

# Scaling Laws for Cognitive Radio Network with Heterogeneous Mobile Secondary Users

Yingzhe Li, Xinbing Wang, Xiaohua Tian  
Dept. of Electronic Engineering  
Shanghai Jiao Tong University, China  
Email: {fox9.1, xwang8, xtian}@sjtu.edu.cn

Xue Liu  
School of Computer Science  
McGill University, CA  
Email: xueliu@cs.mcgill.ca

**Abstract**—In this paper, we study the capacity and delay scaling laws for cognitive radio network (CRN) with static primary users and heterogeneous mobile secondary users coexist in the unit planar area. The primary network consists of  $n$  randomly and uniformly distributed static primary users (PUs) with higher priority to access the spectrum. The secondary network consists of  $m = (h + 1)n^{1+\epsilon}$  heterogeneous mobile secondary users (SUs) which should access the spectrum opportunistically, here  $h = O(\log n)$  and  $\epsilon > 0$ . Each secondary user moves within a circular area centered at its initial position with a restricted speed. The moving area of each mobile SU is  $n^{-\alpha}$ , where  $\alpha$  is a random variable which follows the discrete uniform distribution with  $h + 1$  different values, ranging from 0 to  $\alpha_0$  ( $\alpha_0 > 0$ ). By allowing the secondary users to relay the packets for primary users, we have proposed a routing and scheduling scheme to fully utilize the mobility heterogeneity of secondary users. We show that the primary network and secondary network can achieve optimal capacity and delay scalings if we increase the heterogeneity of secondary users, i.e., the value of  $h$  and  $\alpha_0$ , until  $h = \Theta(\log n)$  and  $\alpha_0 \geq 1 + \epsilon$ . In this optimal condition, both the primary network and part of the secondary network can achieve almost constant capacity and delay scalings except for poly-logarithmic factor.

## I. INTRODUCTION

The throughput scaling laws for large-scale wireless ad hoc networks has been extensively studied since the seminal work of P. Gupta and P. R. Kumar [1]. They studied the random wireless network with  $n$  static nodes randomly located in the unit area and grouped into source-destination (S-D) pairs for transmission. Under the multi-hop relay algorithm, the network could achieve a per-node throughput of  $\Theta(1/\sqrt{n \log n})$ . By using percolation theory, Franceschetti *et al.* [2] showed the capacity performance could be  $\Theta(1/\sqrt{n})$ , even when the nodes are randomly located in the network area.

In contrast to the static wireless network, the capacity performance could be significantly improved when the nodes are mobile. In [3], M. Grossglauser and D. Tse showed the mobile network could achieve a per-node throughput of  $\Theta(1)$  under the 2-hop relay algorithm. However, this significant improvement of throughput capacity has been achieved at the cost of huge delay, which is proved to be  $\Theta(n)$  by Neely *et al.* [4]. Since mobility could improve the capacity of wireless networks, various mobility models and the corresponding impact of mobility on both the capacity and delay performance of the network have been studied in later works. These include the i.i.d. mobility model [5]; random way-point mobility model [6]; random walk

mobility model [7]; and also the recently proposed restricted mobility model [8] and speed-restricted mobility model [9].

The previous works mainly focus on the capacity and delay scaling for a single network. In recent years, the emergence of the cognitive radio technology has motivated the study of capacity and delay scaling laws of cognitive radio networks (CRN). Due to the coexistence of the licensed primary users, as well as the unlicensed secondary users who can only access the spectrum opportunistically, the study of CRN is more challenging than the traditional single ad-hoc network case. In [10], Jeon *et al.* considered a licensed primary network and a cognitive secondary network coexisting in a planar area, where the primary users have higher node density than secondary users. Based on the assumption that SUs know the location of all PUs, they proposed a scheme that both networks can simultaneously achieve the same throughput scaling law as a stand-alone network. In [11], Yin *et al.* studied a more realistic condition, in which the SUs only know the location of primary transmitters. They have shown that when the density of secondary users is higher than primary users, both the primary network and secondary network would achieve the same capacity scaling as the stand-alone network, with zero outage for the secondary users with high probability. Since most former works have reported the CRN can achieve the same throughput scaling as the stand-alone network, Huang *et al.* [12] characterized the general conditions for the cognitive networks to achieve the same throughput and delay scaling as the stand-alone networks.

Those aforementioned results mainly focused on the static CRN and neglected the possible cooperations between two coexisting networks, so the capacity and delay performance may not be further improved. Later, Gao *et al.* [13] proposed a supportive cognitive network, in which the secondary nodes may route packets for the primary network. They have considered the scenario when secondary users are static or mobile. Particularly, when the secondary users are moving according to the i.i.d. mobility model, the primary users could achieve the per-node throughput scaling of  $\lambda_p = \Theta(1/\log n)$  with the delay scaling of  $D_p = \Theta(1)$ , and the secondary users could achieve the per-node throughput scaling of  $\lambda_s = \Theta(1)$  with the delay scaling  $D_s = \Theta(m)$ . However, this cooperative scheme requires the number of supportive secondary users  $m$  should be at least  $\Theta(n^2)$ , thus casting a heavy burden on its

implementation. In order to reduce the number of secondary users, Wang *et al.* [14] derived a cooperation scheme which could achieve the near-optimal capacity and delay scaling for primary network with less supportive mobile secondary users. This is achieved by dividing the secondary users into  $h$  different layers and the mobile secondary users of different layers are associated with different moving areas. However, in order to regulate the moving area of different layer SUs, this network model requires a strict cell partition scheme, which arbitrarily partitioned the whole network into  $h$  different layer cells, and each layer SUs could only move within the corresponding layer cell.

In this paper, motivated by the fact that cooperation among primary users and secondary users could improve the performance of the CRN, as well as the fact that secondary users could be mobile and have heterogeneous moving areas, we focus on the supportive CRN which has static primary users and heterogeneous mobile secondary users. Specifically, all the  $m = (h + 1)n^{1+\epsilon}$  secondary users ( $h = O(\log n)$ ,  $\epsilon > 0$ ) will follow a local speed-restricted mobility model, in which they move according to the i.i.d. mobility model within a fixed circular area of radius  $R$ . The centers of the circular areas are randomly distributed at the beginning. In addition, the moving area of each mobile SU is  $n^{-\alpha}$ , where  $\alpha$  is a random variable which follows the discrete uniform distribution with  $h + 1$  different values, ranging from 0 to  $\alpha_0$  ( $\alpha_0 > 0$ ). We call this mobility model the heterogeneous speed-restricted mobility model (HSRM). Then we will propose a routing and scheduling scheme which has fully utilized the spatial heterogeneity of the secondary users, so that both the primary packets and secondary packets could reach the destination in a step-wise fashion. Our main contributions are summarized as follow:

- We propose and study the heterogeneous speed-restricted mobility model for secondary users in CRN. Compared with the hierarchical mobility model in [14], which requires the secondary users to move within regularly partitioned different layer cells in the network area, the HSRM is more general and representative since: (1) the centers of the moving area for the secondary users in HSRM are randomly distributed in the network; (2) the parameter  $\alpha_0$  is a random positive value in HSRM. And the hierarchical mobility model in [14] is a specified case of HSRM when  $\alpha_0 = 1 + \epsilon'$ , where  $\epsilon' < \epsilon$ .
- Under HSRM, if we increase the heterogeneity of the secondary users (which means to increase the value of  $h$  and  $\alpha_0$ ), we find the primary network can achieve the near-optimal capacity and delay scaling when  $h = \Theta(\log n)$  and  $\alpha_0 \geq 1 + \epsilon$ . Specifically, under the proposed scheme, the primary network can achieve the per-node throughput  $\lambda_p = \Theta(\frac{1}{\log n})$ , with an average delay of  $D_p = \Theta(\log^4 n)$ .
- Under HSRM, the secondary network can also achieve the near optimal capacity scaling if we increase the heterogeneity of SUs until  $h = \Theta(\log n)$  and  $\alpha_0 \geq 1 + \epsilon$ . Specifically, the secondary network can achieve the per-node throughput of  $\lambda_s = \Theta(\frac{1}{\log^3 n})$ , with an average

delay scaling of  $D_{s,j} = O(\log^4 n)$ , if  $j \geq h^*$ ; and  $D_{s,j} = O(n^{(1+\epsilon-\frac{j\alpha_0}{h})} \log^3 n)$ , if  $j < h^*$ . Here  $j$  denotes the type of destination secondary user and  $h^*$  is the critical relay type, both of them will be defined later. Thus, part of the secondary network can also achieve both the near optimal capacity and delay scaling.

The rest of the paper is organized as follows. In Section II, the system model is defined. In section III, we propose the routing and scheduling scheme. In Section IV and Section V, the capacity and delay scalings for primary network and secondary network are derived respectively. In Section VI, the result is discussed. Finally, we conclude our paper in Section VII.

## II. SYSTEM MODEL

### A. Network Topology

We study a static primary network and a mobile secondary network coexist in the unit square area. The primary network consists of  $n$  randomly and evenly distributed static primary users, which are randomly grouped into source-destination (S-D) pairs one by one.

