

## Research Article

# A Two-Dimensional Multiarmed Bandit Approach to Secondary Users with Network Coding in Cognitive Radio Networks

**Yuben Qu, Chao Dong, Dawei Niu, Hai Wang, and Chang Tian**

*College of Communications Engineering, PLA University of Science and Technology, No. 88 Houbiaoying, Yudao Street, Qinhuai District, Nanjing, Jiangsu 210007, China*

Correspondence should be addressed to Yuben Qu; [quyuben@gmail.com](mailto:quyuben@gmail.com) and Hai Wang; [haiwang@ieee.org](mailto:haiwang@ieee.org)

Received 11 February 2015; Revised 8 July 2015; Accepted 14 July 2015

Academic Editor: Francisco Chicano

Copyright © 2015 Yuben Qu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We study how to utilize network coding to improve the throughput of secondary users (SUs) in cognitive radio networks (CRNs) when the channel quality is unavailable at SUs. We use a two-dimensional multiarmed bandit (MAB) approach to solve the problem of SUs with network coding under unknown channel quality in CRNs. We analytically prove the asymptotical-throughput optimality of the proposed two-dimensional MAB algorithm. Simulation results show that our proposed algorithm achieves comparable throughput performance, compared to both the theoretical upper bound and the scheme assuming known channel quality information.

## 1. Introduction

With the fast development of the technology of software-defined radios [1], cognitive radio networks (CRNs) emerge to solve the spectrum underutilization [2] problem. In CRNs, secondary users (SUs) are allowed to simultaneously share the licensed spectrum with primary users (PUs), when it is not occupied by PUs. However, due to the strict low priority, the idle duration of the PU channel (also available transmission time for SUs) is actually not just limited, but uncertain in the sense that SUs' transmissions may be interrupted by PUs at any time. Therefore, how to improve the transmission quality via increasing spectrum utilization under such uncertainty is urgently needed.

Recently, network coding [3] has emerged to an efficient technology to improve SUs' performance in terms of throughput in CRNs, by employing network coding in SUs' data transmissions [4–8], and so forth. The key idea of network coding is to combine multiple input packets into one packet algebraically before forwarding [3]. The number of packets combined together is named as block size, which is a key coding parameter in network coding literatures. It is worth emphasizing that the block size must be determined before data transmission [9–11]. According to previous studies [12–14],

the block size selection usually needs some channel information including the available transmission time and channel quality as a priori.

Nevertheless, in CRNs, similar to the idle duration, the channel quality may also be unknown to SUs. For example, when the SUs choose a new channel, the quality of this channel can not be obtained immediately. Moreover, the channel quality itself may be changing dynamically due to the complex spectrum environment. Unfortunately, most existing works (e.g., [4, 6–8]) assume that the channel quality information is perfectly known and available at SUs. Thus, the block size selection algorithms proposed in these previous studies cannot be applied to the problem of SUs with network coding under unknown channel quality.

In this paper, we propose a two-dimensional multiarmed bandit (MAB) approach to combat the problem of SUs with network coding under unknown channel quality in CRNs. Specifically, we develop a two-dimensional MAB algorithm that sequentially chooses the idle duration and block size jointly. With the help of MAB, the exploration and exploitation on the idle duration-block size pair can ensure the asymptotical-throughput optimality of the proposed algorithm. Unlike previous studies, our proposed algorithm does not need the channel quality in advance.

The novelties of this paper are as follows:

- (1) The problem of SUs with network coding under unknown channel quality in CRNs is defined.
- (2) A novel two-dimensional MAB algorithm will be proposed that chooses the idle duration and block size jointly before data transmission. The proposed algorithm is able to properly handle the case when channel quality cannot be obtained.
- (3) The optimality of the proposed two-dimensional MAB algorithm will be mathematically analyzed, and the algorithm will be theoretically proven as asymptotical throughput optimal.
- (4) The performance of the proposed algorithm is compared to the theoretical upper bound and the scheme assuming known channel quality information, in terms of throughput achieved.

The rest of the paper is organized as follows. In Section 2, the related work will be reviewed. Section 3 introduces the system model and problem formulation. We propose a two-dimensional MAB algorithm for SUs with network coding under unknown channel quality in Section 4. The performance evaluation is conducted via extensive simulations in Section 5. Finally, Section 6 concludes the paper.

## 2. Related Work

Several previous studies have demonstrated the various benefits of network coding in CRNs from different aspects. Jin et al. [4] studied the problem of multicast scheduling with network coding in CRNs and exploited network coding to perform error recovery and reduce overhead. Both Asterjadhi et al. [15] and Baldo et al. [16] utilized network coding to transmit control information reliably, in order to maintain up-to-date information in CRNs. With the help of network coding, Almasaeid and Kamal [6] proposed an algorithm to reduce the negative impact of the channel heterogeneity property on the performance of multicast in CRNs. Although network coding has been widely exploited in CRNs, the above works actually did not consider how to utilize network coding in the SUs' data transmissions. Note that the very recent studies [7, 8] have explored the throughput benefit of network coding over SUs' multicast in CRNs. However, they all assume that the channel quality of the licensed channel is perfectly available at SUs and thus their proposed algorithm cannot be applied to our problem.

On the other hand, our work can be classified as the study of learning the primary user environment in CRNs, since we try to utilize network coding for SUs' data transmission without knowing the idle duration and channel quality information. Among the works about the channel access in CRNs, there are several studies employing MAB to solve the learning-based dynamic spectrum access problem from a sequential decision aspect. Shu and Krunz [17] utilized MAB to develop a throughput optimal decision algorithm for stochastic homogeneous channels. Anandkumar et al. [18] proposed two distributed learning and allocation schemes

by MAB for the case of preallocated ranks for SUs and nonsuch information, respectively. Li et al. [19] formulated the joint channel sensing, probing, and accessing problem as a nonstochastic MAB problem and present an almost-throughput optimal algorithm for nonstochastic channels. To sum up, existing works on MAB-based dynamic spectrum access in CRNs mainly focus on how to access the best channel and merely consider how to transmit data efficiently in the uncertain idle durations under unknown channel quality after accessing the channel.

## 3. Model and Formulation

In this section, we will first introduce the system and network model used in our work and then present the problem formulation.

*3.1. Network and System Model.* We consider a CR network where secondary users (SUs) want to achieve better performance by utilizing network coding. Consider that there are one PU channel and multiple SUs. Without loss of generality, we assume that there are  $N + 1$  SUs, where one SU is the sender, and the rest of  $N$  SUs are the receivers. Before accessing the channel, the SU sender needs to sense the channel to determine whether it is idle or not. Here we assume that the time for sensing the status of the channel is  $t_s$  on average. If the channel is idle, the SU sender then should predict how long the idle status would last and determine the corresponding coding parameter it would use. We assume that the idle duration of the PU channel is nonstationary, in the sense that the idle duration length  $I$  changes with time and does not follow any probability distribution. Therefore, the accessing time of SUs may be unequal, which is more general in CRNs, while most existing studies assume the identical accessing time.

