Secondary User Monitoring in Unslotted Cognitive Radio Networks with Unknown Models

Shanhe Yi[1], Kai Zeng[2], and Jing Xu[1]

[1] Department of Electronics and Information Engineering
Huazhong University of Science and Technology, Wuhan 430074, China
{yishanhe@smail,xujing@mail}.hust.edu.cn
[2] Department of Computer and Information Science
University of Michigan - Dearborn, Michigan 48128, USA
kzeng@umich.edu

Abstract. Cognitive radio networking (CRN) is a promising technology to improve the spectrum utilization by allowing secondary users (unlicensed users) to opportunistically access white space (spectrum holes) in licensed bands. Monitoring the detailed characteristics of an operational cognitive radio network is critical to many system administrative tasks. However, very limited work has been done in this area. In this paper, we study the passive secondary user monitoring problem in an unslotted cognitive radio network, where the users' traffic statistics are unknown in priori. We formulate the problem as a multi-armed bandit (MAB) problem with weighted virtual reward. We propose a dynamic sniffer-channel assignment policy to capture as much as interested user data. Simulation results show that the proposed policy can achieve a logarithmic regret with relative scalability.

Keywords: secondary user monitoring, cognitive radio networks, multi-armed bandit

1 Introduction

The proliferations of wireless communication and ever-increasing wireless traffic demand have put significant pressure on spectrum utilization. On the one hand, the unlicensed spectrum has become over crowded. On the other hand, a large portion of licensed bands are underutilized [1]. The un-balanced spectrum allocation and usage lead to the so-called "spectrum scarcity" problem. The concept of opportunistic spectrum access (OSA) has emerged as a way to dramatically improve spectrum utilization, thus solve this problem. In OSA, the unlicensed users can dynamically access to the licensed band (white space) without interfering the communication among primary users. Cognitive radio [2, 3] is a promising technology to realize OSA, where the secondary users (cognitive radios) can sense the absence or presence of the primary users and opportunistically access to or evacuate from the primary spectrum/channels. A variety of emerging applications, including smart grid, public safety, broadband cellular, and medical applications, are expected to be supported by OSA and cognitive radio networking [4].

Monitoring the detailed characteristics of an operational cognitive radio network is critical to many system administrative tasks, such as spectrum policy enforcement, wireless advisory, fault diagnosis, anomaly detection, attack detection, forensics, resource management, and critical path analysis for network upgrades. However, very limited work has been done in this area [5].

In this paper, we study the passive monitoring problem in cognitive radio networks. Our goal is to capture as much as interested secondary user data. Different from the monitoring in traditional wireless networks [6–8], monitoring a cognitive radio network faces unique challenges: 1) Secondary users' activities are unknown in priori. Due to the agility of the cognitive radio and the activity of the primary user, the secondary users may jump among different channels to seek the best spectrum and communication opportunity or stay in the same channel, lowering its transmission power level or alerting modulation scheme to avoid interference. Furthermore, different secondary users may have different traffic patterns. 2) Cognitive radio networks usually have much wider spectrum than traditional wireless networks. Due to the limitation of hardware technology, each sniffer can only monitor one channel at a time. It becomes difficult or infeasible to deploy a large amount of sniffers to monitor all the channels at all the times.

To solve the above two challenges, we need to learn the characteristics of the primary and secondary users' traffic, meanwhile dynamically assign the limited number of sniffers to the most profit channels where the concerned secondary users may reside in. There exists an interesting tradeoff between assigning sniffers to channels which are already known as the most beneficial based on the current knowledge, versus exploring the channels which are under-observed. This category of learning and decision making under uncertainty defined by a classical tradeoff between exploration and exploitation fits into the multi-armed bandit (MAB) framework [9].