Moreover, there are  $m = (h + 1)n^{1+\epsilon}$  randomly distributed mobile secondary users with  $h = O(\log n)$  and  $\epsilon > 0$ , which move under the heterogeneous local speed-restricted mobility model defined later. All the secondary users are also randomly grouped into S-D pairs one by one.

The unit square area is divided into non-overlapping small square cells, each covers an area of  $\frac{2 \log N}{N}$  with side length  $r = \sqrt{\frac{2 \log N}{N}}$ , here  $N = n^{1+\epsilon}$ . Every node can only communicate with another node inside the same cell, so the transmission range in our scheme is set to be one cell for both the primary network and secondary network.

### B. Transmission Model

In this work, we only consider path loss of wireless channel and ignore the influence of shadowing or fading for simplicity. Thus, the normalized channel power gain  $g(d)$  is given as

$$g(d) = d^{-\gamma}, \quad (1)$$

where  $d$  denotes the distance of transmission and  $\gamma > 2$  represents the path-loss exponent.

We adopt the Gaussian channel model to regulate the transmission rate, which is a continuous function of the Signal to Interference plus Noise Ratio (SINR). Specifically, the data rate from a primary transmitter  $P_i$  to its receiver  $P_{D(i)}$  is determined by:

$$R(P_i, P_{D(i)}) = \log\left(1 + \frac{P_p g(\|P_i - P_{D(i)}\|)}{N_0 + I_p + I_{sp}}\right), \quad (2)$$

Here,  $P_p$  is the transmission power for the primary nodes, and  $N_0$  is the ambient noise power.  $\|\cdot\|$  denotes the distance for two nodes in the unit area. Moreover,  $I_p$  is the sum interference from all the other concurrent primary transmitters

to the receiver  $P_{D(i)}$ ,  $I_{sp}$  is the sum interference from all the current secondary transmitters to  $P_{D(i)}$ . Suppose there are  $N_p$  and  $N_s$  simultaneous primary and secondary transmitters, then  $I_p$  is determined by:

$$I_p = P_p \sum_{k=1, k \neq i}^{N_p} g(\|P_k - P_{D(i)}\|), \quad (3)$$

Further if we set  $P_s$  to be the transmission power for the secondary nodes, and  $S_k$  ( $1 \leq k \leq N_s$ ) to be the secondary transmitters. Then  $I_{sp}$  is determined by:

$$I_{sp} = P_s \sum_{k=1}^{N_s} g(\|S_k - P_{D(i)}\|). \quad (4)$$

Similarly, the data rate from the secondary transmitter  $S_i$  to its receiver  $S_{D(i)}$  is defined as:

$$R(S_i, S_{D(i)}) = \log\left(1 + \frac{P_s g(\|S_i - S_{D(i)}\|)}{N_0 + I_s + I_{ps}}\right), \quad (5)$$

Here  $I_s$  is the sum interference from all the other simultaneous secondary transmitters, and  $I_{ps}$  represents the total interference from the concurrent primary transmitters.

### C. Mobility Model

The secondary users would move under the heterogeneous speed-restricted mobility model (HSRM). Similar to the local speed-restricted mobility model (LSRM) in [9], the secondary users are uniformly and randomly distributed at the beginning. All the SUs would move within their own circular area centered at their initial positions, according to the i.i.d. mobility model. The radius  $R$  of these circles denotes the restricted speed of the secondary users, and the node positions would be totally reshuffled within their own moving area from one time slot to another.

In HSRM, the moving area of mobile SUs is set to be  $n^{-\alpha}$ , where  $\alpha$  is a random variable which follows the discrete uniform distribution with  $h + 1$  possible different values. Specifically,  $\alpha = 0, \frac{\alpha_0}{h}, \frac{2\alpha_0}{h}, \dots, \frac{(h-1)\alpha_0}{h}, \alpha_0$ , with equal probability  $p = \frac{1}{h+1}$ . Here  $\alpha_0$  is a random positive value and  $h = O(\log_2 n)$ . The SUs with  $\alpha = 0$  would move within the whole network area, while the SUs with other values of  $\alpha$  would only move regionally in the network. Particularly, the nodes with  $\alpha = \alpha_0$  will have the smallest moving area. Thus, the value of  $\alpha_0$  could represent how heterogeneous the SUs are, and we will call  $\alpha_0$  the *heterogeneity factor*.

From the definition of HSRM, if we denote the moving area of different types of SUs by  $A_i$ , then  $A_i = n^{-\frac{i\alpha_0}{h}}$ , where  $0 \leq i \leq h$ . Throughout the paper, we call the SUs with moving area  $A_i$  the  $i$ -th type SU. Since all the  $m$  secondary users are evenly divided into  $h + 1$  different types with different moving area, thus each type consists of  $N = n^{1+\epsilon}$  mobile secondary users.

Moreover, we denote the  $k$ -th secondary users of type  $i$  as  $S_{i,k}$  and its corresponding initial position as  $X_{i,k}$ , where  $0 \leq$

$i \leq h$  and  $1 \leq k \leq N$ . Under the HSRM,  $\|S_{i,k} - X_{i,k}\| \leq R_i$ , where  $R_i = \sqrt{\frac{A_i}{\pi}} = \Theta(n^{-\frac{i\alpha_0}{2h}})$ .

Figure 1 illustrates the moving areas of  $i$ -th type SU  $S_{i,k}$  and  $j$ -th type SU  $S_{j,k}$ .

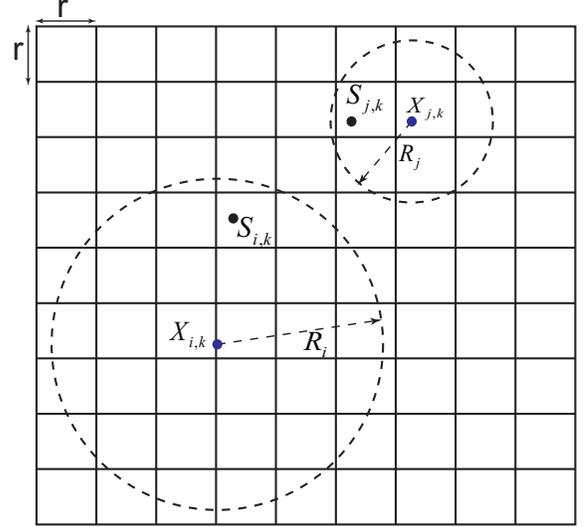


Fig. 1: Cellular structure and moving area of  $S_{i,k}$  and  $S_{j,k}$

### D. Capacity and Delay

The per-node throughput capacity of a S-D pair is defined as the data rate (in bits/time-slot) that each source node can transmit to its destination. For the primary network, we denote its per-node throughput capacity by  $\lambda_p$ , while that of the secondary network is denoted by  $\lambda_s$ .

The delay of S-D pair is defined as the average number of time-slots passed before the packet reaches its destination, after it leaves the source node. For the primary network, we use  $D_p$  to denote its average delay. As for the secondary network, the delay performance will depend on the type of the final destination SU.

Finally, we list some notations in Table I.

TABLE I: Definition of Symbols and Notations

Symbol	Definition
$\alpha_0$	Heterogeneity Factor
$h + 1$	Number of types of SU
$N$	Number of SUs in each type
$h^*$	Critical relay type
$r$	The side length of one cell
$A_i$	The moving area of $i$ -th type SU
$R_i$	The speed or radius of $i$ -th type SU
$\mathcal{P}_i$	The position of $i$ th primary user
$S_{i,k}$	The position of $k$ -th secondary user of type $i$
$X_{i,k}$	The initial position of $k$ -th secondary user of type $i$
Pr	Probability of an event
$\lambda_p$	Per-node throughput of the primary network
$\lambda_s$	Per-node throughput of the secondary network
$D_p$	Delay of primary S-D pairs
$D_{s,k}$	Delay for secondary S-D pairs with a $k$ th-type destination

### III. ROUTING AND SCHEDULING SCHEME

In this section, we would describe the routing and scheduling scheme in our CRN, which would utilize the heterogeneity of mobile secondary users.

#### A. Primary Network Routing Scheme

In our scheme, the secondary users could act as the relay for the primary packets, so there is cooperation among primary users and secondary users. Moreover, since the number of secondary users is larger than primary users in order sense, thus it is reasonable to consider all the primary packets would be relayed by secondary users.

Because the moving area of SUs under HSRM is ranging from globally to regionally, thus our relay algorithm would utilize this mobility heterogeneity to make the packets approach their destination progressively. Specifically, the primary packet would be relayed by a chain of different type secondary users, among which the primary destination node is within the moving area of the relay nodes. This relay procedure starts from a 0 type SU and will continue until the packet is relayed to a particular type SU which is close enough to the primary destination node.

