

Throughput and Delay in Heterogeneous Cognitive Radio Networks with Cooperative Secondary Users

Riheng Jia, Jinbei Zhang, Feng Yang, Xiaoying Gan, Xiaohua Tian, Pengyuan Du, Xinbing Wang
 Dept. of Electronic Engineering, Shanghai Jiao Tong University, China
 Email: {jriaheng, abelchina, yangfeng, ganxiaoying, xtian, dpu, xwang8}@sjtu.edu.cn

Abstract—In this paper¹, we investigate the throughput and delay in heterogeneous cognitive radio networks (HCRN), where the data source and the destination (S-D) is heterogeneously distributed following a rank based model and secondary users (SUs) provide relay service for primary users (PUs). We consider two scenarios: 1) PUs and SUs are both static; and 2) PUs are static and SUs are mobile. For scenario 1, we show that the primary network throughput is the same for different heterogeneous extents of S-D distribution owing to the flexible assistance of SUs, while the throughput of secondary networks is proven to be changing with the S-D heterogeneity exponent α , which depicts the variation of different heterogeneous extents of S-D distribution. In addition, the delay of both primary and secondary networks are shown to be altering with α . Further, we reveal that the number of SUs required to assist PUs can be dramatically reduced when considering the S-D heterogeneity, while achieving the same primary network throughput. For scenario 2, we utilize a modified uniform mobility (MUM) model to depict the motion of SUs and mainly focus on the analysis of throughput and delay for primary networks. It shows that the primary network throughput is also free of the heterogeneous extent of S-D distribution, while the delay changes with α . Due to the mobility of SUs, a better delay-throughput tradeoff of primary networks is achieved compared with that in scenario 1.

Index Terms—Throughput, delay, cognitive radio networks (CRN), heterogeneous S-D distribution, cooperation, economical deployment, modified uniform mobility (MUM) model



1 INTRODUCTION

Recently, there is more and more stress over the already-crowded radio spectrum due to the explosive growth of large-scale wireless applications, which strongly provide motivation of exploring the cognitive radio network (CRN) for its efficient usage of spectrum. Vu *et al.* considered the throughput for single-hop cognitive radio networks and obtained a linear scaling of secondary networks in [2], [3]. Jeon *et al.* studied the throughput of a CRN under general environment and showed that both primary and secondary networks can achieve the same throughput as a stand-alone network in order sense while the secondary network may suffer from a finite outage probability in [4]. Yin *et al.* developed throughput under a similar assumption and adopted transmission protocols which could guarantee zero outage probability for secondary networks in [5].

A common assumption in all the afore-mentioned works of CRN is that the source and destination (S-D) is homogeneously distributed. However, we consider a more realistic situation where the distribution of S-D is no longer homogeneous, but following a rank based model in accordance with the behavior that source would prefer the closer rather than the further

destination to communicate in real life.

The distribution of S-D is developed from the original network architecture [6], [7] to the emerging useful models which capture the characteristic of S-D such as distance based model [8] and rank based model [9]. On foundation of these works, the heterogeneous feature of S-D distribution is brought into wireless networks and their impact on the performance of wireless networks began to draw much attention. Azimdoostv *et al.* considered wireless social networks with social groups consisting of local contacts as well as long-range contacts and the corresponding throughput had also been studied in [10]. Fu *et al.* incorporated the rank based model into wireless networks and explored the impact of traffic locality on capacity over the network under different traffic patterns in [11].

We notice that so far, there is still no research concerning on the impact of heterogeneous distribution of S-D on cognitive radio networks. It remains unknown that how the heterogeneity of S-D distribution will influence the performance of CRN. Gao *et al.* considered the situation where the secondary users (SUs) are assumed to be willing to share the primary users' (PUs) burden of routing packets for utilizing the idle spectrum of PUs in [12]. Thus, the throughput of primary networks could be improved by the aid of SUs. It is also significant to investigate whether

¹The earlier version of this paper is to appear in the proceedings of IEEE INFOCOM 2014 [13]

the SUs could still help improving the throughput of primary networks under the heterogeneous setting.

In this paper, we study the throughput and delay of heterogeneous cognitive radio networks (HCRN) and deeply explore the impact of heterogeneous S-D distribution on network performance. The heterogeneous distribution of S-D is generated by employing a rank based model upon which we propose cooperative routing schemes to fully utilize the S-D heterogeneity. Specifically, we consider the HCRN is operating with cooperative static and mobile SUs, respectively. Our main contributions are summarized as follows:

- **HCRN with cooperative static SUs:**

- 1) We show that the heterogeneous distribution of S-D will not influence the PUs throughput due to the flexible assistance of SUs. With the aid of a varying number of SUs, the primary network can always achieve the near optimal per-node throughput of $\Theta\left(\frac{1}{\log n}\right)$ in whatever ranges of α , where α is the S-D heterogeneity exponent. Moreover, we reveal that the number of SUs required to assist PUs can be dramatically reduced compared with that in [12], which is of great significance to the economical deployment of networks.
- 2) It is testified that both the throughput and delay of secondary networks differs in different ranges of α , which are proven to be affected by the heterogeneous distribution of S-D. And, we prove that providing PUs help will not degrade the performance of secondary networks. Moreover, a better delay-throughput tradeoff can be achieved for both primary and secondary networks as α tends to be larger.

- **HCRN with cooperative mobile SUs:**

- 1) The primary network throughput is proven to be free of the heterogeneous extent of S-D distribution where the almost constant throughput scaling can be achieved except for poly-logarithmic factor. Then, we prove that the delay of primary networks changes with α . And due to the mobility of SUs, the delay performance is largely improved compared with that with cooperative static SUs. Moreover, the delay scaling will be even lower than $\Theta(1)$ when considering the heterogeneity of S-D distribution.

The rest of this paper is organized as follows. In Section 2, the system model is defined. We study traffic locality under the rank based model in Section 3. Throughput and delay of HCRN with cooperative static secondary users or mobile secondary users are derived in Section 4 and 5, respectively. In Section 6, we discuss the results. Our paper is concluded in Section 7.

TABLE 1: Symbols and Definitions

P_p^i	Transmit power of the i -th primary pair
P_s^j	Transmit power of the j -th secondary pair
$X_{p,tx}^i$	Tx location of the i -th primary pair
$X_{p,rx}^i$	Rx location of the i -th primary pair
$X_{s,tx}^j$	Tx location of the j -th secondary pair
$X_{s,rx}^j$	Rx location of the j -th secondary pair

1.1 Related Works

Initiated by Gupta and Kumar's seminal work [1], the researchers started to explore the fundamental performance of large-scale wireless networks using the asymptotic analysis, which provides strong power of perception in network performance as the network size increases. In [1], the per-node capacity is upper bounded by $O\left(\frac{1}{\sqrt{n}}\right)$ and lower bounded by $\Omega\left(\frac{1}{\sqrt{n \log n}}\right)$ for a uniformly distribution of S-D. This gap was later closed by Franceschetti *et al.* [20] using the percolation theory.