In this paper, we study the secondary user monitoring problem in an unslotted cognitive radio networks without prior knowledge of the user traffic statistics. The objective is to maximize the expected captured data of interested secondary users. The problem appears to be much more complicated than a slotted system due to the arbitrary starting and ending times of the transmissions of primary or secondary users. The challenge comes from capturing data of interested secondary users coexisting with primary users and unconcerned secondary users in a highly dynamic wireless network. Sniffers are dynamically assigned to operate on different channels to perform data capturing. During the capture, a sniffer must make a decision whether to keep operating on the current channel or switch to another channel. We formulate the problem as a MAB problem with weighted virtual reward. We propose a dynamic sniffer-channel assignment policy to capture as much as interested user data. Simulation results show that the proposed policy can achieve a logarithmic regret with relative scalability. To the best of our knowledge, we are the first to work on the secondary user monitoring in unslotted cognitive radio networks with unknown models.

The rest of this paper is organized as follows. Related work is discussed in Section 2. Section 3 formulates the problem. Then we present our policy for optimal data capturing in Section 4. In Section 5, simulation results and performance analysis are presented. Finally, we conclude this paper and discuss future works in Section 6.

2 Related Work

2.1 Multichannel Wireless Network Monitoring

Most recently, Arora *et al.* [8] applied MAB to study the optimal sniffer-channel assignment (OSCA) problem in multichannel wireless networks. It considers *sniffer-centric* monitoring [7] that aims to monitor the busiest channels. The core difference between our work and the multichannel wireless network monitoring is that we differentiate secondary users from primary users. When a primary user is found in a channel, our sniffer will switch to other channels to capture the interested secondary user data.

2.2 Cognitive Radio Network Monitoring

Very limited work has been done on secondary user monitoring in cognitive radio networks. Chen *et al.* proposed a secondary user data capturing mechanism applying machine learning technology [5]. The basic idea is to estimate the packet arrival time of interested user data and reuse the sniffer in the time domain. A sniffer can switch to other channels to capture interested data and switch back if it has enough time to do so without missing the next interested packet in the current channel. Dedicated sniffers are used to explore interested packets on different channels. Different from [5], this paper more focuses on the strategy and decision-making study of sniffers to capture data efficiently. We formulate the problem as a MAB problem and no dedicated sniffers are used to sweep channels.

2.3 Opportunistic Spectrum Access

MAB framework has been adopted to study opportunistic spectrum access (OSA) in cognitive radio networks. Liu *et al.* applied UCB1 method proposed in [10] to single user-channel selection in [11]. Liu and Zhao formulated the second user spectrum access problem as decentralized MAB problem and gave logarithmic regret policies in [12,13]. These works are based on slotted systems. Recent work [14] applied MAB to an unslotted primary user system. Our work has fundamentally different objective from OSA. We aim to capture as much as interested secondary user data, while OSA aims to find the channels with least primary user activity or better spectrum access opportunity. Therefore, we cannot directly apply the OSA methods to secondary user monitoring. When a primary emerges, our sniffer will switch to other channels since we are only interested in secondary user data. We also need to learn the statistics of secondary user data in order to make decisions for sniffers, while OSA does not have this component and only concerns the statistics of primary users.

3 Problem Formulation

3.1 System Model

We consider a cognitive radio network with K channels. Each channel is used by one primary user. There are N secondary users and S sniffers. We are interested in the data of M secondary users. We have $M \leq N$ and $S < K$. Multiple secondary users can be in the same channel but they transmit at different times without collision. We assume each secondary user only stays in one channel and they will keep silent when the primary user shows up in the same channel. However, we do not know which secondary users stay in which channels or their traffic statistics in priori. Each sniffer can only monitor one channel at a time, but it can switch channels at any time. The sniffers are assumed to be able to identify the secondary and primary users by examining the packet header or signal feature.

The traffic of PUs is modeled as an on-off renewal process [15]. Similar to [16], Markovian assumption is relaxed on the primary traffic and we assume that the busy period may be arbitrarily distributed while the idle period has an exponential distribution. During the idle period of PU, we also model the traffic of SUs with different on-off renewal processes. Therefore the whole traffic on each channel is actually a combinational on-off renewal process with multi-states indicating which user is occupying the channel at the moment. Note that the channels are occupied temporally from primary users to secondary users. The states of each channel are firstly split into *PU-on* state and *PU-off* state where *on* means the PU is working in the channel and *off* means that PU is absent. Then, the *PU-off* state is split into separately *SU-on* state and *SU-off* state of different SUs since secondary user will use these idle channels for communication.