Time is slotted and synchronized among the channel and multiple SUs. We note that, in our work, one time slot consists of both the time of channel sensing and the time of transmitting one data packet. The slot structure is shown in Figure 1. Suppose there exists a maximum idle duration of the PU channel, that is,  $I_{\max}$  slots. For SUs' application, any data packet that cannot be delivered to all SU receivers is dropped without contributing to the throughput. In the SUs side, we allow each SU receiver to feedback only when it can decode a block or at the end of the current idle duration, in order to reduce the feedback overhead.

In this work, we focus on systematic network coding (SNC) [20] instead of random linear network coding (RLNC) [21] to improve SUs' throughput performance, where SNC is a special version of RLNC. The main difference between SNC and RLNC is that under RLNC all packets sent are coded packets, while under SNC both coded packets and uncoded packets are sent. Since SUs' transmissions may be interrupted by PUs at any time, we believe that SNC is more suitable for SUs than RLNC. The reason is as follows. Under RLNC, all SU receivers can decode all packets of the block simultaneously only when enough independent coded packets are collected. If a SU receiver cannot collect enough

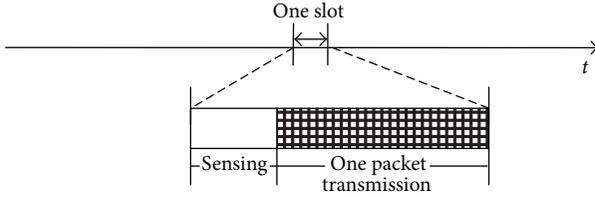


FIGURE 1: SU's slot structure for channel sensing and data transmission.

coded packets to decode the block before PUs' arriving, the number of delivered data packets is zero. Comparatively, under SNC, some SU receivers can still obtain several packets by successfully receiving the uncoded ones, even if they do not receive enough independent packets before PUs' arriving. All in all, SNC can reduce the negative effect from the uncertainty of idle durations while maintaining the throughput benefit by network coding. However, with SNC, the number of time slots for transmitting uncoded packets (the block size), for example,  $k$ , and the number of time slots for transmitting coded packets, for example,  $I - k$ , should be allocated reasonably.

According to several previous studies [12–14], the optimal value of the block size  $k$  can be determined on the estimated idle duration  $I$  and channel quality  $\epsilon$ . In these studies, the channel quality is assumed perfectly known at SUs. However, in CRNs, the channel quality may also be unknown. Therefore, the aforementioned solutions may not be applied to the case when channel quality is unknown. Accordingly, the idle duration  $I$  and block size  $k$  may be determined jointly whenever SUs transmit data by SNC on the PU channel. A reference for the major notations used in this paper is provided in A Summary of Key Notations.

**3.2. Problem Formulation.** In this study, we aim to provide a general sensing/transmitting mechanism for SUs with SNC under unknown channel quality in CRNs. The sequential sensing/transmitting problem is to determine the idle duration of the channel and the block size jointly, without knowing future channel states, for SUs to improve the throughput. We model the problem into a nonstationary multiarmed bandit (MAB) problem. Since there is no gain if no data is transmitted in one time slot, we employ several continuous time slots spent for sensing the channel as our study unit, which is called as a round in this paper. After a round, the SU sender may or may not transmit data, according to the sensing result and previous gains. Note that the transmission time slots are not counted in a round. The strategy for sensing/transmitting is composed of many sequential rounds of sensing.

A sensing/transmitting strategy  $\psi = \langle \alpha_t, i_t, k_t \rangle$  by the SU sender will decide its action. At every round,  $\alpha_t \in \{0, 1\}$  denotes whether the SUs decide to access the channel,  $i_t \in \{1, 2, \dots, I_{\max}\}$  is the idle duration selected, and  $k_t \in \{1, 2, \dots, i_t\}$  denotes the block size used. The value of  $\alpha_t$  is determined according to the sensing result:  $\alpha_t = 1$  if the channel is busy and  $\alpha_t = 0$  otherwise. Note that, for any  $t$ ,

$k_t$  should be strictly no more than  $i_t$ ; otherwise, the block cannot be decoded. At each round  $t$ , the SU sender should choose a strategy vector  $\langle i_t, k_t \rangle$  over strategy space  $\varphi = \{1, 2, \dots, I_{\max}\} \times \{1, 2, \dots, I_{\max}\}$  with the restriction that the second component should be no greater than the first one. The size of  $\varphi$  is thus equal to  $(I_{\max} + 1)I_{\max}/2$ .

Let  $g_\psi(t)$  denote the gain of a strategy  $\psi$  at round  $t$ . In practical, the SU sender will examine the performance of its strategy based on the ACK from the SU receivers after each round. The actual gain of each round can thus be calculated as the total number of successful delivered data packets. The accumulative gain up to round  $T$  of each strategy  $\psi$  is defined as

$$G_\psi(T) = \sum_{t=1}^T g_\psi(t), \quad (1)$$

and the total gain of all chosen strategies accumulating up to round  $T$  is

$$\widehat{G}(T) = \sum_{t=1}^T g_{\psi_t}(t), \quad (2)$$

where the strategy  $\psi_t$  is the SUs' strategy at round  $t$ , which is chosen randomly according to the determined probability distribution over  $\varphi$ .

In the modeled two-dimensional MAB problem for SUs with SNC, each idle duration is considered as an arm of a gamble machine and the block size is the coin the gamble machine bets on that arm. The reward is defined as the number of successfully delivered data packets. Typically, the gambler's performance is measured in terms of regret in the MAB problem. It is defined as the difference between the expected return of the gambler's actions and that of the static optimal strategy. In this study, we aim to design a strategy  $\psi$  that maximizes the expected throughput; that is,

$$\max_{\psi} \frac{\sum_{t=1}^T g_{\psi_t}(t)}{T'}, \quad (3)$$

under the system lifetime constraint,

$$T \cdot t_s + \sum_{t=1}^T \alpha_t \cdot i_t \leq T', \quad (4)$$

where the system lifetime  $T'$  is defined as the total time for both transmitting and sensing at each round.

The regret in our application is the difference between the expected total number of delivered packets using our proposed algorithm and that using the static optimal strategy over lifetime  $T'$ . The static optimal strategy for the SUs is the strategy  $\psi$  to achieve the maximum delivered data packets if the SUs keep using that strategy for all  $T$  rounds. Therefore, we can define the SUs' regret  $\sigma$  after  $T$  rounds of an online strategy  $\psi_t$  as

$$\sigma = \max_{\psi \in \varphi} G_\psi(T) - \widehat{G}(T). \quad (5)$$

A strategy whose average regret per round  $\sigma/T' \rightarrow 0$  with probability 1 when  $T' \rightarrow \infty$  is a zero-regret strategy. Our objective is to design a strategy  $\psi$  with small regret.