Since the secondary users with larger type would correspond to a smaller area, thus the mobility cannot be exploited if the moving area of certain type SUs is small enough. Consequently, we need to find the maximum type of secondary users that can be exploited to relay the primary packets, which leads to the following definition:

**Definition 1:** (Critical Relay Type) The critical relay type  $h^*$  is denoted by:

$$h^* = \max \{i | R_i \geq 2\sqrt{2}r\} \quad (6)$$

, where  $i = 0, 1, \dots, h$ .

From the above definition, we can derive the value of critical relay type will follow:

$$h^* = \begin{cases} h, & \text{if } \alpha_0 < 1 + \epsilon, \\ \lfloor h(\frac{1+\epsilon}{\alpha_0} - \frac{\log \log n + \log 16\pi(1+\epsilon)}{\alpha_0 \log n}) \rfloor, & \text{if } \alpha_0 \geq 1 + \epsilon. \end{cases} \quad (7)$$

Here,  $\lfloor x \rfloor = \max\{n \in \mathbb{Z} | n \leq x\}$ . And it can be seen that  $h^* = \Theta(h)$ .

By introducing the critical relay type, the primary relay algorithm would utilize the relay nodes from type 0 to type  $h^*$ , thus the number of relay step will be  $h^* + 2$  and the primary relay algorithm is shown in Algorithm 1.

From Algorithm 1, the primary packet could reach the primary destination in a step wise fashion after each relay step. And since the initial positions of the  $(k+1)$ -th type relay node satisfies  $\|X_{k+1, u_{k+1}} - P_j\| < R_{k+1} - \sqrt{2}r$ , thus we can guarantee the cell in which  $P_j$  resides is inside the moving area of the eligible  $(k+1)$ -th type relay. This would ensure successful transmission from the final relay  $S_{h^*, u_{h^*}}$  to  $P_j$ .

The relay step from  $S_{i-1, u_{i-1}}$  to  $S_{i, u_i}$  is shown in Figure 2 ( $1 \leq i \leq h^*$ ). In step 1 of Figure 2,  $S_{i-1, u_{i-1}}$  moves until it meets an eligible  $i$ -th type SU  $S_{i, u_i}$  in the same cell; in step

#### Algorithm 1 Relay Algorithm for Primary Packet $B_p$

**Input:** The primary source node  $P_i$  and destination node  $P_j$

**Output:** The  $h^* + 1$  intermediate secondary relay nodes

- 1:  $P_i$  relay  $B_p$  to a 0 type SU  $S_{0, u_0}$ , when  $S_{0, u_0}$  moves to the same cell as  $P_i$ .
- 2: **for**  $k=0$  to  $(h^* - 1)$  **do**
- 3:  $S_{k, u_k}$  moves within its moving area until it meets  $S_{k+1, u_{k+1}}$  in the same cell, whose initial position satisfies  $\|X_{k+1, u_{k+1}} - P_j\| < R_{k+1} - \sqrt{2}r$  and  $\|X_{k+1, u_{k+1}} - X_{k, u_k}\| < R_k$ .
- 4:  $S_{k, u_k}$  relay  $B_p$  to  $S_{k+1, u_{k+1}}$ .
- 5: **end for**
- 6:  $S_{h^*, u_{h^*}}$  moves within its moving area until it arrives at the same cell as  $P_j$ .
- 7:  $S_{h^*, u_{h^*}}$  relay  $B_p$  to  $P_j$ .

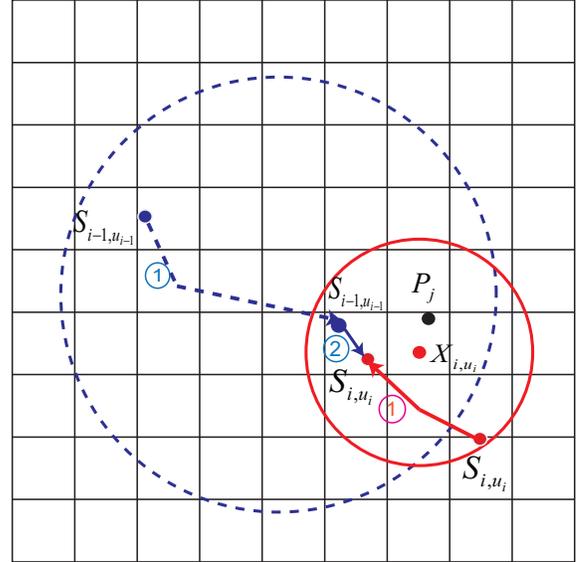


Fig. 2: Relay step from  $S_{i-1, u_{i-1}}$  to  $S_{i, u_i}$  for a primary packet

2, the primary packet is relayed from  $S_{i-1, u_{i-1}}$  to  $S_{i, u_i}$ . The blue dotted circle and red solid circle denotes the moving area of  $S_{i-1, u_{i-1}}$  and  $S_{i, u_i}$  respectively.

After we have derived the primary network relay algorithm, a critical step is to ensure the feasibility of each relay step, which means the number of eligible relay SUs should be larger than 1 with high probability. Before proving the feasibility of the relay algorithm, we would first present the following two lemmas.

**Lemma 1:** Denote  $C_r(x)$  as the circle with radius  $r$ , which is centered at point  $x$ . If two circles  $C_{r_1}(x_1)$  and  $C_{r_2}(x_2)$  satisfies:

$$r_1 \geq r_2, \|x_1 - x_2\| < r_1$$

Then the overlap area of the two circles, denoted as  $C_{r_1}(x_1) \cap C_{r_2}(x_2)$ , satisfies:

$$\frac{\pi}{3}r_2^2 < C_{r_1}(x_1) \cap C_{r_2}(x_2) \leq \pi r_2^2 \quad (8)$$

Lemma 1 can be proved using the basic geometry knowledge and thus we do not provide a specific proof here.

**Lemma 2:** If  $z$  number of nodes are randomly and uniformly distributed in the unit square, then we can ensure at least  $\frac{z\pi R^2}{2\log(1/R)+\pi R^2}$  nodes are placed in a circle with radius  $R$  with high probability.

We refer readers to Theorem 4.1 in [9] for a detailed proof of Lemma 2.

Based on Lemma 1 and Lemma 2, we have the following lemma which guarantees the feasibility of the relay algorithm.

**Lemma 3:** With a random positive heterogeneity factor  $\alpha_0$  and  $h = O(\log_2 n)$ , the number of eligible relay SUs in each step of Algorithm 1 is larger than 1 with high probability.

*Proof:* First we note that each type has  $N = n^{1+\epsilon}$  different secondary users. And since all the type 0 SUs will move globally in the network, thus the eligible number of relay SUs for the first step is  $N$ , which is larger than 1 obviously.

Next, for  $0 \leq k \leq (h^* - 1)$ , consider the step during which the primary packet is relayed from  $S_{k,u_k}$  to a  $(k+1)$ -th type SU  $S_{k+1,u_{k+1}}$ . All the  $(k+1)$ -th type SUs whose initial position satisfy  $\|X_{k+1,u_{k+1}} - P_j\| < R_{k+1} - \sqrt{2}r$  and  $\|X_{k+1,u_{k+1}} - X_{k,u_k}\| < R_k$  can be the eligible relay node. Thus, the initial position for the eligible  $(k+1)$ -th type relay SUs should reside in the overlap area of  $C_{R_k}(X_{k,u_k})$  and  $C_{R_{k+1}-\sqrt{2}r}(P_j)$ , where  $P_j$  is the destination node.

Obviously, for  $0 \leq k \leq (h^* - 1)$ ,  $C_{R_k}(X_{k,u_k})$  and  $C_{R_{k+1}-\sqrt{2}r}(P_j)$  satisfies the condition in Lemma 1. Thus denote  $C_{R_k}(X_{k,u_k}) \cap C_{R_{k+1}-\sqrt{2}r}(P_j)$  as  $A_{k+1}$ , then  $\frac{\pi}{3}(R_{k+1} - \sqrt{2}r)^2 < A_{k+1} \leq \pi(R_{k+1} - \sqrt{2}r)^2$ . Since  $A_{k+1}$  decreases with  $k$ , thus we only need to guarantee there will be at least one eligible  $h^*$ -type relay SU.

The initial positions of eligible  $h^*$ -type relay SUs resides in the region with area  $A_{h^*}$ , where  $\frac{\pi}{3}(R_{h^*} - \sqrt{2}r)^2 < A_{h^*} \leq \pi(R_{h^*} - \sqrt{2}r)^2$ . By the definition of  $h^*$ , it can be proved that:

$$\frac{N \frac{\pi}{3} (R_{h^*} - \sqrt{2}r)^2}{2 \log(1/R_{h^*} - \sqrt{2}r) + \pi (R_{h^*} - \sqrt{2}r)^2} > 1$$

Thus, using Lemma 2, there exists at least one eligible  $h^*$ -type SU resides in  $A_{h^*}$ .

This completes the proof. ■

## B. Secondary Network Routing Scheme

Similar to the primary network routing scheme, the secondary network would also utilize the cooperation among different type of SUs to make the secondary packets approach the destination progressively.