By introducing the mobility into nodes, Grossglauser *et al.* [14] improved the per-node capacity to be $\Theta(1)$, regardless of the unscaled delay. Subsequently, Neely *et al.* [15] considered the throughput/delay tradeoffs in mobile ad hoc networks (MANETs), and enhanced the delay performance compared with [14] using the redundancy mechanism. Intrigued by these breakthrough works, MANETs has been the hotspot of research and many creative works [16], [17], [18] are emerging in multitude within the networking community. In particular, N. Bansal *et al.* [16] put forward the uniform mobility model and established a relatively balanced tradeoffs that $\Omega\left(\frac{o(n)}{n \log^3 n}\right)$ throughput is achieved with $\Theta(1)$ delay.

2 SYSTEM MODEL

2.1 Network Geometry

The network extension is considered to be an unit square where primary and secondary nodes coexist.

HCRN with cooperative static SUs: The static primary and secondary nodes are randomly distributed according to a Poisson Point Process (P.P.P.) of density n and m , respectively.

HCRN with cooperative mobile SUs: The static primary nodes are distributed according to a P.P.P. of density n , and the mobile secondary nodes are following a modified uniform mobility (MUM) model.

The rank based model is utilized to grouped the nodes into one to one S-D pairs. n and m are correlated with $m = n^\beta$, where $\beta > 1$.

2.2 Transmission Model

For the wireless channel in our work, the channel power gain is given as $g(d) = d^{-\gamma}$, where d is the

distance between a transmitter and its receiver, $\gamma > 2$ indicates the path-loss exponent.

We assume that each transmission can achieve the Gaussian channel capacity where the i -th primary Tx-Rx pair can communicate at a rate of

$$R_p^i = \log \left(1 + \frac{P_p^i g(\|X_{p,tx}^i - X_{p,rx}^i\|)}{N_0 + I_p^i + I_{sp}^i} \right), \quad (1)$$

where I_p^i and I_{sp}^i represent the sum interference from all the other concurrent primary transmitters and secondary transmitters to the receiver of the i -th primary pair, respectively. Suppose that N_p primary pairs and N_s secondary pairs communicate simultaneously, I_p^i and I_{sp}^i are given by $I_p^i = \sum_{k=1, k \neq i}^{N_p} P_p^k g(\|X_{p,tx}^k - X_{p,rx}^i\|)$ and $I_{sp}^i = \sum_{k=1}^{N_s} P_s^k g(\|X_{s,tx}^k - X_{p,rx}^i\|)$, respectively. We show the main symbols and their definitions in Table I.

Similarly, the j -th secondary pair can communicate at a rate of

$$R_s^j = \log \left(1 + \frac{P_s^j g(\|X_{s,tx}^j - X_{s,rx}^j\|)}{N_0 + I_s^j + I_{ps}^j} \right). \quad (2)$$

2.3 Mobility Model

In this paper, we define the *Modified Uniform Mobility (MUM) Model*. This model well captures the characteristic of human mobility in real world, and is a modified version of "uniform mobility model".

MUM: In this model, each of the m mobile SUs are moving inside the unit square at speed v , which is a constant and is identical for each SU.

- *Initial State:* At time T_0 , the SUs are distributed uniformly at random within the unit square. In addition, the SUs choose the direction of motion at time T_0 identically and independently in $[0, 2\pi)$.

- *In-Flight State:* After choosing a direction, the SU moves in that direction for a distance s , which is an exponentially distributed random variable with expectation ρ . After each flight of distance s , the SU randomly choose a new direction where it will move for a distance s . In the uniform mobility model [16], ρ is assumed to be of the same order of the diameter of network, which is $\Theta(1)$. However, under our new network setting, there is no need to make the mobile moving such a long distance each time before changing direction as the S-D distance tends to be shorter. The detailed analysis of ρ will be shown later in Section 5.

3 TRAFFIC LOCALITY

It is notable that we adopt a rank based model [9] to determine the destination for each primary or secondary source node in HCRN. This makes the traffic pattern significantly different from the traditional

CRN. Some useful properties of rank based model had been established in [11]. We refer readers to [11] for details and just provide the result here.

Lemma 1. *Considering a single network with $n + 1$ nodes and denote Y_i as the destination selected by the i -th node X_i , Then the average S-D distance $\mathbb{E}(\|X_i - Y_i\|)$ in a unit square can be bounded as:*

$$\mathbb{E}(\|X_i - Y_i\|) = \begin{cases} 1/\sqrt{n} & \alpha > 3/2 \\ \log n/\sqrt{n} & \alpha = 3/2 \\ n^{1-\alpha} & 1 < \alpha < 3/2 \\ 1/\log n & \alpha = 1 \\ 1 & 0 \leq \alpha < 1, \end{cases} \quad (3)$$

where α denotes the S-D heterogeneity exponent. It can be seen that as α increases, each source node tends to choose the closer node as its destination and the traffic flow tends to be localized. The larger value of α results in the shorter distance between the source and destination. Thus, different value of α lead to varying degrees of heterogeneity of S-D distribution.

4 THROUGHPUT AND DELAY OF HCRN WITH COOPERATIVE STATIC SECONDARY USERS

4.1 Primary Protocol

- The unit square is divided into small-square primary cells and the area of each cell satisfy $a_p(n) \geq \frac{\sqrt{2\beta \log n}}{n}$. Every 64 primary cells are supposed to grouped into a primary cluster and we use a round-robin fashion to determine when the cells could be active in each primary cluster. The time is divided into TDMA frames where each frame has 64 primary time slots.

- Define the horizontal data path (HDP) and the vertical data path (VDP) of a S-D pair as the horizon line and the vertical line connecting a source to its destination, respectively. Each source transmits data to its destination by first hopping to the adjacent cells on its HDP and then the VDP.

- When a primary cell is active, each primary source node transmits its own packets in order. Afterwards, each of the S-D paths passing through the cell could be served by the designated relay node for one packet only. Note that when a primary source node finished sending out its own packets, it could be a designated relay node either. Each primary time slot is divided into packet slots according to the fluid model. For each packet, it will be handed over directly to the destination if the destination node happens to be in the adjacent cell. Otherwise, the packets are blindly broadcasted to the nodes in the adjacent cell. At each packet transmission, we denote the transmitting power by $P a_p^{\frac{\gamma}{2}}(n)$, where P is a constant.

- Note that, the above primary protocol is operating independently of whether the secondary network is present or not. We will show later that the primary

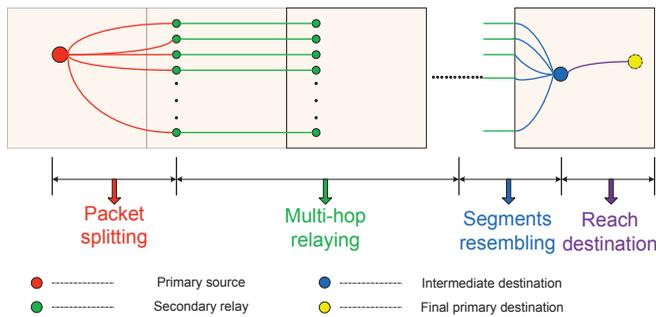


Fig. 1: The process of primary packet transmission with cooperative static SUs.

network can achieve better performance than being a stand-alone network in [1] with the aid of SUs.