#### 4. Two-Dimensional MAB Algorithm for SUs with Network Coding

In this section, we focus on developing a MAB-based algorithm for SUs with network coding under unknown channel quality in CRNs. Specifically, we propose a two-dimensional MAB algorithm to sequentially choose the idle duration  $i_t$  and block size  $k_t$  jointly. It is worth emphasizing that the proposed two-dimensional MAB algorithm does not work for SUs without network coding, since it is unnecessary to determine the block size, when network coding is not employed in SUs' transmissions. In the rest of this section, we will first provide an overview of the algorithm, explain the detailed design of the algorithm, and then theoretically analyze the performance of the proposed algorithm.

**4.1. Overview.** Inspired by the approach in [19], we propose a two-dimensional MAB algorithm for SUs with SNC under unknown channel quality in CRNs, as shown in Algorithm 1. As put in Algorithm 1, three parameters are involved and their meanings are follows.  $\beta$  is to control the bias in the estimation of gain  $g'_{i,k}(t)$ .  $\eta$  is used to control the learning speed. And  $\gamma$  reflects the tradeoff between *exploration* and *exploitation*. The values of  $\beta$ ,  $\eta$ , and  $\gamma$  are very critical to the performance of the proposed algorithm, which will be discussed in Section 4.3.

Generally speaking, the SUs choose the strategy vector  $\langle i_t, k_t \rangle$  over the strategy space  $\varphi$  according to the corresponding probability distribution  $\{p_{i,k}\}_{(i,k) \in \varphi}$ . In the very beginning, the strategy probability is uniformly distributed, that is,  $p_{i,k}(1) = (1/(I_{\max}(I_{\max} + 1)/2)) (\forall (i, k) \in \varphi)$ , since there is no idea about the relationship between the gain and strategy pair initially. This means that the SUs begin to explore the best strategy pair uniformly over the entire strategy space  $\varphi$ , as illustrated in the initialization step. In the following rounds,  $p_{i,k}$  is determined by the corresponding strategy weight  $w_{i,k}$  in the overall strategy weights.

**4.2. Design.** In step 1, the calculation of  $p_{i,k}(t)$  reflects the tradeoff between the exploitation and exploration. Specifically, with probability  $1 - \gamma$ , we will exploit the best idle duration-block size pair in previous rounds, which is shown in the left part of the calculation. It is worth emphasizing that a better idle duration-block size pair  $\langle i, k \rangle$  always implies a greater weight  $w_{i,k}$ . The exploitation will guarantee an asymptotical optimal performance if the previously used strategy is asymptotically optimal. On the other hand, with probability  $\gamma$ , we will explore new idle duration-block size pair with an equal probability  $1/(I_{\max}(I_{\max} + 1)/2)$ , which is illustrated in the right part of the calculation. The exploration would eventually improve our strategy to the static optimal solution.

In steps 2 and 3, the SUs execute the transmission according to the selected  $i_t$  and  $k_t$  and get the scaled output afterwards, respectively. Then, the SUs calculate the gain of the current strategy pair in step 4 and update all weights based on the strategy gain in step 5. Note that, at  $t$ th round, the idle duration-block size pair  $\langle i_t, k_t \rangle$  is chosen according to the corresponding probability distribution  $p_{i,k}(t)$ . Specifically, every idle duration-block size pair  $\langle i, k \rangle$  dominates

a probability  $p_{i,k}$  ( $i = 1, 2, \dots, I_{\max}$ ,  $k = 1, 2, \dots, i$ ). To choose the current idle duration-block size pair, we generate a random value  $\rho \in (0, 1)$  and select  $\langle i^*, k^* \rangle$  if  $\rho$  falls in the corresponding probability  $p_{i^*, k^*}$ . And the output  $g_{i_t, k_t}(t)$  is calculated based on the feedback from all SU receivers. The feedback information includes the packets that are decoded by each SU receiver. According to the feedback, the SU sender calculates the number of common data packets decoded by all SU receivers and then normalizes with the block size to get the scaled output  $g_{i_t, k_t}(t)$ . For instance, if the block size  $k_t = 10$  and the number of common data packets decoded by all SU receivers is 6, then  $g_{i_t, k_t}(t)$  is 0.6.

*Algorithm 1* (two-dimensional MAB algorithm for SUs with SNC under unknown channel quality).

*Parameters.* Real number  $\beta > 0$  and  $0 < \eta, \gamma < 1/2$ .

*Initialization.* Set the initial strategy weights for all the idle duration-block size pairs: for  $i = 1, 2, \dots, I_{\max}$  and  $k = 1, 2, \dots, i$ ,  $w_{i,k}(0) = 1$ .

For round  $t = 1, 2, \dots, T$ , one has the following:

- (1) Calculate the strategy probability distribution:

$$p_{i,k}(t) = (1 - \gamma) \frac{w_{i,k}(t-1)}{\sum_{i=1}^{I_{\max}} \sum_{k=1}^i w_{i,k}(t-1)} + \frac{\gamma}{I_{\max}(I_{\max} + 1)/2}, \quad (6)$$

for  $i = 1, 2, \dots, I_{\max}$ ,  $k = 1, 2, \dots, i$ .

- (2) Choose the strategy  $(i_t, k_t)$  in  $t$ th round according to the above probability distribution, transmit data by SNC to the SU receivers based on the idle duration  $i_t$ , and block size  $k_t$ .
- (3) Get the scaled output  $g_{i_t, k_t}(t) \in [0, 1]$  after the round based on the feedback from all SU receivers.
- (4) For  $i = 1, 2, \dots, I_{\max}$ , and  $k = 1, 2, \dots, i$ , set

$$g'_{i,k}(t) = \begin{cases} \frac{g_{i,k}(t)}{p_{i,k}(t)} + \frac{\beta}{p_{i,k}(t)}, & \text{if } i = i_t, k = k_t, \\ \frac{\beta}{p_{i,k}(t)}, & \text{otherwise.} \end{cases} \quad (7)$$

- (5) Update all the weights as

$$w_{i,k}(t) = w_{i,k}(t-1) \exp(\eta g'_{i,k}(t)), \quad (8)$$

for  $i = 1, 2, \dots, I_{\max}$ ,  $k = 1, 2, \dots, i$ .

**4.3. Theoretic Analysis.** In this part, we analyze the performance of our two-dimensional MAB algorithm in terms of regret. In our application, the regret is the difference between the number of successfully delivered data packets using the static optimal idle duration-block size pair and that using our two-dimensional MAB algorithm.

**Theorem 2.** For any  $\delta \in (0, 1)$ , when  $\beta = \sqrt{(1/LT)\ln(L/\delta)}$ ,  $\eta = \sqrt{\ln L/4TL}$ , and  $\gamma = \sqrt{L\ln L/T}$ , with probability  $1 - \delta$ , the regret of Algorithm 1 is bounded by  $\sigma \leq 6\sqrt{TL\ln L}$ , where  $L = I_{\max}(I_{\max} + 1)/2$ .