The occupancy duration (the *on* state) of PUs and SUs in every appearance follow the corresponding distributions denoted by the following random variables

$$
\begin{aligned}
Y_{pu} &= \{Y_{pu_1}, \cdots, Y_{pu_k} \cdots, Y_{pu_K}\} \\
Y_{su} &= \{Y_{su_1}, \cdots, Y_{pu_j} \cdots, Y_{su_N}\}
\end{aligned}
$$

where $k \in [1, K]$ and $j \in [1, N]$.

We denote the interested secondary user set as T_{su} which is a subset of all the secondary users. Fig. 1 gives an example of $T_{su} = \{su_1, su_3, su_4, su_5, su_7\}$ out of seven SUs ($M = 5, N = 7, K = 3$), where both su_3 and su_4 access to channel 2 and both su_5 and su_7 access to channel 3.

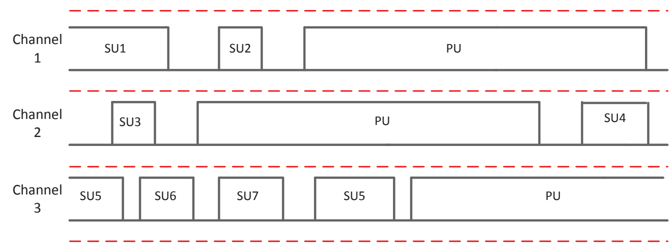

Fig. 1. Traffic Model

According to the statistical characteristics of network traffic, a categorical distribution is introduced to portray the appearance frequency of a set of secondary users in channel k with a probability set $\boldsymbol{P}_k$ which is defined as follows:

$$\boldsymbol{P}_k = \{p_1, \cdots, p_{j_k}, \cdots, p_{n_k}\}$$
$$\forall k \in [1, K], \sum_{j_k=1}^{n_k} p_{j_k} = 1$$
$$\sum_k n_k = N, k \in [1, K]$$

where n_k is the index of secondary users in channel k.

Since the information of appearance probability of SUs is not known in priori, we utilize the observation of SU appearance times to estimate the appearance probability. During the data capture, sniffers collect information to build up a statistical appearance probability of secondary users in channel k denoted by $\boldsymbol{P}_k^o$.

We assume a centralized system where a decision center is used to gather all the information collected from sniffers and make dynamic online sniffer-channel assignment decisions in order to capture interested secondary user data as much as possible.

3.2 MAB Problem Formulation

A typical Multi-Armed Bandit problem consists of a series of components including *player*, *arm*, *reward*, *regret* and *policy* [9, 17, 18]. By mapping sniffers to players and channels to arms, channel assignment of sniffers for data capture without prior knowledge falls into a multi-player multi-armed bandit problem naturally.

Since there is no prior information about secondary user traffic characteristics, sniffers have to identify and monitor the channels which have potential for the most beneficial reward to fulfill the goal of capturing transmission data of interested users as much as possible. To deal with this challenge, we keep recording the length of observed complete idle and busy period length of interested users appeared in channel k in vectors denoted by $\boldsymbol{x}_k^o$, $\boldsymbol{y}_k^o$, respectively. Both $\boldsymbol{x}_k^o$ and $\boldsymbol{y}_k^o$ are updated after each valid capture which succeeds to capture corresponding interested user data. Note that for the need to evaluate the accessibility of channel, the history of $\boldsymbol{x}_k^o$ and $\boldsymbol{y}_k^o$ are recorded. We also count the number of records and keep calculating the average of all the observation values of $\boldsymbol{x}_k^o$ and $\boldsymbol{y}_k^o$ and average them up to the current time (sample mean) which are denoted by $\hat{\boldsymbol{x}}_k^o$ and $\hat{\boldsymbol{y}}_k^o$, respectively.