4.2 Secondary Protocol

We assume that the secondary nodes are capable of pretending to be the primary nodes due to the so called “software-programmability”. So the “fake” primary nodes could be chosen to help primary users to relay packets. Once a secondary node is picked up for relaying packets, it will keep silent unless its related primary cell becomes active.

As we will discuss later that the time-sharing strategy will only incur a constant penalty on the throughput and delay of secondary networks. Thus we divide each secondary frame into three equal-length phases, thus each of them equals to one primary time slot.

Similarly as defined in the primary protocol, the unit area is divided into small-square secondary cells and the area of each cell satisfy $a_s(m) \geq \frac{2 \log m}{m}$. The secondary cells are grouped into secondary clusters with each secondary cluster of 64 secondary cells. Accordingly, each secondary cluster follows a 64-TDMA pattern to communicate. A preservation region is defined as a square containing 9 primary cells with the current active transmission in the center cell. Only the secondary transmitters in an active secondary cell outside all the preservation regions can transmit data, otherwise, the data has to be temporarily stored in buffer until the preservation region is clear.

Now we show the *three phases as follows*:

Phase 1 : When an active secondary cell is outside the preservation regions, each secondary source node is allowed to transmit one own packet and then serve each S-D path passing through the cell. The routing of secondary packets is similar to that defined in the primary protocol. At each packet transmission, the transmitting power is denoted by $P a_s^{\frac{\gamma}{2}}(m)$.

Phase 2 : Only the secondary nodes who buffered the primary packets are permitted to access the spectrum. We notice that the throughput per secondary S-D pair is less than that of primary S-D pair which is caused by the different node densities, so packet splitting is needed to prevent the bottleneck effect

from happening in relaying primary packets through the secondary network. As shown in Fig.1, suppose primary source node pours all the split primary segments to the neighboring cell, then in each primary cell, we assume that there are N secondary nodes including the designated relay node who are preparing to relay along the data paths passing through the cell. Each secondary node is assigned to relay $1/N$ portion of the primary packet to the intermediate destination by using a multi-hop scheme. Due to traffic locality, the value of N may vary in different ranges of α , which is much more complicated compared with that in [12]. Based on different traffic patterns, the value of N is set as follows:

$$N = \begin{cases} \Theta(\sqrt{\log n}) & \alpha = \frac{3}{2}, \beta > 1 \\ \Theta\left(n^{(\frac{3}{2}-\alpha)\beta} / \sqrt{\log n}\right) & 1 < \alpha < \frac{3}{2}, \beta \geq \frac{1}{\alpha-\frac{1}{2}} \\ \Theta\left(n^{\frac{\beta}{2}} / \log^{\frac{3}{2}} n\right) & \alpha = 1, \beta \geq 2 \\ \Theta\left(n^{\frac{\beta}{2}} / \sqrt{\log n}\right) & 0 \leq \alpha < 1, \beta \geq 2. \end{cases} \quad (4)$$

We can easily find that the heterogeneous distribution of S-D do make an impact on the number of SUs required to help relaying the primary packets. It should point out that the range of β here is to guarantee that there are more than N secondary users in each primary cell. When $0 \leq \alpha \leq 1$, we can calculate the range of $\beta \geq 2$, which is the same result as that in [12]. But when $1 < \alpha < \frac{3}{2}$ or $\alpha = \frac{3}{2}$, the results turn to be $\beta \geq \frac{1}{\alpha-\frac{1}{2}}$ or $\beta > 1$, respectively, which means the node density of secondary networks can be sharply reduced compared with that in [12]. The situation $\alpha > \frac{3}{2}$ is ignored in Equation (4) as the transmission tends to be a single-hop fashion which we will show the details in next subsection. The transmission scheme is similar as that in first phase.

Phase 3 : A recycling region consisted by nine primary cells and a layer of secondary cells around them is defined as to recycle the primary packet from the intermediate destination. Noting that each recycling region is located between two preservation regions and they are not overlapped with each other. The recycling process is assumed to happen in the center cell of the recycling region. At each packet transmission, the transmitting power of the intermediate destination node is the same as that of primary node.

4.3 Throughput Analysis

Before the throughput analysis, we first consider the transmission mode over the network in term of the average hops for each packet to go through. We denote \mathbb{E}_h as the average number of hops per packet for a single network, then $\mathbb{E}_h \sim \frac{\mathbb{E}(\|X_i - Y_i\|)}{\sqrt{a}}$, where a denotes the area of each cell in the network. By

definition, in primary networks we have

$$\mathbb{E}_h \sim \begin{cases} 1/\sqrt{\log n} & \alpha > 3/2 \\ \sqrt{\log n} & \alpha = 3/2 \\ n^{\frac{3}{2}-\alpha}/\sqrt{\log n} & 1 < \alpha < 3/2 \\ \sqrt{n/\log^3 n} & \alpha = 1 \\ \sqrt{n/\log n} & 0 \leq \alpha < 1. \end{cases} \quad (5)$$

It is easy to see that $\mathbb{E}_h \rightarrow \infty$ when $0 \leq \alpha \leq \frac{3}{2}$, and $\mathbb{E}_h \rightarrow 0$ when $\alpha > \frac{3}{2}$. It means that when $0 \leq \alpha \leq \frac{3}{2}$, we can still use our cellular TDMA multi-hop scheme to analyze per-node throughput of the primary and secondary network. However when the value of α exceeds $\frac{3}{2}$, a particular S-D pair is within the cell and it tends to be a single-hop transmission, which would be discussed in Section 4.3.3. In the rest of this paper, we denote $E(\|X_p^i - Y_p^i\|)$ as $E_{sd}(n)$ for short. Similarly, $E(\|X_s^i - Y_s^i\|)$ is denoted as $E_{sd}(m)$.

4.3.1 Primary Network

In order to obtain the throughput results, we first give the following lemmas.