*Proof.* The detailed proof is provided in the Appendix.  $\square$

Based on the above regret analysis, we prove that Algorithm 1 is asymptotically throughput optimal as shown in the following theorem.

**Theorem 3.** Algorithm 1 is asymptotically throughput optimal when  $T$  is sufficiently large.

*Proof.* According to Theorem 2, we have  $\sigma/T' = \sigma/(Tt_s + \sum_{t=1}^T i_t) \leq 6\sqrt{TL\ln L}/(Tt_s + \sum_{t=1}^T i_t) \leq 6\sqrt{TL\ln L}/(Tt_s + T) = 1/\sqrt{T} \cdot 6\sqrt{L\ln L}/(t_s + 1)$ . Therefore, when  $T$  is sufficiently large, the throughput of our algorithm is asymptotically optimal.  $\square$

In Theorem 3, the value of  $T$  is critical to the asymptotical optimality of Algorithm 1. To guarantee the asymptotical optimality, we may give  $\theta$  a small value, for example,  $\theta = 0.01$ , to derive the approximate value of  $T$ . Specifically, according to the proof of Theorem 3, we only need to solve the inequality  $1/\sqrt{T} \cdot 6\sqrt{L\ln L}/(t_s + 1) \leq \theta$ . The solution of the aforementioned inequality is easily obtained as follows:  $T \geq 36L\ln L/\theta^2(t_s + 1)^2$ .

## 5. Results and Discussion

In this section, we evaluate the performance of our proposed two-dimensional MAB algorithm via simulations. Note that, when channel quality is known in advance, the optimal coding block size of SNC can be determined according to the current idle duration [7]. We call this block size selection scheme ‘‘OSNC.’’ We take the algorithm that assumes both perfect idle duration information and known channel quality and employs OSNC as a reference labeled as ‘‘Full Information + OSNC’’ in the figures, whose performance is the theoretical upper bound of all schemes. Furthermore, to show the efficiency of the two-dimensional MAB algorithm under uncertain channel quality, we compare it with the one-dimensional MAB algorithm with OSNC proposed in [7] labeled as ‘‘MAB + OSNC’’ in the figures, which assumes the known channel quality information and chooses the idle duration by MAB.

Besides, we define a new metric named ‘‘utilization,’’ which is the ratio of the performance of our two-dimensional MAB algorithm to that of ‘‘Full Information + OSNC’’ or ‘‘MAB + OSNC.’’ This metric can not only precisely characterize the spectrum utilization of SUs in CRNs but also show the throughput performance of our proposed algorithm under unknown channel quality. Lastly, we present the practical regrets of our two-dimensional MAB algorithm to validate the regret bound analysis.

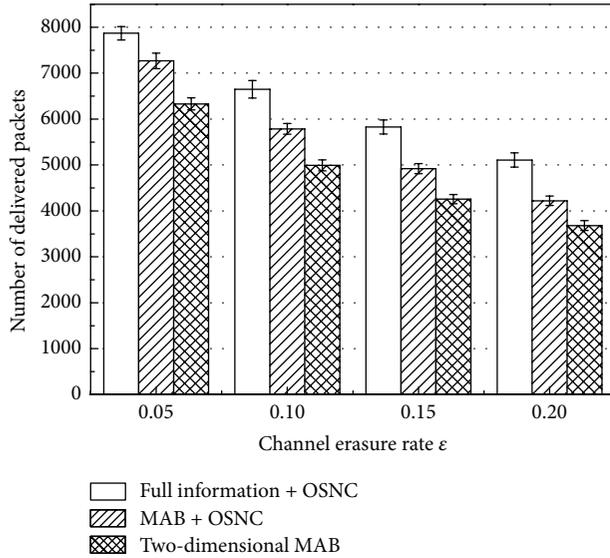
**5.1. Simulation Results.** In the simulation, we fix the number of SU receivers  $N$  at 10 and vary the identical channel erasure

rate  $\epsilon$  from 0.05 to 0.20. The number of total rounds  $T$  is fixed at 1000. We define the total number of packets delivered over the system life time as the main performance metric, since we mainly focus on the throughput performance in our study.

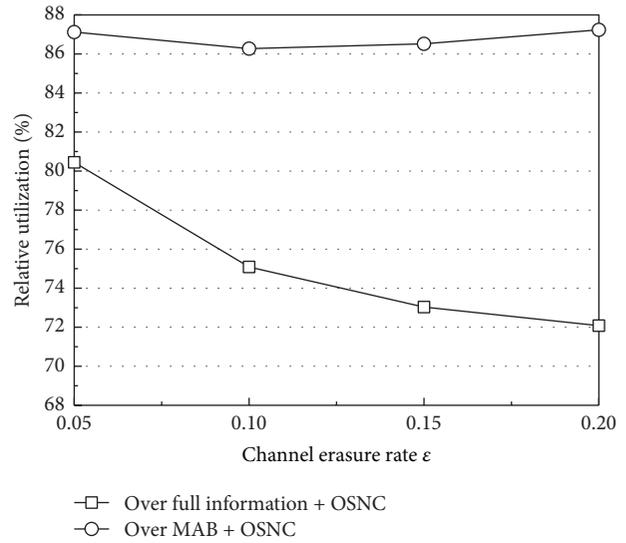
To simulate a trace of nonstationary idle durations, we mix several different probability distributions. In this work, we only present two representative cases due to space limitation. One consists of a Poisson distribution  $P(\lambda)$ , an uniform distribution  $U[A, B]$ , and a geometric distribution  $G(p)$ . Another is composed of a hypergeometric distribution  $H(n, M, N)$ , a normal distribution  $N(\mu, \sigma^2)$ , and a binomial distribution  $B(K, q)$ . To be specific, we generate three sequences of the distributions and then everyone of the trace data is chosen with equal probability from these distributions. Accordingly, the generated sequence does not follow any probability distribution. We set the distribution parameters  $\lambda = 50$ ,  $A = 30$ ,  $B = 50$ ,  $p = 0.025$ ,  $n = 60$ ,  $M = 30$ ,  $N = 40$ ,  $\mu = 25$ ,  $\sigma = 10$ , and  $K = 40$ , as an illustration in this study. Moreover, we also generate a trace of nonstationary idle durations by applying a Poisson distribution with round-dependent parameter; that is,  $r$  continuous idle durations constitute a Poisson distribution with a specific parameter  $\lambda'$ . Specifically, the  $T$  idle durations are divided into  $T/r$  groups, where the idle durations in  $i$ th group follow the Poisson distribution  $P(20|\sin(2ir))$  ( $i = 1, 2, \dots, T/r$ ). In the simulations, we set the traces with  $r = 50$  (fast changing) and  $r = 200$  (slowly changing) as nonstationary idle durations III and IV, respectively.