The total captured data of sniffers is summed up by the length of every captured data of all the interested secondary users denoted by

$$V_o(t - t_0) = \sum_{n(t_0)}^{n(t)} \sum_k \sum_m y_{k,m}^o(n) \tag{1}$$

where $y_{k,m}^o$ is the element in $\boldsymbol{y}_k^o$, $k \in [1, K]$ is the index of channel and $m \in [1, M]$ is the index of interested user. t_0 and t are the start time and end time, respectively. $n(t)$ is the count number of $\boldsymbol{y}_k^o$ history from time t_0 to t.

Different interesting secondary users appear in different channels with different probabilities and their occupancy durations have different distributions. Take these two influential factors into consideration, a virtual reward of channel k is proposed as a weighted length of captured data which is given by

$$V_k = \sum_m \frac{\hat{y}_{m,k}^o}{\hat{y}_{m,k}^o + \hat{x}_{m,k}^o} p_{m,k}^o \tag{2}$$

where $\hat{y}_{m,k}^o$ and $\hat{x}_{m,k}^o$ is the average length of busy/idle period recorded in vector $\hat{\boldsymbol{y}}^o$ and $\hat{\boldsymbol{x}}^o$ and m is the index of interested secondary users. $p_{m,k}^o$ is the estimation of interested secondary user appearance probability which is the element in $\boldsymbol{P}_k^o$.

In order to measure the performance of strategies or policies dealing with MAB problems, *regret* is introduced as a key metric. If the reward model, user traffic parameters, and other useful prior knowledge are known, it is easy to infer that sniffers should always make the right decision to capture interested user data as much as possible by utilizing these prior knowledge. The total data captured by a "genie" is denoted by $V^*(T)$ which is similar to $V_o(T)$ in Eq. (1) but using prior knowledge. Thus the regret is given as a difference of expected value of gained reward between the "genie" and the proposed method:

$$\mathbb{E}V^*(T) - \mathbb{E}V_o(T)$$

4 Optimal data capture policy

In this section, we introduce an optimal strategy for sniffers to perform efficient data capture. Our proposed policy is centralized, which can be divided into two parts: monitor policy and decision policy.

4.1 Monitor Policy

We modify the sensing and transmission scheme for single secondary user in [14] into our monitor policy for multiple sniffers to fulfill our goal.

Similar to [14], when a sniffer is assigned to a given channel, it is required to keep working for at least two successive complete idle or busy period before leaving. However, there is an exception when a sniffer senses that the assigned channel is occupied by a primary user. Since the duration of a primary user occupation is usually longer than the secondary users' and we are interested in secondary user monitoring, when a sniffer encounters such a situation, it switches to another channel immediately to perform monitor policy again. Thus the utilization of sniffers is improved and the tensity of limited monitoring resource gets relieved to some extent. As mentioned above, during the busy period of interested users, sniffers capture the data, collect the complete periods information, and then send it to the decision center, which will decide if the sniffer should stay on the current channel or switch to another one according to the decision policy which will be proposed in Section 4.2.

As to the initialization, sniffers follow the same rule to capture data of each channel in sequence in order to generate initial estimations of unknown parameters. Initialization process with a single sniffer abiding the monitoring rule is illustrated in Fig. 2.