Lemma 2. *The probability that a randomly selected designated relay node is a secondary node in a primary cell approaches one w.h.p..*

Proof: Denoting the probability by γ , from the proof of Theorem 1, we know that in a primary cell, the number of primary and secondary nodes are $\Theta(n a_p(n))$ and $\Theta(m a_p(n))$, respectively. Then the probability can be expressed as $\gamma = \frac{\Theta(m a_p(n))}{\Theta(m a_p(n)) + \Theta(n a_p(n))}$, which approaches one as $n \rightarrow \infty$, where $\beta > 1$. \square

Lemma 3. *With the primary protocol defined previously, an active primary cell can support a constant data rate of L_1 , the intermediate destination node can deliver the primary packets to the intended primary destination at a constant rate of L_2 , where $L_1, L_2 > 0$ and are independent of n and m .*

We refer readers to [12] for detailed proof of Lemma 3. Based on Lemma 2 and Lemma 3, we have

Theorem 1. *With the given protocols, the primary network can achieve the following per-node throughput w.h.p.:*

$$\lambda_p(n) = \begin{cases} \Theta\left(\frac{1}{n a_p(n)}\right) & 0 \leq \alpha \leq 1, \beta \geq 2 \\ \Theta\left(\frac{1}{n a_p(n)}\right) & 1 < \alpha < \frac{3}{2}, \beta \geq \frac{1}{\alpha - \frac{1}{2}} \\ \Theta\left(\frac{1}{n a_p(n)}\right) & \alpha = \frac{3}{2}, \beta > 1, \end{cases} \quad (6)$$

where $a_p(n) \geq \sqrt{2}\beta \log n/n$ and $a_p(n) = o(1)$.

Proof: In this proof, an upper-bound of the throughput per S-D pair is firstly derived and then we provide a lower-bound by using the secondary protocol in Section 4.2.

From Lemma 3, we know that the primary source node can transmit all its packets to the secondary network with a constant data rate of L_1 . Similarly, the primary destination node can receive all its packets with a constant data rate of L_2 . Since only primary source nodes transmit packets in each active primary cell while the secondary relay nodes keep silent. As the size of a primary cell equals to $a_p(n)$ and the node density of PUs is n , we can easily get the number of primary source nodes in each primary cell as $\Theta(n a_p(n))$ w.h.p., then the throughput is upper-bounded by $\Theta\left(\min\left(\frac{L_1}{n a_p(n)}, \frac{L_2}{n a_p(n)}\right)\right) = \Theta\left(\frac{1}{n a_p(n)}\right)$.

As to the lower-bound, we calculate the achievable throughput per S-D pair. In the analysis of secondary protocol, each primary source node transmits all its packets to the secondary network w.h.p. by dividing data into N secondary data paths, each of them at a rate of $\Theta\left(\frac{1}{m E_{sd}(m) \sqrt{a_s(m)}}\right)$ given in Theorem 2. By setting proper $\sqrt{a_s(m)}$ corresponding to different traffic patterns, which satisfy $a_s(m) \geq 2 \log m/m$. Thus, a throughput of $N \Theta\left(\frac{1}{m E_{sd}(m) \sqrt{a_s(m)}}\right) = \Theta\left(\frac{1}{n a_p(n)}\right)$ can be achieved for each primary source node, which is considered as the lower-bound.

By combining the two bounds, Theorem 1 is derived and thus completing the proof. \square

Noting that the condition of $\beta \geq 2$, $\beta \geq \frac{1}{\alpha - \frac{1}{2}}$ and $\beta > 1$ corresponding to different ranges of α are needed in the proof to guarantee that there are more than N secondary nodes in each primary cell w.h.p.. By setting $a_p(n) = \sqrt{2}\beta \log n/n$, the primary network can achieve the following per-node throughput w.h.p.:

$$\lambda_p(n) = \begin{cases} \Theta\left(\frac{1}{\log n}\right) & 0 \leq \alpha \leq 1, \beta \geq 2 \\ \Theta\left(\frac{1}{\log n}\right) & 1 < \alpha < \frac{3}{2}, \beta \geq \frac{1}{\alpha - \frac{1}{2}} \\ \Theta\left(\frac{1}{\log n}\right) & \alpha = \frac{3}{2}, \beta > 1. \end{cases} \quad (7)$$

From Equation (7) we can see that the per-node throughput of primary networks can be improved from $\Theta(1/\sqrt{n \log n})$ as in the stand-alone network to $\Theta(1/\log n)$ with the help of SUs. We can also find that the density of SUs can be sharply reduced to $\beta \geq \frac{1}{\alpha - \frac{1}{2}}$ ($1 < \alpha < \frac{3}{2}$) or $\beta > 1$ ($\alpha = \frac{3}{2}$) compared with the results $\beta \geq 2$ in [12], while achieving the same primary network throughput in order sense.

4.3.2 Secondary Network

Using the similar way as the analysis of Equation (5), we can easily get that the transmission also tends to be a single-hop fashion in the secondary network when $\alpha > \frac{3}{2}$. In the following, we mainly focus on the throughput for the secondary network operating in a multi-hop fashion.

In order to obtain the throughput results, we first give the following lemmas.

Lemma 4. Let n_{st} denote the number of total secondary nodes in the unit square, then we have $\frac{m}{2} < n_{st} < em$ w.h.p..

The proof can be found in [5].

Lemma 5. With the secondary protocol defined previously, each TX in a secondary cell can support a constant data rate of K which is independent of m .

Proof: Since the 64-TDMA scheme and the preservation region are adopted, the interference caused by the other concurrent primary or secondary transmitters could be bounded by a constant. \square

We note that the interference at a particular receiver only depends on the scheduling scheme in this situation since all of the interference comes from the concurrent transmitters. Therefore we can use the same method to bound the interference although the S-D distance in our model is different from that in [5].

The next lemma shows the bound of the average number of S-D paths passing through or originating from each secondary cell.

Lemma 6. The number of secondary S-D paths (including both HDPs and VDPs) passing through or originating from each secondary cell is bounded as $\Theta\left(m E_{sd}(m) \sqrt{a_s(m)}\right)$ w.h.p., for $0 \leq \alpha \leq \frac{3}{2}$.

Proof: According to Lemma 1, the distance of secondary S-D pair is bounded by $\Theta(E_{sd}(m))$ w.h.p., so it is intuitive that the corresponding HDP (or VDP) has a distance of the same order.

Based on this corollary, we affirm that all the secondary sources of HDPs that pass through or originate from a particular cell should be located in an area of size $\Theta\left(E_{sd}(m) \sqrt{a_s(m)}\right)$ w.h.p.. We assume $\Theta\left(E_{sd}(m) \sqrt{a_s(m)}\right)$ could be bounded by $C_1 E_{sd}(m) \sqrt{a_s(m)}$ where C_1 is a constant independent of m . Now we proceed the proof as follow.