Pick a random secondary source node  $S_{i,k_i}$ , whose corresponding destination node is  $S_{j,k_j}$ , and the packet from  $S_{i,k_i}$  to  $S_{j,k_j}$  is denoted by  $B_{s,j}$ . The number of relay steps for  $B_{s,j}$  will depend on the value of  $j$  and  $h^*$ . Specifically, when  $j > h^*$ ,  $B_{s,j}$  will be relayed from a 0-type relay SU down to a  $h^*$ -type relay SU. On the other hand, when  $j \leq h^*$ ,  $B_{s,j}$  will only be relayed from a 0-type SU down to a  $j$ -th type SU.

Thus, if we denote  $h' = \min(j, h^*)$ , the relay process of  $B_{s,j}$  will take up  $h' + 2$  steps. Due to the mobility of secondary

transmitters and receiver, the secondary relay algorithm is different from the primary network, as shown in Algorithm 2. From Lemma 3, we can also guarantee the feasibility of the secondary relay algorithm.

---

## Algorithm 2 Relay Algorithm for Secondary Packet $B_{s,j}$

---

**Input:** The source node  $S_{i,k_i}$  and destination node  $S_{j,k_j}$

**Output:** The  $h' + 1$  intermediate secondary relay nodes

- 1:  $S_{i,k_i}$  moves within its moving area until it meets a 0-type SU  $S_{0,u_0}$  in the same cell.
  - 2:  $S_{i,k_i}$  relay  $B_{s,j}$  to  $S_{0,u_0}$ .
  - 3: **for**  $k=0$  to  $(h' - 1)$  **do**
  - 4:  $S_{k,u_k}$  moves within its moving area until it meets  $S_{k+1,u_{k+1}}$  in the same cell, whose initial position satisfies  $\|X_{k+1,u_{k+1}} - X_{j,k_j}\| < R_{k+1} - \sqrt{2}r$  and  $\|X_{k+1,u_{k+1}} - X_{k,u_k}\| < R_k$ .
  - 5:  $S_{k,u_k}$  relay  $B_{s,j}$  to  $S_{k+1,u_{k+1}}$ .
  - 6: **end for**
  - 7:  $S_{h',u_{h'}}$  moves within its moving area until it encounters  $S_{j,k_j}$  in the same cell.
  - 8:  $S_{h',u_{h'}}$  relay  $B_{s,j}$  to  $S_{j,k_j}$ .
- 

## C. Primary Scheduling Scheme

The scheduling scheme would choose which primary or secondary S-D pair to be activated in each time slot. In our scheme, the successful transmission is restricted to be within a cell, and each cell can support at most one transmission during one time slot.

After we have defined the routing scheme, we would build up the scheduling scheme for the primary and secondary network. In our scheme, the primary network and secondary network share the same time frame structure. In order to ensure equal opportunity for all the cells to be activated and also limit the interference among concurrent transmissions, a 25-TDMA scheme is adopted which is similar to the scheme in [11]. Specifically, all the cells are divided into 25 subsets according to a  $5 \times 5$  pattern. Each time slot is divided into 25 sub-slots, and the cells of different subsets would be activated during one sub-slot with a round-robin fashion.

From the primary relay algorithm, all the primary packets would be relayed by SUs. Thus the scheduling scheme for the primary network only needs to select the primary transmitter in each time slot, which is described as follow:

*Primary Scheduling Scheme:* During the active period of each cell, randomly select a source PU  $P_i$  in this cell (if there is any). Let  $P_i$  relay a primary packet  $P_p$  to a random 0-type SU  $S_{0,u_0}$ , if there exists 0-type SU in this cell.

## D. Secondary Scheduling Scheme

The secondary network shares the same time frame structure with the primary network, and a 25-TDMA scheme is also adopted by the secondary network.

Despite secondary users are required to relay the primary packets, they still need to control the interference for the

primary users. We would define the *preservation region* so as to keep the interference from the SUs to PUs. To be specific, the preservation region is a square area that contains 9 cells, with the current active primary transmitter at the center cell. Only the secondary users outside any current preservation region can they transmit or relay packets, as shown in Figure 3.

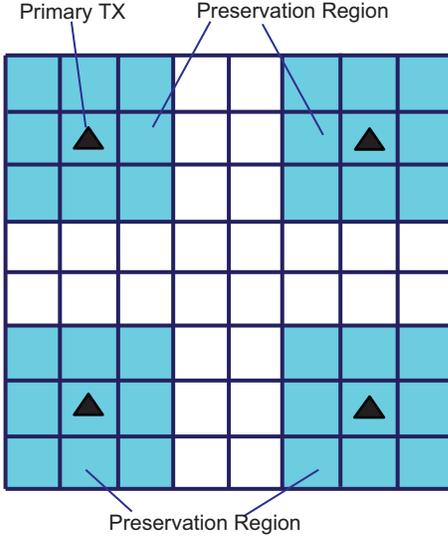


Fig. 3: Illustration of the preservation region

Since the secondary users are required to relay not only secondary packets, but also primary packets, thus we have proposed a scheduling scheme which would guarantee transmission opportunity for both the primary packets and secondary packets.

Specifically, the secondary scheduling scheme would consist of  $2h^* + 3$  phases and each phase consumes one time slot. Note during each phase, if the SUs are inside the current preservation region, they can only buffer the packets, or the SU of 0-type may receive the primary packets when it is in the same cell as a primary user. Otherwise, the secondary network would operate under the following phases:

For  $k = 1, 3, \dots, h^*$ ,

**Phase  $k$ :** During the active period of each cell, all pairs of nodes  $(S_{k-1, u_{k-1}}, S_{k, u_k})$  residing in this cell are eligible for transmission in this phase, if  $S_{k-1, u_{k-1}}$  contains a primary packet  $B_p$  to relay and  $S_{k, u_k}$  can act as the  $k$ -th type relay SU for this  $B_p$ . One of such node pairs would be randomly selected to transmit if the eligible transmission node pairs is non-empty in this cell.

**Phase  $(h^* + 1)$ :** During the active period of each cell, all pairs of nodes  $(S_{h^*, u_{h^*}}, P_j)$  residing in this cell are eligible for transmission in this phase, if  $S_{h^*, u_{h^*}}$  contains a primary packet  $B_p$  destined to  $P_j$ . One of such node pairs would be randomly selected to transmit if the eligible transmission node pairs is non-empty in this cell.

**Phase  $(h^* + 2)$ :** During the active period of each cell, randomly select a source SU  $S_{i, k_i}$  in this cell (if there is any). Let  $S_{i, k_i}$  relay a secondary packet  $B_{s, j}$  to a random 0-type SU  $S_{0, u_0}$ , if there exists any 0-type SU in this cell.

For  $k = 1, 3, \dots, h^*$ ,

**Phase  $(h^* + 2 + k)$ :** During the active period of each cell, two types of SU pairs residing in this cell are eligible for transmission in this phase: (1) node pair  $(S_{k-1, u_{k-1}}, S_{k-1, j_{k-1}})$  which satisfies:  $S_{k-1, u_{k-1}}$  contains a secondary packet  $B_{s, k-1}$  destined to  $S_{k-1, j_{k-1}}$ ; or (2) node pair  $(S_{k-1, u_{k-1}}, S_{k, u_k})$  which satisfies:  $S_{k-1, j_{k-1}}$  contains a secondary packet  $B_{s, k'}$  ( $k \leq k' \leq h^*$ ) to relay and  $S_{k, u_k}$  can act as the  $k$ -th type relay SU for  $B_{s, k'}$ . One of such node pairs would be randomly selected to transmit if the eligible transmission node pairs is non-empty in this cell.

**Phase  $(2h^* + 3)$ :** During the active period of each cell, all pairs of nodes  $(S_{h^*, u_{h^*}}, S_{j, k_j})$  ( $h^* \leq j \leq h$ ) residing in this cell are eligible for transmission in this phase, if  $S_{h^*, u_{h^*}}$  contains a secondary packet  $B_{s, j}$  destined to  $S_{j, k_j}$ . One of such node pairs would be randomly selected to transmit if the eligible transmission node pairs is non-empty in this cell.

#### IV. CAPACITY AND DELAY SCALINGS FOR THE PRIMARY NETWORK

In this section, we would first evaluate the capacity scaling of the primary network, and then study the delay performance for the primary network.

##### A. Capacity Performance

In this part, we first give the following lemmas:

**Lemma 4** (*Ji et al. [15]*): Assume  $x$  nodes are placed into  $y$  equal-sized areas randomly, evenly and independently. Let  $Z(x, y)$  be the random variable that

$$Z(x, y) = \begin{cases} \Theta\left(\frac{x}{y}\right), & \text{if } x \gg y \log y, \\ \Theta(\log y), & \text{if } x = cy \log y \text{ for some constant } c, \\ \Theta\left(\frac{\log y}{\log \frac{y \log y}{x}}\right), & \text{if } \frac{y}{\text{poly}(\log(y))} \leq x \ll y \log y, \\ \Theta\left(\frac{\log y}{\log \frac{y}{x}}\right), & \text{if } x < \frac{y}{\log y}. \end{cases} \quad (9)$$

**Lemma 5:** At any moment, there are at most  $\Theta(1)$  PUs in any cell with high probability.

*Proof:* In our scheme,  $n$  PUs are randomly and evenly distributed into  $\frac{1}{r^2} = \frac{N}{2 \log N}$  cells. According to the last condition in Lemma 4, the maximum number of PUs in any lattice should be  $\Theta(1)$  ■

**Lemma 6:** At any moment, for  $0 \leq i \leq h$ , the number of  $i$ -th type secondary users in each cell is at most  $\Theta(\log n)$  with high probability.