Figure 2 shows the performance of different algorithms under the first nonstationary idle durations when channel quality deteriorates. From Figure 2(a), we can see that the performance of all algorithms degrades with the deterioration of channel quality, while the proposed two-dimensional MAB algorithm achieves fairly comparable performance, compared with the theoretic upper bound and the ‘‘MAB + OSNC’’ assuming the known channel quality in advance. Specifically, the relative utilization of the two-dimensional MAB algorithm over ‘‘Full Information + OSNC’’ and ‘‘MAB + OSNC’’ is up to 80.2% and 87%, respectively. This is because although we do not know enough channel information to select the optimal block size, we are able to approximate the optimal transmission scheme by the tradeoff between exploiting the past best idle duration-block size pair and exploring the new possible better idle duration-block size pair. This shows the great advantage of our proposed two-dimensional MAB algorithm, which promises a wide range of applications when faced with more channel uncertainties. For example, when SUs enter a new CRN, the channel quality information cannot be learned at once. In this situation, our proposed two-dimensional MAB algorithm can be used immediately and achieves a comparable performance. Besides, this two-dimensional MAB algorithm has flexible expansibility compared to the ‘‘MAB + OSNC,’’ since different channel may have different quality and the quality of one single channel may be changing with time.

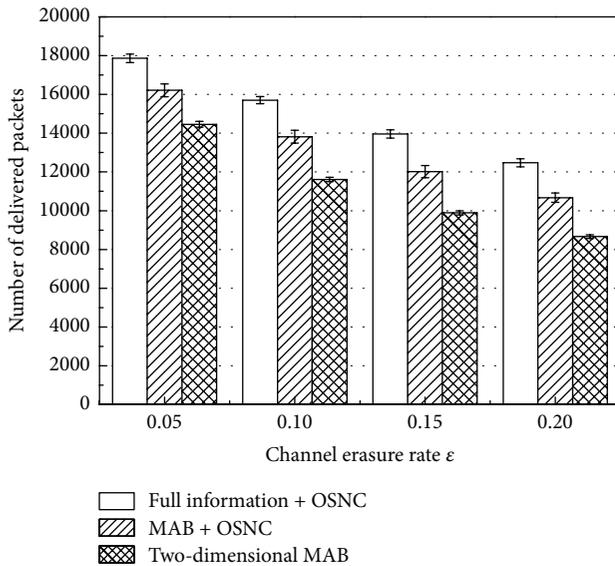
Moreover, from Figure 2(b), we observe that the utilization of the proposed algorithm over ‘‘MAB + OSNC’’ is relatively stable when channel quality varies. This is because although the performance of both the proposed algorithm



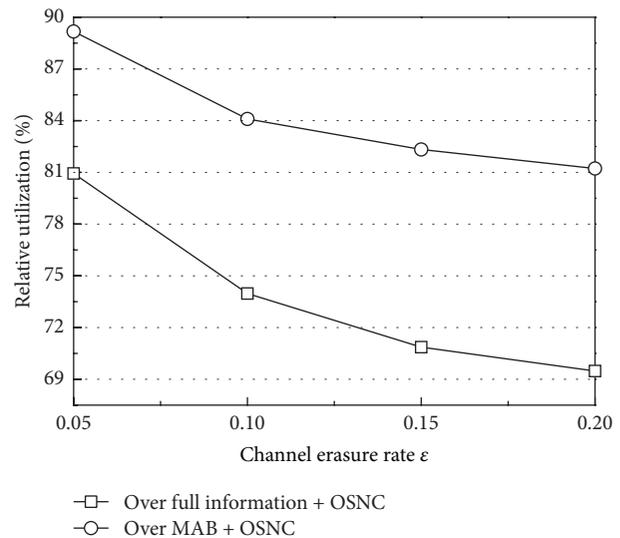
(a) Performance of different schemes under different channel qualities



(b) Relative utilization of two-dimensional MAB Algorithm

FIGURE 2: Performance comparison under *nonstationary idle duration I*.

(a) Performance of different schemes under different channel qualities

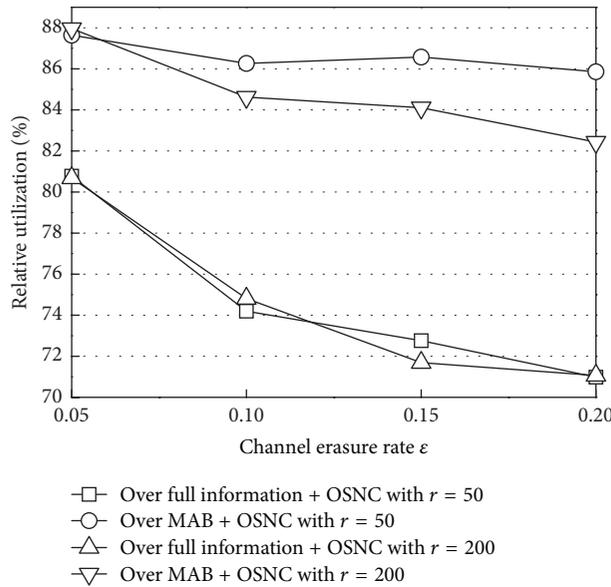
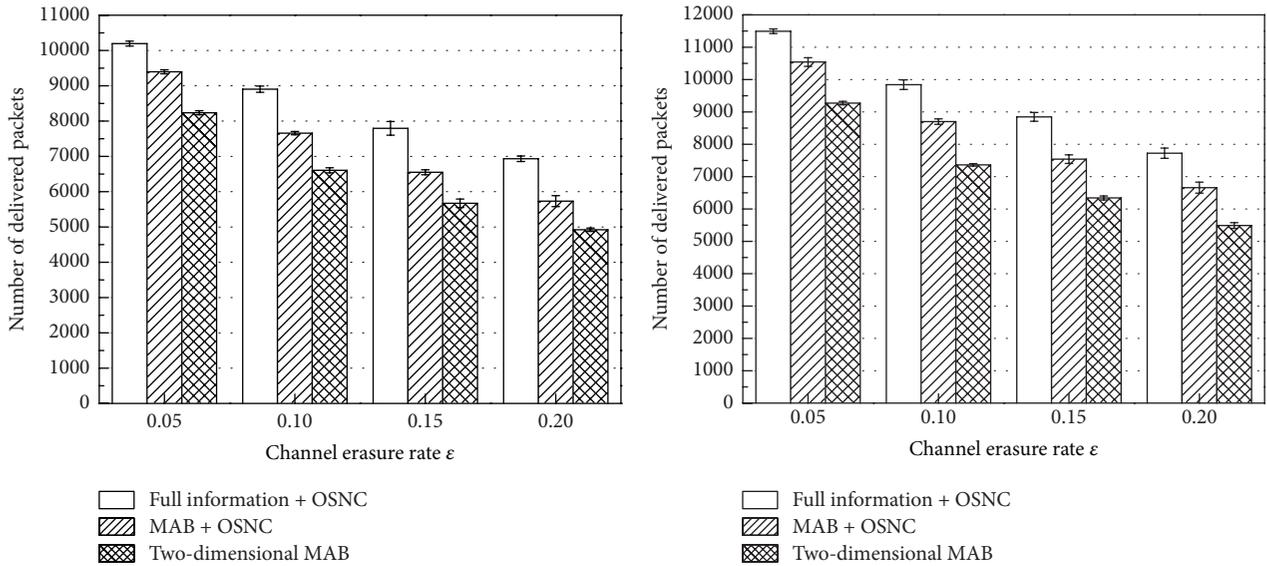


(b) Relative utilization of two-dimensional MAB algorithm

FIGURE 3: Performance comparison under *nonstationary idle duration II*.

and “MAB + OSNC” degrades with the deterioration of channel erasure rate, our proposed algorithm jointly selects the idle duration and block size to maximize the throughput rather than only estimate the idle duration. Also, as illustrated in Figure 2(b), the utilization of our proposed algorithm over “Full Information + OSNC” decreases when channel erasure rate increases. The reason is as follows. When channel erasure rate increases, the uncertainty of the channel is more complex besides the idle duration, which can make the learning of the two-dimensional MAB algorithm harder.

