which translates into lower total revenue of the secondary BTS and lower total average payoff of the SUs as well. But the evaluation of the normalized revenue and normalized average payoff have proved that by applying the OR rule provide better per-auction round results, than the

AND rule. Finally, the results have demonstrated that optimum sensing times maximizing revenue and payoff are the same but optimum sensing time depends to the level of risk present in the system.

For the sake of simplicity, only a single-channel system is considered, but it can be extended to a multi-channel case with possibly heterogeneous channels. The convergence properties of the proposed learning algorithm and its detailed parameters settings haven't been addressed and they remain to be an open issue for the future research.

## ACKNOWLEDGMENT

This work was supported by the Scientific Grant Agency of the Ministry of Education, science, research and sport of the Slovak Republic under the contract No. 1/0766/14.

This paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

We support research activities in Slovakia/This project is being co-financed by the European Union.

## REFERENCES

- [1] E. Hossain, D. Niyato, and Z. Han, *Dynamic Spectrum Access and Management in Cognitive Radio Networks*. Cambridge University Press, 2009.
- [2] R. Tandra, S. Mishra, and A. Sahai, "What is a spectrum hole and what does it take to recognize one?" *Proceedings of the IEEE*, vol. 97, no. 5, pp. 824–848, May 2009.
- [3] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access," *Signal Processing Magazine, IEEE*, vol. 24, no. 3, pp. 79–89, May 2007.
- [4] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *Communications Surveys Tutorials, IEEE*, vol. 11, no. 1, pp. 116–130, First 2009.
- [5] E. Axell, G. Leus, E. Larsson, and H. Poor, "Spectrum sensing for cognitive radio : State-of-the-art and recent advances," *Signal Processing Magazine, IEEE*, vol. 29, no. 3, pp. 101–116, May 2012.
- [6] K. Letaief and W. Zhang, "Cooperative communications for cognitive radio networks," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 878–893, May 2009.
- [7] D. Niyato and E. Hossain, "Spectrum trading in cognitive radio networks: A market-equilibrium-based approach," *Wireless Communications, IEEE*, vol. 15, no. 6, pp. 71–80, December 2008.
- [8] S. Sengupta and M. Chatterjee, "Designing auction mechanisms for dynamic spectrum access," *Mobile Networks and Applications*, vol. 13, no. 5, pp. 498–515, 2008. [Online]. Available: <http://dx.doi.org/10.1007/s11036-008-0081-1>
- [9] S. Gandhi, B. Chiranjeeb, L. Cao, H. Zheng, and S. Suri, "Towards real-time dynamic spectrum auctions," *Computer Networks*, vol. 52, no. 4, pp. 879–897, 2008.
- [10] T. Wang, J. Wu, and C. Long, "Flexibly cooperative spectrum sensing and allocation in cognitive radio networks: An auction-based framework," in *Computer Science and Network Technology (ICCSNT), 2011 International Conference on*, vol. 2, Dec 2011, pp. 1250–1254.
- [11] M. Tehrani and M. Uysal, "Auction based spectrum trading for cognitive radio networks," *Communications Letters, IEEE*, vol. 17, no. 6, pp. 1168–1171, June 2013.
- [12] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *Journal of Artificial Intelligence Research* 4, pp. 237–285, 1996.
- [13] Y.-C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 4, pp. 1326–1337, April 2008.