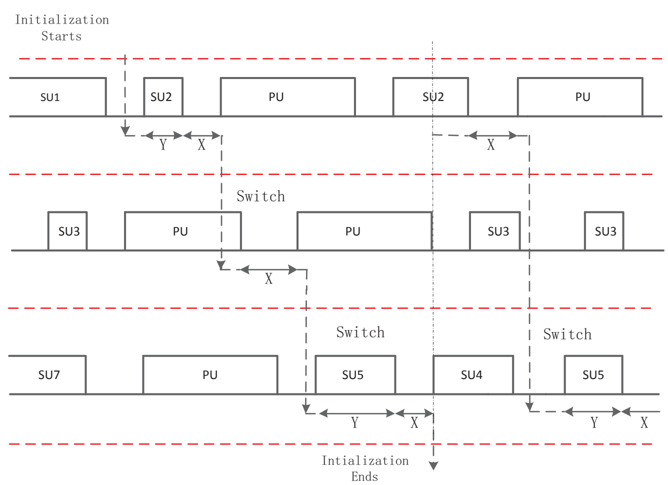

Fig. 2. Initialization and Monitor Rule

As shown in the figure, since the first idle/busy period is incomplete and the following user is an interested secondary user, three consecutive idle/busy periods are monitored of which the latter two are complete. Then the sniffer switches to the next channel to perform initialization monitoring. After finding the user in this channel is a primary user, the sniffer switches to the next channel immediately. When it comes to a multi-sniffer case, each sniffer chooses a random but different sequence of channels to conduct initialization separately as shown above. The initialization is finished after gaining information from every channel and followed by regular monitoring.

4.2 Decision Policy

We adopt the policy named UCB1 (Upper Confidence Bound) proposed in [10], which is able to achieve logarithmic regret for reward distribution with finite support. In UCB1, a decision index denoted by $I(t)$ is computed for each arm before making decisions. In our system model, $I(t)$ can be calculated from observations x^o, y^o and observed appearance probabilities $\hat{P}^o$. The index has the following form:

$$I_k(t) = \overline{V_k}(t) + \sqrt{\frac{2\log(t)}{\sum\limits_m n_{k,m}(t)}}$$

where $\overline{V_k}(t)$ is the sample mean of virtual reward defined in Eq. (2) up to time t and $\sum\limits_m n_{k,m}(t)$ is a counting number of successful capture of interested users in channel i.

The update rule follows

$$\overline{V_k}(t+\Delta t) = \begin{cases} \frac{\overline{V_k}(t)+V_k(t,t+\Delta t)}{n_{k,m}(t)+1} & \text{if SU m is captured at channel k;} \\ \overline{V_k}(t) & \text{else} \end{cases}$$

$$n_{k,m}(t+\Delta t) = \begin{cases} n_{k,m}(t)+1 & \text{if SU m is captured at channel k;} \\ n_{k,m}(t) & \text{else} \end{cases}$$

where $V_k(t, t+\Delta t)$ is the incremental virtual reward from time t to time $t+\Delta t$.

The index is updated after each valid capture and available channel(arm) k with the highest index is chosen for available sniffer:

$$k = \arg\max_k I_k(t)$$

As proved in [10], there exists an optimal logarithmic regret in UCB1 method.

5 Performance Analysis

In order to evaluate the performance of our proposed policy, we developed a simulator using MATLAB. We examine the performance under different number of channels, sniffers and users. In the simulation, the number of channels K is set from 4 to 6 and the number of interesting users M is set from 4 to 12. We consider both single sniffer and multiple sniffer cases. In all the cases, the busy/idle periods of users follow exponential distributions. The busy/idle periods of primary users are much longer than those of secondary users.Both $\mathbb{E}Y_{pu}^{on}$ and $\mathbb{E}Y_{pu}^{off}$ is set as 1. The busy periods of secondary users follow different exponential distributions while the idle periods follow the same. $\mathbb{E}Y_{su}^{on}$ is random selected from 0.2 to 0.4 and $\mathbb{E}Y_{su}^{off}$ is set as 0.1. The appearance probabilities of SUs in channels are also randomly generated. The simulation results shown in each figure are the averages over 200 runs.

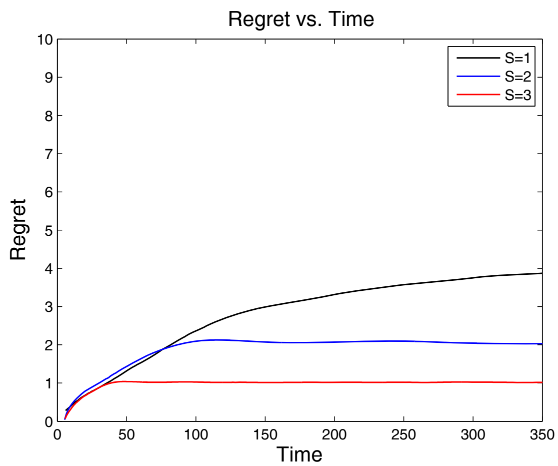

Fig. 3. Regret *vs.* Time (S=[1,2,3] M/N=8/12 K=4)

Fig. 3 gives the regret in which the logarithmic regret order of the proposed policy can be observed. Due to the monitor rule, as more sniffers participate in the monitoring, the information about the interested user traffic characteristics can be estimated more accurately and the regret converges more quickly. Therefore, the regret is diminished and converges faster when the number of sniffers increases.