Figure 3 describes the performance of different algorithms under the second nonstationary idle durations when channel quality deteriorates. The performance of the two-dimensional MAB algorithm under the second nonstationary idle durations is similar to that under the first nonstationary idle durations. The utilization of the two-dimensional MAB algorithm over “Full Information + OSNC” and “MAB + OSNC” is 81% and 87% at most, respectively. Also, as shown in Figure 3(b), the utilization of our proposed algorithm over “MAB + OSNC” under the second idle durations changes



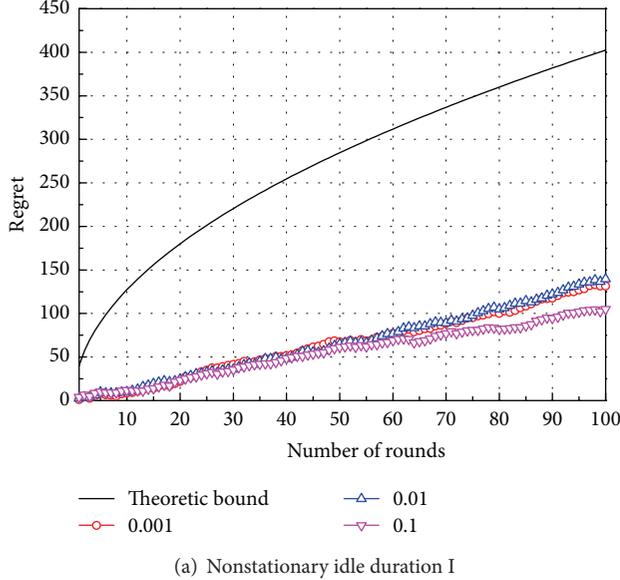
(c) Relative utilization under nonstationary idle durations III and IV

FIGURE 4: Performance under *nonstationary idle durations III and IV*.

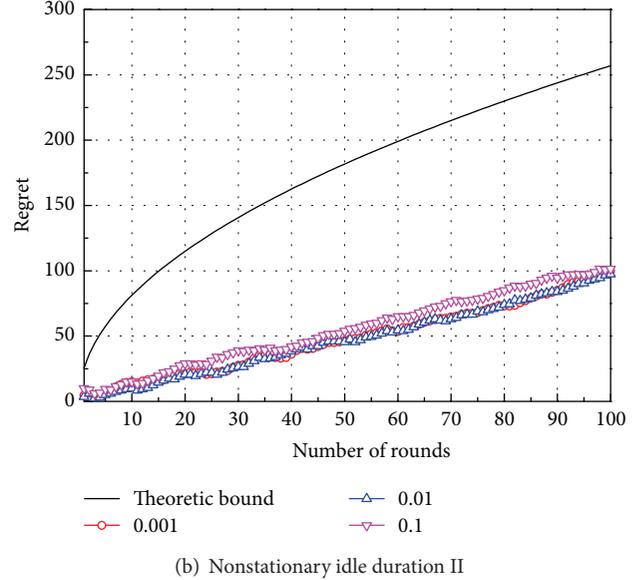
from 87% to 81% when channel erasure rate varies from 0.05 to 0.2. Although the utilization of the two-dimensional MAB algorithm over “MAB + OSNC” in the second idle durations is not as stable as that in the first idle durations, the utilization is as least 81%. This shows the effectiveness of our proposed algorithm choosing idle duration and block size jointly when faced with uncertain channel quality.

Figure 4 presents the performance results of our proposed two-dimensional MAB algorithm under nonstationary idle durations III and IV, compared with the “Full Information + OSNC” and “MAB + OSNC.” From Figures 4(a) and 4(b), we can see that the two-dimensional MAB algorithm achieves

similar relative throughput performance under nonstationary idle durations III and IV, with the deterioration of channel quality. Furthermore, according to Figure 4(c), the relative utilizations of the two-dimensional MAB algorithm over “Full Information + OSNC” and “MAB + OSNC” are up to 81% and 88%, respectively. And the relative utilization of the proposed algorithm over “Full Information + OSNC” is stable, in the sense that the two relative utilization curves almost coincide when  $r$  varies from 50 to 200. This shows that the stability of the two-dimensional MAB algorithm is relative independent of the type of nonstationary idle duration length. This is because, with the help of MAB,



(a) Nonstationary idle duration I



(b) Nonstationary idle duration II

FIGURE 5: Practical regrets of two-dimensional MAB algorithm.

the two-dimensional MAB algorithm can efficiently handle the uncertainty of both idle duration and channel quality, no matter whether the idle duration is slowly changing or fast changing.

**5.2. Analysis of the Practical Regrets.** To validate the regret bound analysis, we conduct simulations on the practical regrets achieved by the proposed two-dimensional MAB algorithm, compared with the corresponding theoretic bound. Specifically, we fix the number of SU receivers  $N$  and the channel erasure rate  $\epsilon_n$  ( $n \in \{1, 2, \dots, N\}$ ) at 10 and 0.1, respectively, and increase the number of rounds  $T$  from 1 to 100. To better investigate the regret, we let the parameter  $\delta$  in Theorem 2 change from 0.001 to 0.1, which represents different confidential levels on the achieved regret bound.

Figures 5(a) and 5(b) illustrate the practical regrets of the two-dimensional MAB algorithm compared with the theoretic bound derived by Theorem 2, under nonstationary idle durations I and II, respectively. As shown in the figures, the practical regrets of the proposed two-dimensional MAB algorithm are strictly bounded by the corresponding theoretic bound, no matter what the value of  $\delta$  is chosen. These results validate our theoretic analysis on the regret bounds.

## 6. Conclusion

In this paper, we investigate how to utilize network coding to improve the throughput performance of SUs in CRNs when the PU channel quality information is not available. We formulate the problem of SUs with network coding under unknown channel quality as a two-dimensional MAB problem. We propose a two-dimensional MAB algorithm that chooses the idle duration and block size jointly. The performance of the proposed algorithm is analyzed in terms of regrets and analytically proven asymptotically throughput

optimal, compared to the static optimal scheme. Our extensive simulation results show that the throughput performance of our proposed algorithm is close to the theoretic bound and the network coding algorithm assuming known channel quality information, by achieving up to 87% utilization.