*Proof:* There are  $N = n^{1+\epsilon}$  number of SUs in each type. Since their initial positions are randomly and evenly distributed in the network and all the SUs moving according to the i.i.d. mobility model in their moving area, thus we can use the result of Lemma 4. According to condition 2 of Lemma 4, the maximum number of  $i$ -th type SU in each cell is  $\max(n_2) = \Theta(\log N) = \Theta(\log n)$ . ■

**Lemma 7:** During the routing process of primary packets, the primary transmitters and secondary relay nodes can support a constant data rate in each cell.

*Proof:* In order to prove that every primary transmitter or secondary relay node could support a constant data rate, we divide the whole routing process into three parts: input, relay and output.

First consider the input process, during which the primary transmitters would relay the primary packet to a 0-type SU  $S_{0,u_0}$  in the same cell. Since a 25-TDMA scheme is adopted, so the data rate of the primary transmitter could be given as:

$$R(P_i, S_{0,u_0}) = \frac{1}{25} \log\left(1 + \frac{P_p g(\|P_i - S_{0,u_0}\|)}{N_0 + I_p + I_{sp}}\right), \quad (10)$$

Where  $P_i$  is the primary transmitter and  $S_{0,u_0}$  is the 0-type secondary relay node.  $\frac{1}{25}$  is introduced by dividing the primary time slot into 25 TDMA sub-slots. Since we have restricted the transmission range to be within a cell, so the transmission power for the primary transmitter is  $P_p = Pr^\gamma$ , where  $P$  is a constant and  $r$  is the side length of a cell.

Since the transmission is within a lattice, so  $\|P_i - S_{0,u_0}\| \leq \sqrt{2}l$ , thus:

$$\begin{aligned} P_p g(\|P_i - S_{0,u_0}\|) &= Pr^\gamma (\|P_i - S_{0,u_0}\|)^{-\gamma} \\ &\geq Pr^\gamma (\sqrt{2}r)^{-\gamma} \\ &= P(\sqrt{2})^{-\gamma} \end{aligned} \quad (11)$$

Then we should consider  $I_p$ , which is the sum interference from the concurrent primary transmitters. Since the 25-TDMA scheme is adopted, there would be at most 8 concurrent primary transmitters with a distance of at least  $4r$  from  $S_{0,u_0}$ , and 16 primary transmitters with a distance at least  $9r$  and so on. So the  $I_p$  is bounded by:

$$\begin{aligned} I_p &= \sum_{k=1, k \neq i}^{N_p} P_p g(\|P_k - S_{0,u_0}\|) \\ &\leq \sum_{k=1, k \neq i}^{\infty} 8k[(5k-1)r]^{-\gamma} Pr^\gamma \\ &= 8P \sum_{k=1, k \neq i}^{\infty} \frac{k}{(5k-1)^\gamma} \\ &= R_1 \end{aligned} \quad (12)$$

Here  $R_1$  is a constant. This is because the path-loss exponent  $\gamma$  is larger than 2, so the infinite series of Equation (12) converges to a finite number.

Next we should bound the sum interference from the concurrent secondary transmitters. In the input process, a minimum distance of  $r$  could be guaranteed from all the concurrent secondary transmitters to  $S_{0,u_0}$ . And since the secondary network also employs the 25-TDMA scheme, thus  $I_{sp}$  is bounded by:

$$\begin{aligned} I_{sp} &= \sum_{k=1, k \neq i}^{N_s} P_s g(\|S_k - S_{0,u_0}\|) \\ &\leq \sum_{k=1, k \neq i}^{\infty} 8k[(5k-4)r]^{-\gamma} Pr^\gamma \\ &= 8P \sum_{k=1, k \neq i}^{\infty} \frac{k}{(5k-4)^\gamma} \\ &= R_2 \end{aligned} \quad (13)$$

Here  $R_2$  is a constant and the reason that equation (13) converges is the same as equation (12).

Combining all the three terms together, we could show that the transmission rate for the input process is bounded by

$$\begin{aligned} R(P_i, S_{0,u_0}) &= \frac{1}{25} \log\left(1 + \frac{P_p g(\|P_i - S_{0,u_0}\|)}{N_0 + I_p + I_{sp}}\right) \\ &\geq \frac{1}{25} \log\left(1 + \frac{P(\sqrt{2})^{-\gamma}}{R_1 + R_2}\right) \\ &= C_1 \end{aligned} \quad (14)$$

After we have considered the input process, we would turn to the relay process for the primary packet, which correspond to Phase 1 through Phase  $h^*$  of the secondary scheduling scheme. Specifically, in the relay process, the data rate is regulated by:

$$R(S_{i,u_i}, S_{i+1,u_{i+1}}) = \frac{1}{25} \log\left(1 + \frac{P_s g(\|S_{i,u_i} - S_{i+1,u_{i+1}}\|)}{N_0 + I_s + I_{ps}}\right), \quad (15)$$

Here, for  $0 \leq i \leq h^* - 1$ ,  $S_{i,u_i}$  is the intermediate relay node for the primary packet. Since we have adopted the 25-TDMA scheme and the preservation region, thus by using similar method as the input process,  $I_s$  and  $I_{ps}$  can be bounded by a constant. Consequently, the data rate for the relay process is also a constant  $C_2$ .

Finally, we would consider the output process, during which the  $h^*$ -type layer relay SU would transmit the packet to the primary destination node  $P_j$ . The transmission rate is regulated by:

$$R(S_{h^*,u_{h^*}}, P_j) = \frac{1}{25} \log\left(1 + \frac{P_s g(\|S_{h^*,u_{h^*}} - P_j\|)}{N_0 + I_s + I_{ps}}\right), \quad (16)$$

Like the previous processes, the transmission rate in the output process is also a constant  $C_3$ .

This finishes the proof.  $\blacksquare$

According to the primary scheduling scheme and the results of Lemma 5 and Lemma 7, we can derive the following theorem that counts the per-node throughput of the primary network.

**Theorem 1:** Under the proposed primary relay algorithm, the primary network can achieve the following average per-node throughput with high probability:

$$\lambda_p = \Theta\left(\frac{1}{h}\right), \quad (17)$$

*Proof:* During the routing process of the primary packet, since every primary transmitter and secondary relay nodes could support a constant data rate, thus we assume any node could transmit or relay at a rate of  $R$  bits per time-slot. We divide the whole routing process into three parts: input, relay and output.

During input process, the primary packet is relayed from the primary transmitter to a 0 type relay SU. Since there are at most  $\Theta(1)$  within each cell, thus the per-node time-average input rate of PUs should be  $\Theta(1)$  with high probability. Since there are  $\Theta(n)$  S-D pairs, thus the aggregated input rate of all PUs should be  $\Theta(n)$ .

During the relay process, the primary packet is relayed from the  $(i-1)$ -th type relay SU to an  $i$ -th type SU, for  $1 \leq i \leq h^*$ . Since the transmission range is set to be  $r^2$  and each relay SU can support a constant data rate in each cell, thus the whole network can support a maximum aggregated relay rate of  $\Theta(\frac{N}{2 \log N})$ , which is larger than the aggregated input rate. This indicates the relay process could forward an aggregated  $\Theta(n)$  bits per time-slot. Since each relay process consumes  $\frac{1}{2h^*+3}$  fraction of the secondary scheduling cycle, thus the primary network can achieve per-node time-average relay rate of  $\Theta(\frac{1}{h^*})$ .

Finally, during the output process, the primary packet is transmitted to the destination primary user. The output process is similar to the input process, and  $\Theta(n)$  bits can be forwarded to the primary destination nodes during the output process. Since each output process consumes  $\frac{1}{2h^*+3}$  fraction of the secondary scheduling cycle, thus the output process can achieve the per-node time-average output rate of  $\Theta(\frac{1}{h^*})$ .

Consequently, the primary network can achieve the per-node throughput of  $\lambda_p = \min(\Theta(\frac{1}{h^*}), \Theta(1)) = \Theta(\frac{1}{h^*})$ . ■

### B. Delay Performance

In order to derive the delay performance of the primary network, we will first give the following lemmas.

**Lemma 8:** For random positive value  $l$  and  $N$  such that  $0 < l < 1$  and  $N \geq \Theta(1)$ , the following equation holds with high probability:

$$1 - (1-l)^N = \min(\Theta(1), \Theta(Nl)) \quad (18)$$

*Proof:* We consider the following three conditions depending on the value of  $l$  and  $N$ :

(1)  $l = \Theta(1)$ . In this case,  $1 - (1-l)^N > l = \Theta(1)$  holds, thus Equation (18) is satisfied.