Fig. 4 shows the proportion of captured interested user data under different number of sniffers. The proportion increases asymptotically which indicates that under the proposed policy, the data captured by the limited number of sniffers is able to catch up with the "genie" as long as the time goes on. The more sniffers we have, the higher ratio of data is captured.

Fig. 5 gives the reward of using different number of sniffers which is positively correlated to the number of sniffers More involved sniffers result in more data captured. The accumulated gain increases linearly against time.

In Fig. 6, we vary M, the number of interested users. The simulation result of *regret* shows that our proposed policy is relatively scalable against the number of interested secondary user. While halving or doubling the number of interested users, the increment or decrement percentage of regret is about 25%.

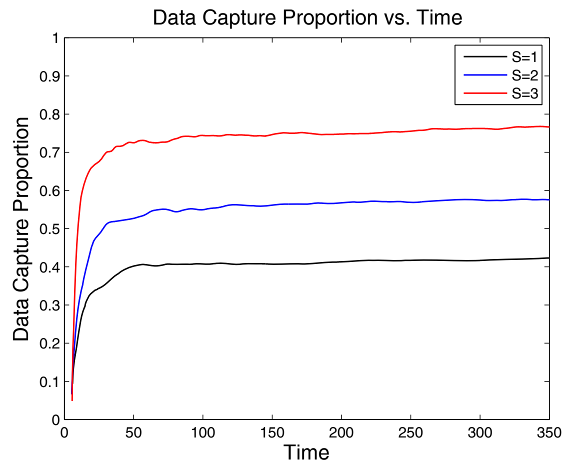

Fig. 4. Data Capture Proportion *vs.* Time (S=[1,2,3] M/N=8/12 K=4)

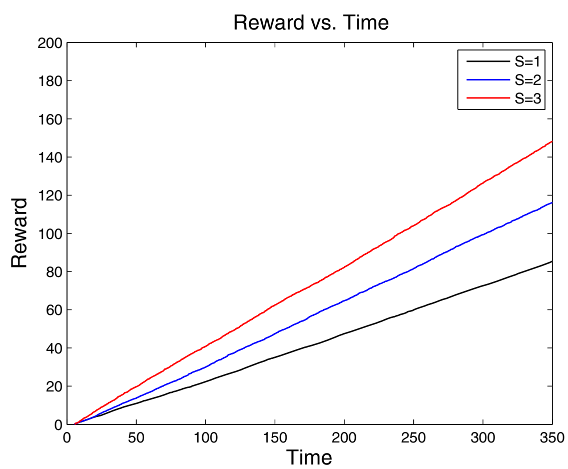

Fig. 5. Reward *vs.* Time (S=[1,2,3] M/N=8/12 K=4)

6 Conclusions and Future Work

In this paper, we study secondary user monitoring problem in unslotted cognitive radio networks with unknown user traffic statistics. We formulate the problem as a MAB problem with weighted virtual reward and propose a dynamic sniffer-channel assignment policy. Simulation results show that the proposed policy can achieve a logarithmic regret with relative scalability. Our future work will be to study the secondary user monitoring problem in more complicated scenarios by considering secondary user channel switching and its influence on our policy design.