## Appendix

### Proof of Theorem 2

*Proof.* Define  $G_{i,k}(T) = \sum_{t=1}^T g_{i,k}(t)$ ,  $G'_{i,k}(T) = \sum_{t=1}^T g'_{i,k}(t)$  as the actual gain and virtual gain for strategy  $(i, k)$  up to round  $T$ , respectively. And the total gain up to round  $T$  of the chosen strategy sequence  $(i_t, k_t)_{t=1,2,\dots,T}$  is as follows:  $\widehat{G}(T) = \sum_{t=1}^T g_{i_t, k_t}(t)$ .

Define  $W(t) = \sum_{i=1}^{I_{\max}} \sum_{k=1}^i w_{i,k}(t)$ . Since  $w_{i,k}(t) = \prod_{t'=0}^{t-1} (w_{i,k}(t'+1)/w_{i,k}(t')) = \prod_{t'=1}^t e^{\eta g'_{i,k}(t')}$ , we have  $W(T) = \sum_{i=1}^{I_{\max}} \sum_{k=1}^i \prod_{t'=1}^T e^{\eta g'_{i,k}(t')} = \sum_{i=1}^{I_{\max}} \sum_{k=1}^i e^{\eta G'_{i,k}(T)}$  and  $W(0) = I_{\max}(I_{\max} + 1)/2$ . We derive the bound of the regret  $\sigma$  by using the quantity  $\ln(W(T)/W(0))$ . For simplicity, we define  $L = I_{\max}(I_{\max} + 1)/2$ .

For the lower bound, according to the definitions, we have

$$\begin{aligned} \ln \frac{W(T)}{W(0)} &= \ln \sum_{i=1}^{I_{\max}} \sum_{k=1}^i e^{\eta G'_{i,k}(T)} - \ln L \\ &\geq \eta \max_{(i,k) \in \Psi} G'_{i,k}(T) - \ln L. \end{aligned} \quad (\text{A.1})$$

For the upper bound,

$$\begin{aligned} \ln \frac{W(t)}{W(t-1)} &= \ln \sum_{i=1}^{I_{\max}} \sum_{k=1}^i \frac{w_{i,k}(t)}{W(t-1)} \\ &= \ln \sum_{i=1}^{I_{\max}} \sum_{k=1}^i \frac{w_{i,k}(t-1)}{W(t-1)} e^{\eta g'_{i,k}(t)} \end{aligned}$$

$$\begin{aligned}
 &\leq \ln \left\{ \sum_{i=1}^{I_{\max}} \sum_{k=1}^i \frac{w_{i,k}(t-1)}{W(t-1)} \left[ \eta g'_{i,k}(t) + \eta^2 g_{i,k}^{\prime 2}(t) \right] \right\} \\
 &= \ln \left\{ 1 + \sum_{i=1}^{I_{\max}} \sum_{k=1}^i \frac{p_{i,k}(t)}{1-\gamma} \left[ \eta g'_{i,k}(t) + \eta^2 g_{i,k}^{\prime 2}(t) \right] \right\} \\
 &\leq \frac{\eta}{1-\gamma} \sum_{i=1}^{I_{\max}} \sum_{k=1}^i p_{i,k}(t) g'_{i,k}(t) \\
 &\quad + \frac{\eta^2}{1-\gamma} \sum_{i=1}^{I_{\max}} \sum_{k=1}^i p_{i,k}(t) g_{i,k}^{\prime 2}(t).
 \end{aligned} \tag{A.2}$$

On the one hand, we have

$$\sum_{i=1}^{I_{\max}} \sum_{k=1}^i p_{i,k}(t) g'_{i,k}(t) = g_{i_t, k_t}(t) + L\beta. \tag{A.3}$$

On the other hand,

$$\sum_{i=1}^{I_{\max}} \sum_{k=1}^i p_{i,k}(t) g_{i,k}^{\prime 2}(t) = (1+\beta) \sum_{i=1}^{I_{\max}} \sum_{k=1}^i g'_{i,k}(t), \tag{A.4}$$

which is due to  $p_{i,k}(t)g'_{i,k}(t) \leq g_{i,k}(t) + \beta \leq 1 + \beta$ .

Thus,

$$\begin{aligned}
 \ln \frac{W(t)}{W(t-1)} &\leq \frac{\eta}{1-\gamma} (g_{i_t, k_t}(t) + L\beta) \\
 &\quad + \frac{\eta^2(1+\beta)}{1-\gamma} \sum_{i=1}^{I_{\max}} \sum_{k=1}^i g'_{i,k}(t).
 \end{aligned} \tag{A.5}$$

Summing for  $t = 1, 2, \dots, T$ , we have the following inequality:

$$\begin{aligned}
 \ln \frac{W(T)}{W(0)} &\leq \frac{\eta}{1-\gamma} (\widehat{G}(T) + L\beta T) \\
 &\quad + \frac{\eta^2(1+\beta)}{1-\gamma} \sum_{i=1}^{I_{\max}} \sum_{k=1}^i G'_{i,k}(T) \\
 &\leq \frac{\eta}{1-\gamma} (\widehat{G}(T) + L\beta T) \\
 &\quad + \frac{\eta^2(1+\beta)}{1-\gamma} L \max_{(i,k) \in \Psi} G'_{i,k}(T).
 \end{aligned} \tag{A.6}$$

Combining the lower bound and the upper bound, we have

$$\begin{aligned}
 \frac{\eta}{1-\gamma} (\widehat{G}(T) + L\beta T) + \frac{\eta^2(1+\beta)}{1-\gamma} L \max_{(i,k) \in \Psi} G'_{i,k}(T) \\
 \geq \eta \max_{(i,k) \in \Psi} G'_{i,k}(T) - \ln L.
 \end{aligned} \tag{A.7}$$

That is also

$$\begin{aligned}
 \widehat{G}(T) &\geq (1-\gamma-\eta(1+\beta)L) \max_{(i,k) \in \Psi} G'_{i,k}(T) \\
 &\quad - \frac{1-\gamma}{\eta} \ln L - L\beta T.
 \end{aligned} \tag{A.8}$$

For any fixed  $u > 0$  and  $v > 0$ , according to the Chernoff bound, we have  $\Pr[G_{i,k}(T) > G'_{i,k}(T) + u] \leq e^{-uv} E[e^{v[G_{i,k}(T) - G'_{i,k}(T)]]]$ . Let  $v = \beta$  and  $u = (\ln(L/\delta))/\beta$ . Then,  $\Pr[G_{i,k}(T) > G'_{i,k}(T) + (1/\beta)\ln(L/\delta)] \leq \delta/L$ .