(2)  $l = o(1)$  and  $Nl \geq \Theta(1)$ . Since  $l$  can be arbitrarily small when  $n$  is large enough, thus  $-2l < \ln(1-l) < -l$  will hold. Consequently:

$$1 - \exp(-Nl) < 1 - (1-l)^N < 1 - \exp(-2Nl),$$

Since  $Nl \geq \Theta(1)$ , as  $1 - \exp(-Nl) > 1 - e^{-1}$ . So it is clear that  $1 - (1-l)^N = \Theta(1)$ .

(3)  $l = o(1)$  and  $Nl = o(1)$ . In this case, it can be proved that  $x - \frac{x^2}{2} < 1 - \exp(-x) < x$  always holds when  $x > 0$ . Thus,  $Nl - \frac{(Nl)^2}{2} < 1 - (1-l)^N < Nl$  will be satisfied. Thus,  $1 - (1-l)^N = \Theta(Nl)$ .

This completes the proof. ■

By using the results of Lemma 8 and Lemma 6, we can have the following lemma that counts the delay for the primary network.

**Theorem 2:** Under the proposed primary relay algorithm, the primary network can achieve the following average delay with high probability:

$$D_p = \Theta(hn^{(1+\epsilon-h^*\frac{\alpha_0}{n})} + h^2n^{\frac{\alpha_0}{n}} \log^2 n) \quad (19)$$

*Proof:* Consider the routing process for a primary packet  $B_p$ , which is sent from  $P_i$  to  $P_j$ . Similar to Section VI-C [8], we would evaluate the average number of time slots for  $B_p$  to be successfully relayed in each process. During the routing process, we would ignore the possible contention for transmission among different packets in each node. The routing process is also divided into input, relay and output process.

(a) First we analyze the input process, in which primary source node  $P_i$  will transmit to an eligible 0 type SU. A successful relay from  $P_i$  to the eligible 0 type relay node will happen when the two conditions hold: (1) an eligible 0 type relay node  $S_{0,u_0}$  is found in the same cell as  $P_i$ ; (2) among all the eligible transmission pairs in this cell,  $P_i$  is the transmitter of the selected pair and an eligible 0 type relay SU is the selected receiver. If we denote the conditions as  $a_0$  and  $b_0$ , then we will calculate the probability of them respectively.

Since all the  $N = n^{1+\epsilon}$  0 type SUs can be eligible relay for  $P_i$ , thus for an arbitrary 0 type SU  $S_{0,u_0}$ , we first calculate the probability that  $S_{0,u_0}$  moves to the cell in which  $P_i$  resides, denoted by  $p_s^0$ . Since  $S_{0,u_0}$  moves in the whole network with i.i.d. mobility model, thus  $p_s^0 = r^2$ . If we consider all the 0 type SUs together and use the result of Lemma 8, we have  $\Pr(a_0) = 1 - (1 - p_s^0)^N = \Theta(1)$ .

In addition, since there are at most  $\Theta(1)$  primary users and  $\Theta(\log n)$  0 type SUs in any cell, thus  $\Pr(b_0|a_0) = \Theta(\frac{1}{\log n})$ . So the probability for successful relay is  $p_0 = \Pr(a_0)\Pr(b_0|a_0) = \Theta(\frac{1}{\log n})$ . Thus, the delay for the input process is  $D_0 = \Theta(1/p_0) = \Theta(\log n)$ .

(b) Next we consider the relay and output process. For  $1 \leq i \leq h^* + 1$ , suppose  $B_p$  is currently held by  $S_{i-1,u_{i-1}}$ , then a successful relay of  $B_p$  during the relay and output process will happen when the following three conditions happen: (1) the current slot is assigned to Phase  $i$  of the secondary scheduling scheme; (2) an eligible  $i$ -th type relay node, or the destination node resides in the same cell as  $S_{i-1,u_{i-1}}$  in the current time slot; (3) For  $1 \leq i \leq h^*$ , among all the eligible transmission pairs in this cell,  $S_{i-1,u_{i-1}}$  is the transmitter of the selected pair and an eligible  $i$ -th type relay SU is the selected receiver. While for  $i = h^* + 1$ ,  $(S_{h^*,u_{h^*}}, P_j)$  is the selected transmission pair. We denote the three conditions as  $t_i$ ,  $a_i$  and  $b_i$  respectively.

Since a complete scheduling cycle of the secondary network consists of  $2h^* + 3$  phases, thus  $\Pr(t_i) = \frac{1}{2h^*+3} = \Theta(\frac{1}{h})$ .

Then we will calculate  $\Pr(a_i)$  for relay process, i.e.,  $1 \leq i \leq h^*$ . If we denote  $\mathcal{T}_i$  as the set of  $i$ -th type SUs that can relay  $B_p$  from  $S_{i-1,u_{i-1}}$ , and  $\mathcal{Q}$  as the set of cells inside the moving

range of the SUs in  $\mathcal{T}_i$ , i.e., the cells inside  $C_{R_{i-1}}(X_{i-1, u_{i-1}}) \cap C_{2R_{i-1}-\sqrt{2}r}(P_j)$ . Thus,

$$\Pr(a_i) = \sum_Q \Pr(S_{i-1, u_{i-1}} \in Q) \left(1 - \prod_{S_{i,k} \in \mathcal{T}_i} (1 - \Pr(S_{i,k} \in Q))\right) \quad (20)$$

Since  $R_i \geq 2\sqrt{2}r$ , so that  $\Theta(2R_i - \sqrt{2}r) = \Theta(R_i)$ . The upper bound of  $p_a^i$  can be calculated as follow:

$$\begin{aligned} \Pr(a_i) &\leq \sum_Q \Pr(S_{i-1, u_{i-1}} \in Q) \\ &= \Theta\left(\frac{(2R_i - \sqrt{2}r)^2}{r^2}\right) \Theta\left(\frac{r^2}{R_{i-1}^2}\right) \\ &= \Theta\left(\frac{R_i^2}{r^2}\right) \Theta\left(\frac{r^2}{R_{i-1}^2}\right) \\ &= \Theta\left(n^{-\frac{\alpha_0}{h}}\right) \end{aligned} \quad (21)$$

Next, if we denote  $K$  as the set of cells inside  $C_{R_{i-1}}(X_{i-1, u_{i-1}}) \cap C_{R_{i-1}-\sqrt{2}r}(P_j)$ , then  $K \in Q$ . For any cell  $a \in K$ , we can guarantee that on average,  $a$  is within the moving area of  $\Theta(N(R_i - \sqrt{2}r)^2)$  SUs inside  $\mathcal{T}_i$ . Thus we can calculate the lower bound of  $\Pr(a_i)$  as:

$$\begin{aligned} \Pr(a_i) &\geq \sum_K \Pr(S_{i-1, u_{i-1}} \in K) \left(1 - \prod_{S_{i,k} \in \mathcal{T}_i} (1 - \Pr(S_{i,k} \in K))\right) \\ &\geq \sum_K \Theta\left(\frac{r^2}{R_{i-1}^2}\right) \min(\Theta(1), \Theta(N(R_i - \sqrt{2}r)^2 \frac{r^2}{R_i^2})) \\ &= \Theta\left(\frac{(R_i - \sqrt{2}r)^2}{r^2}\right) \Theta\left(\frac{r^2}{R_{i-1}^2}\right) \Theta(1) \\ &= \Theta\left(n^{-\frac{\alpha_0}{h}}\right) \end{aligned} \quad (22)$$

From Equation (21) and Equation (22), we can conclude  $\Pr(a_i) = \Theta\left(n^{-\frac{\alpha_0}{h}}\right)$ , for  $1 \leq i \leq h^*$ .

As to the output process,  $\Pr(a_{h^*+1}) = r^2 n^{\frac{h^* \alpha_0}{h}}$ , since the cell in which  $P_j$  resides is totally in the moving area of  $S_{h^*, u_{h^*}}$ .

Finally we will calculate the conditional probability  $\Pr(b_i | t_i a_i)$ . During the relay process, since the number of each type SUs in every cell does not exceed  $\Theta(\log n)$ , thus under the pessimistic assumption that all the  $(i-1)$ -th type SUs are holding the primary packet to transmit and all the  $i$ -th type SUs are able to receive the packet, then  $\Pr(b_i)$  is lower bounded by  $\Pr(b_i | t_i a_i) = \Theta\left(\frac{1}{\log^2 n}\right)$ , for  $1 \leq i \leq h^*$ . During the output process, since there are  $\Theta(\log n)$  number of  $h^*$ -type relay SUs and  $\Theta(1)$  number of receiving PUs. Thus the probability for  $(S_{h^*, u_{h^*}}, P_j)$  to be the selected transmission pair is  $\Pr(b_{h^*+1} | t_{h^*+1} a_{h^*+1}) = \Theta\left(\frac{1}{\log n}\right) \Theta(1) = \Theta\left(\frac{1}{\log n}\right)$ .