Let m_h denote the number of HDPs passing through or originating from each secondary cell, by assuming all secondary nodes are sources, the resulting upper bound on m_h follows Poisson distribution $\left(\lambda = m C_1 E_{sd}(m) \sqrt{a_s(m)}\right)$. Using the well known inequality $\mathbb{P}(X \geq x) \leq \frac{e^{-\lambda} (\lambda)^x}{x^x}$, for $x > \lambda$ [20], where X is a Poisson random variable with parameter λ , we obtain $P(m_h \geq 2\lambda) \leq \left(\frac{e}{4}\right)^{m C_1 E_{sd}(m) \sqrt{a_s(m)}}$. Similarly, we denote m_v as the number of corresponding VDPs, and we also have $P(m_v \geq 2\lambda) \leq \left(\frac{e}{4}\right)^{m C_1 E_{sd}(m) \sqrt{a_s(m)}}$. According to the union bound, we have

$$\mathbb{P}(m_h + m_v \geq 4\lambda) \leq 2 \left(\frac{e}{4}\right)^{m C_1 E_{sd}(m) \sqrt{a_s(m)}}$$

Therefore the probability that there is at least one cell supporting more than 4λ data paths is upper bounded by $2m \left(\frac{e}{4}\right)^{m C_1 E_{sd}(m) \sqrt{a_s(m)}}$ since there are at

most m cells in the network. It is not difficult to prove that $2m \left(\frac{e}{4}\right)^{m C_1 E_{sd}(m) \sqrt{a_s(m)}} \rightarrow 0$ w.h.p., when $E_{sd}(m)$ varies according to Equation (3) for $0 \leq \alpha \leq \frac{3}{2}$. which completes the proof. \square

Note that Lemma 6 is new result for the HCRN setup. Based on Lemma 4-6, we have the following theorem.

Theorem 2. With the predefined model and secondary protocol, the secondary network can achieve the following per-node throughput w.h.p.:

$$\lambda_s(m) = \begin{cases} \Theta\left(1/\log^{\frac{3}{2}} m\right) & \alpha = \frac{3}{2} \\ \Theta\left(m^{\alpha-\frac{3}{2}}/\sqrt{\log m}\right) & 1 < \alpha < \frac{3}{2} \\ \Theta\left(\sqrt{\log m}/\sqrt{m}\right) & \alpha = 1 \\ \Theta\left(1/\sqrt{m \log m}\right) & 0 \leq \alpha < 1. \end{cases} \quad (8)$$

Proof: From Lemma 5 we know that a given TX node in a secondary cell can support a constant data rate of at least K , so each secondary S-D pair can achieve a data rate of at least K divided by the maximum number of data paths that pass through or originate from the secondary cell. We already proved that the number of data paths handled by a particular secondary cell is $\Theta\left(m E_{sd}(m) \sqrt{a_s(m)}\right)$ w.h.p..

Thus the per-node throughput $\Theta\left(\frac{1}{m E_{sd}(m) \sqrt{a_s(m)}}\right)$ is achieved w.h.p. \square

It should point out that since the given time sharing strategy only incurs a constant penalty (i.e., $1/3$) on the achievable throughput and delay within the secondary network and thus will not hurt the scaling law performance of secondary networks. Therefore, Theorem 2 still holds when the secondary network operates as a stand-alone network.

4.3.3 Throughput Analysis for Single-hop Fashion

When $\alpha > \frac{3}{2}$, each S-D pair locates within a particular cell w.h.p. for both primary and secondary networks. In this case, we could easily bound the per-node throughput of primary and secondary networks by adapting the routing scheme to a single-hop fashion.

We have studied the interference for primary and secondary networks in a cellular TDMA model. Since we remain the scheduling scheme unchanged, Lemma 3 and Lemma 5 still hold in this case. Therefore the S-D pair in both networks can transmit data at a constant rate. We give the achievable per-node throughput of the single-hop fashion through the following theorem.

Theorem 3. In our proposed single-hop scheme, the primary and secondary network can achieve the following per-node throughput w.h.p.:

$$\lambda_p(n) = \Theta\left(\frac{1}{\log n}\right), \lambda_s(m) = \Theta\left(\frac{1}{\log m}\right). \quad (9)$$

Proof: Since we adopt a single-hop scheme, each S-D pair can transmit data within a particular cell at

a constant rate according to Lemma 3 and Lemma 5. Moreover, as stated in the proof of Theorem 1, there are at most $\Theta(\log n)$ nodes in a primary cell and $\Theta(\log m)$ nodes in a secondary cell. Therefore, the achievable throughput is at least the transmission rate divided by the number of nodes in a particular cell for both primary and secondary networks. \square

4.4 Delay Analysis

Similar to the throughput analysis, first, we consider the situation when $0 \leq \alpha \leq \frac{3}{2}$. Then we focus on the single-hop fashion, where $\alpha > \frac{3}{2}$.

4.4.1 Primary Network

First of all, we calculate the traveling time for the N segments of a primary packet to reach the corresponding intermediate destination node within the secondary network. Since the data paths for the N segments are along the route and an active secondary cell (outside all the preservation regions) transmits one packet for each data path passing through it within a secondary time slot. Assuming that there exists a central entity to coordinate the transmission of N packet segments, thus we can guarantee that the N segments depart from the N nodes, move hop by hop along the data paths, and finally reach the corresponding intermediate destination in a synchronized fashion. According to the fluid model, the N segments experience the same delay in Theorem 5.

Let T_p and T_s denote the durations of the primary and secondary time slot, respectively. According to the proposed protocols, we have $T_p = 64 T_s$. Since each secondary time frame is split into three phases and one of them is used for relaying the primary packet, each primary packet suffers from the following delay

$$D_p(n) = \frac{3}{64} D_s(m) + C_1 + C_2 = \Theta\left(\frac{E_{sd}(m)}{\sqrt{a_s(m)}}\right), \quad (10)$$

where the secondary network delay $D_s(m)$ is defined in Theorem 5. C_1 denotes the average time for a primary packet traveling from the primary source node to the N secondary relay nodes, C_2 denotes the average time for a primary packet traveling from the intermediate destination node to the final destination node, which C_1 and C_2 are both constants. In order to obtain a better delay performance, we should make $a_s(m)$ as large as we could. However, a larger $a_s(m)$ results in a decreased throughput per S-D pair in the secondary network and consequently a decreased throughput for the primary network, for the primary traffic traverse over the secondary network w.h.p..

Lemma 7. *The relationship between $a_p(n)$ and $a_s(m)$ based on the predefined protocols is as follows:*

$$a_s(m) = \left(\frac{N n a_p(n)}{m E_{sd}(m)}\right)^2. \quad (11)$$

Proof: According to Theorem 1, the maximum throughput per S-D pair for the primary network is $\Theta\left(\frac{1}{n a_p(n)}\right)$. Since a primary packet is divided into N segments and then routed by N parallel S-D paths within the secondary network, the supported rate for each secondary S-D pair is required to be $\Theta\left(\frac{1}{N} \times \frac{1}{n a_p(n)}\right)$. And based on Theorem 2, the secondary cell size needs to be set as $\left(\frac{N n a_p(n)}{m E_{sd}(m)}\right)^2$. \square

Substituting Equation (4) into Equation (10), we have the following theorem.

Theorem 4. *Based on the proposed protocols in Subsection A and B, the primary network can achieve the following delay w.h.p.*

$$D_p(n) = \Theta\left(\frac{n^{\beta-1} E_{sd}(m)}{N a_p(n)}\right). \quad (12)$$

4.4.2 Secondary Network

The delay performance of the secondary network is given by the following theorem.