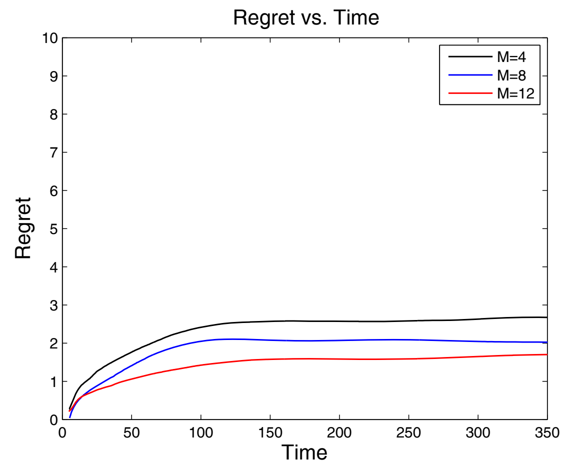

Fig. 6. Regret *vs.* Time (S=2 M/N=[4,8,12]/12 K=4)

7 Acknowledgement

This work was partially supported by the US NSF CAREER award under grant CNS-1149500 and National Key Technologies R&D Program of China 2011BAK08B01.

References

1. FCC: Unlicensed operations in the tv broadcast bands, second memorandum opinion and order. FCC 10-174 (Sept. 2010)
2. Haykin, S.: Cognitive radio: Brain-empowered wireless communications. IEEE JSAC **23**(2) (Feb. 2005) 201–220
3. Akyildiz, I.F., Lee, W., Vuran, M.C., Mohanty, S.: Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey. Computer Networks (Elsevier) **50**(13) (Sep 2006) 2127–2159
4. Wang, J., Ghosh, M., Challapali, K.: Emerging cognitive radio applications: A survey. Communications Magazine, IEEE **49**(3) (March 2011) 74 –81
5. Chen, S., Zeng, K., Mohapatra, P.: Efficient data capturing for network forensics in cognitive radio networks. In: Network Protocols (ICNP), 2011 19th IEEE International Conference on, IEEE (2011) 176–185
6. Yeo, J., Youssef, M., Agrawala, A.: A framework for wireless lan monitoring and its applications. In: Proceedings of the 3rd ACM workshop on Wireless security, ACM (2004) 70–79
7. Chhetri, A., Nguyen, H., Scalosub, G., Zheng, R.: On quality of monitoring for multi-channel wireless infrastructure networks. In: Proceedings of the eleventh ACM international symposium on Mobile ad hoc networking and computing, ACM (2010) 111–120

8. Arora, P., Szepesvari, C., Zheng, R.: Sequential learning for optimal monitoring of multi-channel wireless networks. In: INFOCOM, 2011 Proceedings IEEE, IEEE (2011) 1152–1160
9. Robbins, H.: Some aspects of the sequential design of experiments. Bulletin of the American Mathematical Society **58**(5) (1952) 527–535
10. Auer, P., Cesa-Bianchi, N., Fischer, P.: Finite-time analysis of the multiarmed bandit problem. Machine learning **47**(2) (2002) 235–256
11. Lai, L., El Gamal, H., Jiang, H., Poor, H.: Cognitive medium access: Exploration, exploitation, and competition. Mobile Computing, IEEE Transactions on **10**(2) (2011) 239–253
12. Liu, K., Zhao, Q., Krishnamachari, B.: Decentralized multi-armed bandit with imperfect observations. In: Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on, IEEE (2010) 1669–1674
13. Liu, K., Zhao, Q.: Distributed learning in multi-armed bandit with multiple players. Signal Processing, IEEE Transactions on **58**(11) (2010) 5667–5681
14. Tehrani, P., Zhao, Q., Tong, L.: Multi-channel opportunistic spectrum access in unslotted primary systems with unknown models. In: Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2011 4th IEEE International Workshop on, IEEE (2011) 157–160
15. Zhao, Q., Sadler, B.: A survey of dynamic spectrum access. Signal Processing Magazine, IEEE **24**(3) (2007) 79–89
16. Chen, S., Tong, L.: Low-complexity distributed spectrum sharing among multiple cognitive users. In: MILITARY COMMUNICATIONS CONFERENCE, 2010-MILCOM 2010, IEEE (2010) 2274–2279
17. Lai, T., Robbins, H.: Asymptotically efficient adaptive allocation rules. Advances in applied mathematics **6**(1) (1985) 4–22
18. Agrawal, R.: Sample mean based index policies with o (log n) regret for the multi-armed bandit problem. Advances in Applied Probability (1995) 1054–1078