Consequently, for  $1 \leq i \leq h^* + 1$ , the relay or output step will take an average delay of

$$\begin{aligned} D_i &= \Theta\left(\frac{1}{\Pr(t_i a_i b_i)}\right) \\ &= \Theta\left(\frac{1}{\Pr(t_i) \Pr(a_i) \Pr(b_i | t_i a_i)}\right) \\ &= \begin{cases} \Theta(hn^{\frac{\alpha_0}{h}} \log^2 n), & \text{for } 1 \leq i \leq h^*, \\ \Theta(hn^{(1+\epsilon-\frac{h^* \alpha_0}{h})}), & i = h^* + 1. \end{cases} \end{aligned} \quad (23)$$

Combining all the delay for each relay step, and consider the value of  $h^*$ , which is shown in Equation (7), we can derive the delay performance for the primary network is:

$$\begin{aligned} D_p &= \sum_{i=0}^{h^*+1} D_i \\ &= \Theta(hn^{(1+\epsilon-\frac{h^* \alpha_0}{h})} + h^2 n^{\frac{\alpha_0}{h}} \log^2 n) \end{aligned} \quad (24)$$

## V. CAPACITY AND DELAY SCALINGS FOR THE SECONDARY NETWORK

In this section, we would evaluate the capacity and delay performance for the secondary network.

The secondary network is different from the primary network because the secondary users should access the spectrum opportunistically, i.e., only when the secondary users are outside the current preservation regions can they transmit or relay the packets. According to our scheme, a portion of secondary users would inevitably jump into the preservation region, and the 0-type SUs which jumps into the preservation region may receive the primary packets. But due to the node density of each type secondary users is higher than primary users in order sense, as well as the choice of cell size, the portion of SUs that fall into preservation region will approach 0 with high probability. This fact is already observed in many previous works, such as [10]. Thus, the introduction of preservation region will not degenerate the transmission opportunity for secondary users with high probability.

### A. Capacity Performance

**Lemma 9:** During the routing process of secondary packets, the secondary transmitters and relay nodes can support a constant data rate in each cell.

The proof of Lemma 9 is similar to that of Lemma 7, thus we do not repeat it here.

Combining the result of Lemma 9 and Lemma 6, we can derive the following theorem:

**Theorem 3:** Under the proposed secondary relay algorithm, the secondary network can achieve the following per-node throughput with high probability:

$$\lambda_s = \Theta\left(\frac{1}{h^2 \log n}\right) \quad (25)$$

*Proof:* Since the number of each type secondary users does not exceed  $\Theta(\log n)$ , thus the number of source SUs in the input process does not exceed  $\Theta(h \log n)$ . And since the input process consumes  $\frac{1}{2h^*+3}$  fraction of the secondary scheduling cycle, thus the time-average per-node input rate of SUs is  $\Theta\left(\frac{1}{h^2 \log n}\right)$ .

During the relay and output process, since at most  $\Theta(\log n)$  SUs of each type reside in the cell, thus a time-average relay and output rate of  $\Theta\left(\frac{1}{h \log n}\right)$  is achievable.

Consequently, the secondary network can achieve the per-node throughput of  $\lambda_s = \min(\Theta\left(\frac{1}{h^2 \log n}\right), \Theta\left(\frac{1}{h \log n}\right)) = \Theta\left(\frac{1}{h^2 \log n}\right)$ .

## B. Delay Performance

In this part, we will have the following theorem that counts the average delay for the secondary network.

**Theorem 4:** Under the proposed secondary relay algorithm, the secondary network can achieve the following average delay with high probability:

$$D_{s,j} = O(h^2 \log n + h^2 n^{\frac{\alpha_0}{h}} \log^2 n + h^2 n^{(1+\epsilon-\frac{h'\alpha_0}{h})} \log n), \quad (26)$$

where  $j$  denotes the type of the destination node, and  $h' = \min(h^*, j)$ , which is defined in Algorithm 2.

*Proof:* Consider the relay process of the secondary packet  $B_{s,j}$ , which is sent from  $S_{i,k_i}$  to  $S_{j,k_j}$ . We will also calculate the probability for  $B_{s,j}$  to be successfully relayed. The relay process of  $B_{s,j}$  consumes  $h' + 2$  steps, for  $0 \leq i \leq h' + 1$ , a successful relay of  $B_{s,j}$  will happen if the following three conditions hold simultaneously: (1) the current time slot is assigned to the Phase ( $h^* + 2 + i$ ) of the secondary scheduling scheme; (2) an eligible  $i$ -th type relay SU, or the destination node is found in the same cell as the secondary users which holds the packet at the current time slot; (3) for  $0 \leq i \leq h'$ , among all the eligible transmission pairs in the cell, the SU which holds  $B_{s,j}$  is the transmitter of the selected pair and an eligible  $i$ -th type relay SU is the receiver of the selected pair; while for  $i = h' + 1$ ,  $(S_{h',u_{h'}}, S_{j,k_j})$  is the selected transmission pair. Denote the three conditions by  $t_i$ ,  $a_i$  and  $b_i$ . Clearly,  $\Pr(t_i) = \frac{1}{2h^*+3}$ .

Next we consider the value of  $\Pr(a_i)$ . Obviously the intermediate relay process is the same as the primary network, and the input and output is different from the primary user due to the SU mobility. Using the similar method as the proof of Lemma 2, we can calculate the value of  $\Pr(a_i)$  as follow:

$$\Pr(a_i) = \begin{cases} \Theta(1), & i = 0, \\ \Theta(n^{-\frac{\alpha_0}{h}}), & \text{for } 1 \leq i \leq h', \\ \Theta(n^{\frac{h'\alpha_0}{h} - (1+\epsilon)} \log n), & i = h' + 1. \end{cases} \quad (27)$$

Finally we will calculate the conditional probability  $\Pr(b_i|t_i a_i)$ . By the definition of condition  $b_i$ , and considering the number of each type SUs do not exceed  $\Theta(\log n)$  in every cell, we can calculate  $\Pr(b_i|t_i a_i)$  as;

$$\Pr(b_i|t_i a_i) = \begin{cases} \Theta(\frac{1}{h \log n}), & i = 0, \\ \Omega(\frac{1}{\log^2 n}), & \text{for } 1 \leq i \leq h', \\ \Omega(\frac{1}{h \log^2 n}), & i = h' + 1. \end{cases} \quad (28)$$

Consequently, we can get:

$$\Pr(t_i a_i b_i) = \begin{cases} \Theta(\frac{1}{h^2 \log n}), & i = 0, \\ \Omega(\frac{1}{h \log^2 n} n^{-\frac{\alpha_0}{h}}), & \text{for } 1 \leq i \leq h', \\ \Omega(\frac{1}{h^2 \log n} n^{\frac{h'\alpha_0}{h} - (1+\epsilon)}), & i = h' + 1. \end{cases} \quad (29)$$

Since the delay for each step is  $D_i = \Theta(\frac{1}{\Pr(t_i a_i b_i)}) = \Theta(\frac{1}{\Pr(t_i)})\Theta(\frac{1}{\Pr(a_i)})\Theta(\frac{1}{\Pr(b_i|t_i a_i)})$ , for  $0 \leq i \leq h' + 1$ . Thus by adding them up, we can get the result of the  $D_{s,j}$ . ■

## VI. DISCUSSION

In this section, we will discuss how mobility heterogeneity of the secondary users can affect the capacity and delay scalings of the CRN. Specifically,  $\alpha_0$  measures the range of the mobility heterogeneity and  $h$  measures the diversity of mobility heterogeneity. The increase in  $\alpha_0$  and  $h$  will lead to the increase of mobility heterogeneity of SUs. Thus we can find which value of  $\alpha_0$  and  $h$  will lead to the optimal capacity and delay performance of the primary and secondary network.

### A. Optimal Performance of Primary Network

In our scheme,  $\alpha_0$  is a random positive value, and  $h = O(\log n)$ .

If  $h = \Theta(1)$ , denote  $\alpha_{th} = \frac{1+\epsilon-\frac{2 \log \log n}{\log n}}{1+1/h}$ , then primary network could achieve the following average delay:

$$D_p = \begin{cases} \Theta(n^{\frac{\alpha_0}{h}} \log^2 n), & \text{if } \alpha_0 \geq \alpha_{th}, \\ \Theta(n^{(1+\epsilon-\alpha_0)}), & \text{if } \alpha_0 < \alpha_{th}. \end{cases} \quad (30)$$

with the per-node throughput  $\lambda_p = \Theta(1)$ .

In this case, the primary network can achieve the constant per-node throughput, which is theoretically optimal. As for the delay performance, different  $\alpha_0$  leads to different delay. However, from Equation (30),  $D_p$  can only achieve sub-optimal condition, since  $D_p = \omega(\text{polylog}(n))$ . The delay-capacity tradeoff curve when  $h = \Theta(1)$  is shown in the blue dotted curve of Figure 4, and note that  $\alpha_{th} = \frac{1+\epsilon}{1+1/h}$  with high probability.