Theorem 5. *Based on the predefined secondary protocol, the delay is given by*

$$D_s(m) = \begin{cases} \Theta(\sqrt{\log m}) & \alpha = \frac{3}{2} \\ \Theta\left(m^{\frac{3}{2}-\alpha} / \sqrt{\log m}\right) & 1 < \alpha < \frac{3}{2} \\ \Theta\left(\sqrt{m} / \log^{\frac{3}{2}} m\right) & \alpha = 1 \\ \Theta\left(\sqrt{m} / \sqrt{\log m}\right) & 0 \leq \alpha < 1. \end{cases} \quad (13)$$

Proof: We first calculate the average number of hops for each packet to reach the destination along the secondary S-D path. Then we use the fact that each hop costs one secondary packet a constant time to derive the average delay for each secondary S-D pair. We denote E_h^s as the average number of hops for each secondary S-D pair, then we have $E_h^s \sim \frac{E_{sd}(m)}{\sqrt{a_s(m)}}$.

Since we use the fluid model, the packet size of each secondary transmission scales proportionally to the throughput $\lambda_s(m)$, each packet arriving at a secondary cell will be sent to the next active time slot of the cell.

So, w.h.p. $D_s(m) = \frac{2}{n_{st}} \sum_{j=1}^{\frac{n_{st}}{2}} \frac{\|X_s^j - Y_s^j\|}{\sqrt{a_s(m)}} = \Theta\left(\frac{E_{sd}(m)}{\sqrt{a_s(m)}}\right)$. \square

4.4.3 Delay Analysis for Single-hop Fashion

We provide the delay performance of primary and secondary networks in a single-hop fashion (when $\alpha > \frac{3}{2}$) through the following theorem.

Theorem 6. *Based on the predefined protocols, the primary and secondary network can achieve the following delays,*

$$D_p(n) = \Theta(\log n), \quad D_s(m) = \Theta(\log m). \quad (14)$$

Proof: According to the analysis of Equation (5), each S-D pair is located in a particular cell w.h.p. in this case. Therefore the packet delay is the time for each packet to reach the destination through

single-hop transmission. Recall that there are at most $\Theta(\log n)$ nodes in a primary cell, the probability that a particular S-D pair is chosen to transmit data during its active time slot is $P_{Delay,P} = \Theta\left(\frac{1}{\log n}\right)$. Thus we have $D_p(n) = \frac{1}{P_{Delay,P}} = \Theta(\log n)$.

For the secondary network, as the preservation regions will not influence the delay in order sense. Hence, we have $D_s(m) = \Theta(\log m)$. \square

5 THROUGHPUT AND DELAY OF HCRN WITH COOPERATIVE MOBILE SECONDARY USERS

In this section, we mainly focus on the impact of heterogeneous S-D distribution on the throughput and delay of primary networks with cooperative mobile SUs when $0 \leq \alpha \leq \frac{3}{2}$. For primary networks, the S-D is still heterogeneously distributed following the rank based model. For secondary networks, since the SUs are mobile and the distance of secondary S-D is changing all the time, the investigation of the impact of heterogeneous S-D distribution on mobile secondary networks is fruitless. Moreover, from the results obtained in Section 4, we can easily get that SUs can achieve the same scaling performance as a stand-alone network when transmitting secondary packets and relaying the primary packets simultaneously. And, this problem was extensively studied in [14]-[16]. Therefore, in the following we assume that all the data are produced by primary networks, and secondary networks only provide relay service for simplicity.

5.1 Primary Protocol

In our protocol, all the primary packets are assumed to be relayed by SUs in a multi-hop fashion. The unit square is partitioned the same way as that in Section 4.1, and a 25-TDMA scheduling scheme is applied here. The detailed routing process of primary packets is divided into three phases as follow:

• **Primary Source to SU:** As showed in Fig. 2, once a particular primary source node S meets some SUs M_0 moving in the direction of δ such that $\delta \in [\theta - \frac{\pi}{6}, \theta + \frac{\pi}{6}]$ within the same cell (θ denotes the direction of the destination R relative to S), S will deliver the packet to M_0 if it has packets to transmit. In fact, the choice of angle $\frac{\pi}{6}$ is arbitrary and any SU moving within an angle of $(0, \frac{\pi}{3}]$ would satisfy.

Note that M_0 should also satisfy the property that it moves in the direction of δ for a distance of at least $\frac{2}{\sqrt{3}}|SR|$ ($|SR|$ denotes the S-D distance), which guarantees that M_0 can successfully reaches the handoff region. Since the number of SUs are larger than that of PUs in order sense, a constant fraction of SUs satisfy this property. In [16], the average S-D distance is $\Theta(1)$ for the uniform distribution of S-D pairs, and ρ (see Section 2.3) is set to be $\Theta(1)$. However, in HCRN

the average S-D distance tends to be smaller with the variation of α . Therefore, there is no need for the SU to move such a long distance in each flight.

Proposition 1. *Under our new network setting, the expectation of s (distance of each flight before changing direction) is set as follow:*

$$\rho = \begin{cases} \frac{2 \log n}{\sqrt{3n}} & \alpha = 3/2 \\ \frac{2 n^{1-\alpha}}{\sqrt{3}} & 1 < \alpha < 3/2 \\ \frac{\sqrt{3}}{2} & \alpha = 1 \\ \frac{\sqrt{3} \log n}{\sqrt{3\pi}} & 0 \leq \alpha < 1 \end{cases} \quad (15)$$

• **SU to SU:** As this phase is completely done by the SUs, which we will show the details later.

• **SU to Primary Destination:** As the relay process goes on, the primary packet is getting closer and closer to the corresponding destination R . When the SU falls into the primary cell containing R , the primary packet is finally handed over to R .

5.2 Secondary Protocol

In this part, we mainly show the details of secondary relaying process, which corresponds to the phase of SU to SU defined above.

The unite square is divided into small-square secondary cells the same way as that in Section 4.2, and a 25-TDMA scheduling scheme is applied here. In addition, the secondary network share the same time frame structure with primary networks.

• **SU to SU:** After the primary packet being delivered to M_0 successfully, M_0 keeps flying in the direction of δ . As M_0 will receive multiple primary packets which are designated to different destinations. At any time when coming across another proper SU, M_0 should make a decision which packet should be delivered first. If this issue is not handled well, a large delay may occur due to the limited contact time of two mobile SUs. In the following, we first establish the *Handoff Region* denoted as quadrilateral G_{AERC} in Fig. 2. Then the *Earliest Deadline First* (EDF) scenario in [16] is utilized to determine which packet should be delivered first.