Applying this bound, with probability at least  $1 - \delta$ , we have

$$\begin{aligned}
 \widehat{G}(T) &\geq (1-\gamma'-\eta(1+\beta)L) \left[ \max_{(i,k) \in \Psi} G_{i,k}(T) - \frac{1}{\beta} \ln \frac{L}{\delta} \right] \\
 &\quad - \frac{1-\gamma}{\eta} \ln L - L\beta T.
 \end{aligned} \tag{A.9}$$

By doing some transpositions and using the fact that  $\max_{(i,k) \in \Psi} G_{i,k}(T) \leq T$ , with probability  $1 - \delta$ , we have

$$\begin{aligned}
 \max_{(i,k) \in \Psi} G_{i,k}(T) - \widehat{G}(T) &\leq (\gamma + \eta(1+\beta)L)T \\
 &\quad + (1-\gamma-\eta(1+\beta)L) \frac{1}{\beta} \ln \frac{L}{\delta} + \frac{1-\gamma}{\eta} \ln L \\
 &\quad + L\beta T \\
 &\leq \gamma T + 2\eta TL + \frac{1}{\beta} \ln \frac{L}{\delta} + \frac{1-\gamma}{\eta} \ln L + L\beta T.
 \end{aligned} \tag{A.10}$$

If we set  $\beta = \sqrt{(1/LT) \ln(L/\delta)}$ ,  $\eta = \sqrt{\ln L/4TL}$ , and  $\gamma = \sqrt{L \ln L/T}$ , then

$$\max_{(i,k) \in \Psi} G_{i,k}(T) - \widehat{G}(T) \leq 6\sqrt{TL \ln L} \tag{A.11}$$

with probability  $1 - \delta$ .  $\square$

## A Summary of Key Notations

$N$ :	# of SU receivers
$i_t$ :	# of time slots determined by SUs at round $t$
$T$ :	# of total rounds
$k_t$ :	Block size determined at round $t$
$I$ :	Estimated idle duration
$t_s$ :	Average sensing time
$I_{\max}$ :	Maximum idle duration
$\varphi$ :	Set of idle duration and block size pairs
$\mathbb{N}$ :	Set of SU receivers
$\psi(t)$ :	SUs' strategy at round $t$
$\alpha_t$ :	SUs' decision to access the channel or not at round $t$
$g_\phi(t)$ :	Gain for strategy $\phi$ at round $t$
$G_s(T)$ :	Total gain over a fixed strategy $s$ up to round $T$
$\widehat{G}(T)$ :	Total gain over chosen strategies up to round $T$ .

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work is supported by the National NSF of China under Grants no. 61371124 and no. 61472445 and the NSF of Jiangsu Province under Grant no. BK20140076.

## References

- [1] J. Mitola III, *Cognitive radio: an integrated agent architecture for software defined radio [Ph.D. dissertation]*, KTH Royal Institute of Technology, 2000.
- [2] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Computer Networks*, vol. 50, no. 13, pp. 2127–2159, 2006.
- [3] R. Ahlswede, N. Cai, S. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.
- [4] J. Jin, H. Xu, and B. Li, "Multicast scheduling with cooperation and network coding in cognitive radio networks," in *Proceedings of the IEEE Conference on Computer Communications (IEEE INFOCOM '10)*, pp. 1–9, San Diego, Calif, USA, March 2010.
- [5] L. Geng, Y.-C. Liang, and F. Chin, "Network coding for wireless ad hoc cognitive radio networks," in *Proceedings of the 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '07)*, pp. 1–5, September 2007.
- [6] H. M. Almasaeid and A. E. Kamal, "Exploiting multichannel diversity for cooperative multicast in cognitive radio mesh networks," *IEEE/ACM Transactions on Networking*, vol. 22, no. 3, pp. 770–783, 2014.
- [7] Y. Qu, C. Dong, S. Tang, C. Chen, H. Wang, and C. Tian, "Towards near optimal network coding for secondary users in cognitive radio networks," in *Proceedings of the 11th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON '14)*, pp. 522–530, July 2014.
- [8] Y. Qu, C. Dong, H. Dai et al., "Network coding-based multicast in multi-hop CRNs under uncertain spectrum availability," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM '15)*, Hong Kong, April-May 2015.
- [9] S. Chachulski, M. Jennings, S. Katti, and D. Katabi, "Trading structure for randomness in wireless opportunistic routing," in *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM '07)*, pp. 169–180, 2007.
- [10] P. A. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proceedings of the Annual Allerton Conference on Communication Control and Computing*, pp. 40–49, Urbana, Ill, USA, October 2003.
- [11] C. Fragouli, D. Lun, M. Médard, and P. Pakzad, "On feedback for network coding," in *Proceedings of the 41st Annual Conference on Information Sciences and Systems (CISS '07)*, pp. 248–252, Baltimore, Md, USA, March 2007.
- [12] L. Yang, Y. E. Sagduyu, and J. H. Li, "Adaptive network coding for scheduling real-time traffic with hard deadlines," in *Proceedings of the 13th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '12)*, pp. 105–114, June 2012.
- [13] I.-H. Hou and P. R. Kumar, "Broadcasting delay-constrained traffic over unreliable wireless links with network coding," in *Proceedings of the 12th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '11)*, May 2011.
- [14] R. Li, H. Gangammanavar, and A. Eryilmaz, "Optimal dynamic coding-window selection for serving deadline-constrained traffic over time-varying channels," *IEEE Transactions on Information Theory*, vol. 58, no. 10, pp. 6556–6571, 2012.
- [15] A. Asterjadhi, N. Baldo, and M. Zorzi, "A distributed network coded control channel for multihop cognitive radio networks," *IEEE Network*, vol. 23, no. 4, pp. 26–32, 2009.
- [16] N. Baldo, A. Asterjadhi, and M. Zorzi, "Dynamic spectrum access using a network coded cognitive control channel," *IEEE Transactions on Wireless Communications*, vol. 9, no. 8, pp. 2575–2587, 2010.
- [17] T. Shu and M. Krunz, "Throughput-efficient sequential channel sensing and probing in cognitive radio networks under sensing errors," in *Proceedings of the 15th Annual ACM International Conference on Mobile Computing and Networking (MobiCom '09)*, pp. 37–48, ACM, September 2009.
- [18] A. Anandkumar, N. Michael, A. K. Tang, and A. Swami, "Distributed algorithms for learning and cognitive medium access with logarithmic regret," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 731–745, 2011.
- [19] X.-Y. Li, P. Yang, Y. Yan, L. You, S. Tang, and Q. Huang, "Almost optimal accessing of nonstochastic channels in cognitive radio networks," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM '12)*, pp. 2291–2299, Orlando, Fla, USA, March 2012.
- [20] B. Shrader and N. M. Jones, "Systematic wireless network coding," in *Proceedings of the IEEE Military Communications Conference (MILCOM '09)*, pp. 1–7, October 2009.
- [21] T. Ho, M. Médard, R. Koetter et al., "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, 2008.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