If  $h = \Theta(\log n)$ , denote  $\alpha'_{th} = 1 + \epsilon - \frac{3 \log \log n}{\log n}$ . We should note that in this case,  $\Theta(n^{\frac{\alpha_0}{h}}) = \Theta(c^{\alpha_0}) = \Theta(1)$ , where  $c$  is a constant. Thus the primary network could achieve the following average delay:

$$D_p = \begin{cases} \Theta(\log^4 n), & \text{if } \alpha_0 \geq \alpha'_{th}, \\ \Theta(n^{(1+\epsilon-\alpha_0)} \log n), & \text{if } \alpha_0 < \alpha'_{th}. \end{cases} \quad (31)$$

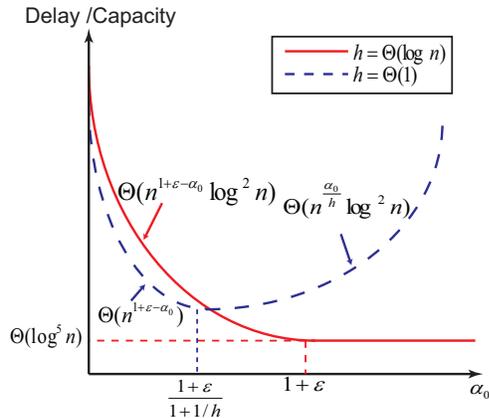
with the per-node throughput  $\lambda_p = \Theta(\frac{1}{\log n})$ .

In this case, the primary network can achieve the near-optimal capacity performance. And the delay performance can be improved when  $\alpha_0$  increases from 0 to  $\alpha'_{th}$ . Specifically, for  $\alpha_0 \geq \alpha'_{th}$ , the delay performance is near optimal, which is  $D_p(\min) = \Theta(\log^4 n)$ .

The delay-capacity tradeoff curve when  $h = \Theta(\log n)$  is plotted in the red-solid curve in Figure 4. Since  $\alpha'_{th} = 1 + \epsilon$  when  $n$  approaches infinity. It is clearly that when  $\alpha_0 \geq 1 + \epsilon$  and  $h = \Theta(\log n)$ , the primary network can achieve a near-optimal delay-capacity tradeoff. From the perspective of delay-capacity tradeoff, the previous results indicate the increase of heterogeneity of SUs will improve the delay-capacity tradeoff, until the near-optimal tradeoff is obtained.

TABLE II: Comparison of optimal scalings of HSRM with other mobility models for SU

Reference	SU Mobility	PU Throughput	PU Delay	SU Throughput	SU Delay	Number of SUs
S. Cui [13]	i.i.d.	$\Theta(1/\log n)$	$\Theta(1)$	$\Theta(1)$	$\Omega(n^2)$	$\Omega(n^2)$
X. Wang [14]	hierarchical	$\Theta(1/\log n)$	$\Theta(\log^2 n)$	$\Theta(1/n^\delta)$	$\Omega(\log^2 n)$	$O(n^{1+\delta'})$
This paper	HSRM	$\Theta(1/\log n)$	$\Theta(\log^4 n)$	$\Theta(1/\log^3 n)$	$O(\log^4 n)$	$O(n^{1+\delta''})$

Fig. 4: Relation between Delay/Capacity and the heterogeneity factor  $\alpha_0$  for primary network

### B. Optimal Performance of Secondary Network

Using similar method as the previous part, we can show the secondary network can achieve the optimal performance when  $h = \Theta(\log n)$  and  $\alpha_0 \geq 1 + \epsilon$ . Specifically, the secondary network could achieve the following average delay:

$$D_{s,j} = \begin{cases} O(\log^4 n), & \text{if } j \geq h^*, \\ O(n^{(1+\epsilon-\frac{j\alpha_0}{h})} \log^3 n), & \text{if } j < h^*. \end{cases} \quad (32)$$

with per-node throughput of  $\lambda_s = \Theta(\frac{1}{\log^3 n})$ .

Under optimal condition, the secondary network also has the potential to achieve the near-optimal delay-capacity tradeoff, when the type of destination SU  $j$  satisfies  $j \geq h^*$ . As for the nodes with  $j < h^*$ , the delay can be reduced if  $\alpha_0$  increases.

Combining the result of the primary network and secondary network, we can conclude that the increase of mobility heterogeneity for SUs will improve the delay-capacity tradeoff for both primary and secondary network, so that both network can achieve near-optimal capacity and delay scaling laws.

Finally, the optimal capacity and delay scalings of HSRM is compared with the previous works in Table II, where  $\delta, \delta', \delta'' > 0$ . From Table II, a near-constant capacity and delay scalings for the primary network can be achieved in all these works. Compared with [13], the HSRM achieves better delay scaling for secondary network and requires less secondary users. While compared with [14], HSRM can achieve better capacity scalings for secondary network and is more general and flexible than the hierarchical mobility model in [14].

## VII. CONCLUSION

In this paper, we study the impact of mobility heterogeneity of secondary users in cognitive radio network. We propose the

heterogeneous speed-restricted mobility model for the SUs, and put forward a corresponding routing and scheduling scheme which exploits the heterogeneity of SUs. In particular, we show that when the heterogeneity of secondary users increases, the delay-capacity tradeoff for both primary and secondary network can be improved. Under the optimal condition, the primary network and part of the secondary network can achieve near-constant (except for poly-logarithmic factor) capacity and delay scalings. This result indicates the mobility heterogeneity of SUs can be utilized to significantly improve the capacity and delay scalings of the cognitive radio network.

## REFERENCES

- [1] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transaction on Information Theory*, vol. 46, no. 2, pp. 388-404, 2000.
- [2] M. Franceschetti, O. Dousse, D. N. C. Tse, and P. Thiran, "Closing the Gap in the Capacity of Wireless Networks via Percolation Theory," *IEEE Transactions on Information Theory*, vol. 53, no. 3, pp. 1009-1018, 2007.
- [3] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Transactions on Networking*, vol. 10, pp. 477-486, 2002.
- [4] M. J. Neely and E. Modiano, "Capacity and Delay Tradeoffs for Ad-Hoc Mobile Networks," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 1917-1937, 2005.
- [5] X. Lin and N. B. Shroff, "The fundamental capacity-delay tradeoff in large mobile wireless networks," *Presented at the Third Annual Mediterranean Ad Hoc Networking Workshop Bodrum, Turkey, 2004*.
- [6] G. Sharma, R. Mazumdar, and N. B. Shroff, "Delay and Capacity Trade-offs in Mobile Ad Hoc Networks: A Global Perspective," *IEEE Transactions on Networking*, vol. 15, no. 5, pp. 981-992, October 2007.
- [7] L. Ying, S. Yang, R. Srikant, "Optimal Delay-Throughput Trade-Offs in Mobile Ad Hoc Networks," *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 4119-4143, 2008.
- [8] M. Garetto, and E. Leonardi, "Restricted Mobility Improves Delay-Throughput Trade-offs in Mobile Ad-Hoc Networks," *IEEE Transactions on Information Theory*, Vol.56, No.10, pp.5016-5029, October 2010
- [9] X. Wang, Y. Bei, Q. Peng, and L. Fu, "Speed improves Delay-Capacity Trade-Off in MotionCast," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 5, pp. 729 - 741, May 2011.
- [10] S. Jeon, N. Devroye, M. Vu, S. Chung, and V. Tarokh, "Cognitive Networks Achieve Throughput Scaling of A Homogeneous Network," *submitted to IEEE Transactions on Information Theory*, 2009. Available at <http://arxiv.org/abs/0801.0938>.
- [11] C. Yin, L. Gao, and S. Cui, "Scaling Laws for Overlaid Wireless Networks: A Cognitive Radio Network versus a Primary Network," *IEEE/ACM Transactions on Networking*, vol. 18, no. 4, pp. 1317-1329, August 2010.
- [12] W. Huang and X. Wang, "Throughput and Delay Scaling of General Cognitive Networks," in *Proceedings of IEEE INFOCOM, 2011*.
- [13] L. Gao, R. Zhang, C. Yin, and S. Cui, "Throughput and Delay Scaling in Supportive Two-tier Networks," *submitted to JSAC, 2011*. Available at <http://arxiv.org/abs/0905.3407>.
- [14] X. Wang, F. Luo, Y. Li, Z. Cao, and X. Gan, "Mobility Reduces the Number of Secondary Users in Cognitive Radio Network," in *Proceedings of IEEE GLOBECOM, 2011*.
- [15] M. Ji, Z. Wang, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves, "The Capacity of Ad Hoc Networks with Heterogeneous Traffic Using Cooperation," in *Proceedings of IEEE INFOCOM, 2010*.
- [16] Y. Li, X. Wang, and X. Tian, "Scaling laws for Cognitive Radio Network with Heterogeneous Mobile Secondary Users," Technical Report, 2011. Available at <http://iwct.sjtu.edu.cn/Personal/xwang8/paper/CRHetero.pdf>.