Handoff Algorithm: Let's focus on the handoff region G_{AERC} . Line AE , line FG and line CR are all perpendicular to line SR . Line FG is the central line of G_{AERC} . When M_0 enters the handoff region (flying across line AE), each of the buffered packets randomly selects a static node as its deadline within the area of G_{FGRC} . For instance, we assume that packet P_1 has a deadline D_1 at a distance of d away. Then, P_1 would be discarded if not being successfully handed off to another proper SU until M_0 reaches D_1 . Suppose M_0 wants to transmit the primary packet to R which is located in the direction of θ^* . Thus M_0 will deliver the packet to some SUs M_1 which moves in the direction of δ^* such that $\delta^* \in [\theta^* - \frac{\pi}{6}, \theta^* + \frac{\pi}{6}]$. The whole process repeats thus the primary packet

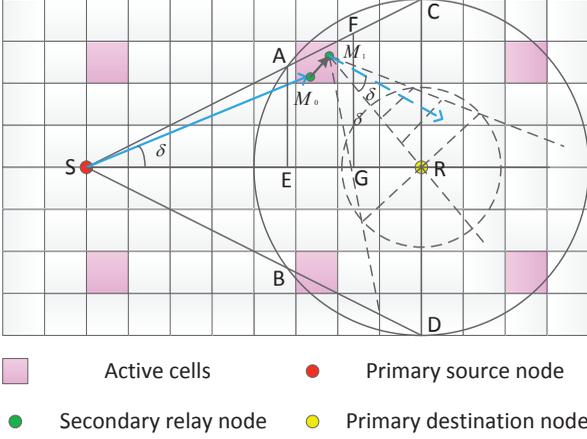


Fig. 2: The relay process of primary packet from step k to $k + 1$ with cooperative mobile SUs.

is transferred among the SUs from source to destination. In the following lemma, we give the maximum number of handoff for any primary packet, which is of importance to the delay analysis later.

Lemma 8. *Based on Proposition 1 and the analysis of SU to SU, for any primary packet, the maximum number of handoff N^* is:*

$$N^* = \begin{cases} \log_{\sqrt{3}} \sqrt{n} & \alpha = 3/2 \\ 2(1 - \alpha) \log_{\sqrt{3}} (n/\log n) & 1 < \alpha < 3/2 \\ \log_{\sqrt{3}} (n/\log^2 n) & \alpha = 1 \\ \log_{\sqrt{3}} (n/\log n) & 0 \leq \alpha < 1 \end{cases} \quad (16)$$

which can be bounded as $\log n$ in any case of α .

Proof: Let D denote the maximum primary S-D distance, then according to Lemma 1 D can be bounded as $E(\|X_i - Y_i\|)$. Considering the extreme situation where M_0 delivers the primary packet to M_1 successfully at point A in Fig. 2, then the distance between the packet and its destination is shortened to be $D/\sqrt{3}$. Therefore, the distance between a primary packet and its destination decreases by at least a factor of $\sqrt{3}$ after each handoff. We define a packet as stage k packet if the temporary position of the packet is at a distance between $D3^{-\frac{k}{2}}$ and $D3^{-\frac{k-1}{2}}$ from the destination. Since the distance between the primary packet and its destination ranges from at most $E(\|X_i - Y_i\|)$ and at least $\log n/n$, N^* can be obtained from the equation $\frac{E(\|X_i - Y_i\|)}{(\sqrt{3})^k} = \frac{\log n}{n}$. \square

5.3 Throughput Analysis

As mentioned in the secondary protocol, when the SU enters the handoff region, numerous primary packets buffered in the SU are competing for being handed off. If the primary source node produces data too fast, then the number of packets which needs to be handed off within the handoff region would be larger. Thus, the severe packet loss phenomenon would happen due to the finite transmission ability of SUs. In the

following, we will consider these issues and obtain the throughput results.

We assume that each primary source node produces data at a rate at most $\lambda_p(n)$, then the whole network can produce data at a total rate of $n\lambda_p(n)$. We say that there is an input into stage k if a packet becomes a stage k packet. Since there are totally m SUs, we expect that each SU will receive stage k packets at a rate of at most $\frac{n\lambda_p(n)}{m}$. The receiving rate of SU can be bounded as $\frac{6n\lambda_p(n)}{m}$ even if all the packets need to be sent in a particular direction. The reason is that they will be spread out to SUs moving in a direction of $[\theta - \frac{\pi}{6}, \theta + \frac{\pi}{6}]$.

Lemma 9. *Consider a particular primary packet Z_p in its handoff region with a deadline at a distance of d away. Let $A(d)$ denotes the amount of data whose deadlines are earlier than that of Z_p , then we have:*

$$E[A(d)] < [108\beta + 24] \frac{n\lambda_p(n)d}{m\|v\|} \log n. \quad (17)$$

Proof: Consider a particular primary packet Z_p and let M denote the SU who carries Z_p . As illustrated in Fig. 3, the handoff region of Z_p begins at point B and its deadline point C is located at a distance of d away. For simplicity, we denote T_0 as the time when Z_p enters its handoff region.

According to the handoff algorithm, the handoff takes place according to the EDF. Therefore, the handoff of Z_p would be delayed by the other packets which have a earlier deadline. In the following, we target on bounding the amount of data which have a earlier deadline than that of Z_p . We divided these packets into two types: 1) Packets which arrive at M after T_0 ; 2) Packets which arrive at M before T_0 .

For the packets of the first type, until reaching the deadline C , M has moved a distance d . Thus the expected amount of data which M receives can be bounded as $\frac{6n\lambda_p(n)}{m} \frac{d}{\|v\|} \log m$. In the worst case we assume that all these packets have a deadline before that of Z_p , then the expected number of packets which need to be handed off before Z_p can be bounded as $6n \log m \frac{\lambda_p(n)d}{m\|v\|}$.

For the packets of the second type, the analysis will be much more complicated. We define the packets of the second type as Z_{pp} . Back to Fig. 3, we assume that M receives Z_{pp} at point A , which located at a distance X from point B . we further consider Z_{pp} in two circumstances.

- If the distance between the destination of Z_{pp} and A is less than $\frac{\sqrt{3}}{2}X$, then the deadline of Z_{pp} is located at a distance of at most X from point A , which means the deadline of Z_{pp} will not fall into segment BC .
- If the destination of Z_{pp} is located at point E , thus line AE equals to $\frac{2\sqrt{3}}{3}(X + d)$ (can be calculated using some geometry knowledge), which is ap-

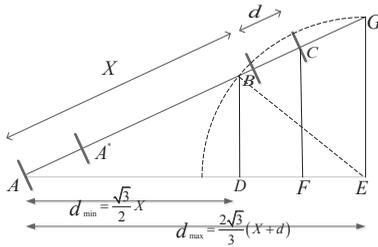


Fig. 3: Figure explaining the proof of Lemma 9.

proximately $\frac{4}{3}(X+d)$. In this setting, the handoff region of Z_{pp} is shown in Fig. 3 as quadrilateral $GBDEG$, then the deadline of Z_{pp} is located at a distance at least $X+d$ from point A. In other words, the deadline of Z_{pp} will not fall into segment BC .

Packet Z_p would not be delayed by Z_{pp} in either case above. Thus, if the destination of Z_{pp} lies between the distances $d_{\min} = \frac{\sqrt{3}}{2}X$ and $d_{\max} = \frac{4}{3}(X+d)$ from point A.

Calculate the ratio of d_{\max} to d_{\min} , which is $\frac{8(X+d)}{3\sqrt{3}x}$. When $X < 8d$, the ratio is larger than $\sqrt{3}$, then the expected number of packets which need to be handed off before Z_p can be bounded as $48d \log m \frac{n \lambda_p(n)}{m \|v\|}$. When $X > 8d$, the ratio is smaller than $\sqrt{3}$, which implies that M may receive at most two kinds of stages packet. Thus, the expected arrival rate of such packets is at most $\frac{12n \lambda_p(n)}{m}$. For any packet ($X > 8d$), the size of its deadline interval is at least $\frac{X}{4}$, then the probability that Z_{pp} has a deadline in the segment BC equals to $d/\frac{X}{4} = \frac{4d}{X}$. Due to the heterogeneous distribution of primary S-D, the expected number of conflicting packets relative to Z_p will be bounded as:

$$\begin{aligned} & \int_{X=8d}^{E\|X_i - Y_i\|} 12n \lambda_p(n) \frac{4d}{Xm \|v\|} dX \\ &= \frac{48n \lambda_p(n) d}{m \|v\|} \log \left(\frac{nE \|X_i - Y_i\|}{8d} \right) \\ &< \frac{24n \lambda_p(n) d}{m \|v\|} \log (nE \|X_i - Y_i\|), \end{aligned} \quad (18)$$

Substituting the value of $E\|X_i - Y_i\|$ into Equation (18), we get the bound as:

$$\begin{cases} \frac{24n d \lambda_p(n)}{m \|v\|} \log(\sqrt{n} \log n) & \alpha = 3/2 \\ \frac{(2-\alpha)24n d \lambda_p(n)}{m \|v\|} \log n & 1 < \alpha < 3/2 \\ \frac{24n d \lambda_p(n)}{m \|v\|} \log \left(\frac{n}{\log n} \right) & \alpha = 1 \\ \frac{24n d \lambda_p(n)}{m \|v\|} \log n & 0 \leq \alpha < 1, \end{cases} \quad (19)$$

which is at most $\frac{24n d \lambda_p(n)}{m \|v\|} \log n$ at any case of α .

To sum up, the expected amount of data that needs to be handed off before a packet with deadline at a distance of d is delivered can be bounded as $[108\beta + 24] \frac{n \lambda_p(n) d}{m \|v\|} \log n$ \square

Now that we have obtained the bound of the

amount of data which a SU needs to handoff during the time $\frac{d}{\|v\|}$. In the following, we investigate how much data a SU can transmit during the time $\frac{d}{\|v\|}$.

Assuming that two SUs can communicate only when they are within a distance of at most $r_c = \sqrt{2} \sqrt{\frac{2 \log m}{m}}$ ($\sqrt{2}$ times the side length of a primary cell). We say that two SUs contact each other if the distance between them is less than $\frac{r_c}{2}$. First, we show the amount of data a SU can transmit during a contact.

Lemma 10. *Conditioned on the fact that the bandwidth allocated for mobile-to-mobile data transmission is $\frac{1}{25}W$, then the expected amount of data a SU can transmit successfully during a contact w.h.p. is $\frac{W r_c}{50 \|v\|}$.*

Proof: Consider two SUs M_1 , M_2 , and M_2 is moving toward M_1 with a relative velocity at most $2\|v\|$. Thus, during a contact M_1 and M_2 are within a distance of r_c for a duration of at least $\frac{r_c}{2\|v\|}$. \square

Next, we calculate the number of contacts which a SU has in time t .

Suppose a SU M prepares to transmit data to a destination in the direction ω . According to the routing protocol, M will deliver the packet to some other SU M_i moving in the direction ζ , such that $\zeta \in [\omega - \frac{\pi}{6}, \omega + \frac{\pi}{6}]$.

Given M , we define the SU M_i as contactable if M_i satisfy: 1) M_i move in the direction ζ ; 2) M_i contacts M during the time t . Let $N(\omega, t) = |\{i | M_i = \text{contactable}\}|$, which denotes the number of contacts which a SU has in time t . We use the following lemma to show the expectation of $N(\omega, t)$, of which the proof is similar to that in [16]. However, we include it for completeness.

Lemma 11. *Consider a particular SU M , for $N(\omega, t)$ defined above. Then we have:*

$$E[N(\omega, t)] > \frac{1}{24} m \|v\| t r_c, \quad (20)$$

for any $\omega \in [0, 2\pi)$.

Proof: It is obvious that $N(\omega, t)$ is a sum of $m-1$ i.i.d. Bernoulli random variables N_i , where $N_i = 1$ if M_i is contactable. Then, to obtain $E[N(\omega, t)]$, we just need to calculate $E[N_i]$ for $E[N(\omega, t)] = (m-1)E[N_i]$.

Firstly, we investigate the condition of $N_i = 1$. Consider the SU moving in the direction ζ , thus we get the velocity of M_i as $|2\|v\| \sin(\frac{\zeta}{2} + \frac{\pi}{2})|$. And $N_i = 1$ is true only if M_i lies in a region of area $t \times r_c \times |2\|v\| \sin(\frac{\zeta}{2})|$. Conditioned on the moving direction of M_i , we get the expectation of N_i as:

$$E[N_i] = \frac{1}{2\pi} \int_{\zeta=\omega-\frac{\pi}{6}}^{\zeta=\omega+\frac{\pi}{6}} t \times r_c \times |2\|v\| \sin(\frac{\zeta}{2})| d\zeta, \quad (21)$$

$E[N_i]$ is minimized when $\omega = 0$. In this case, we get

the lower bound of $E[N_i]$ as:

$$E[N_i] = \frac{2\|v\|tr_c}{2\pi} \times 2 \int_0^{\frac{\pi}{6}} \sin\left(\frac{\zeta}{2}\right) d\zeta > \frac{1}{24} \|v\|tr_c. \quad (22)$$

Thus, $E[N(\omega, t)] > \frac{1}{24} (m-1) \|v\|tr_c \approx \frac{1}{24} m \|v\|tr_c$.

Substituting $r_c = \sqrt{2} \sqrt{\frac{2 \log m}{m}}$ and using the secondary Chernoff bound [21] with $\lambda = \frac{1}{2}$, we get that:

$$\Pr \left[N(\omega, t) < \frac{1}{2} \cdot \frac{1}{24} \cdot 2 \sqrt{\frac{\log m}{m}} \cdot t \|v\| m \right] \leq e^{-\frac{1}{48} (\log m)^{\frac{3}{2}}} \quad (23)$$

where $t > \frac{2 \log m}{\|v\| \sqrt{m}}$. Equation (23) implies that a SU can contact at least $\Theta(\sqrt{\log m})$ other SUs if it travels a distance $\frac{2 \log m}{\sqrt{m}}$ with probability at least $1 - \frac{1}{(\log m)^2}$. \square

Based on Lemma 10 and Lemma 11, we directly give the following results:

Lemma 12. Denote $N(t)$ as the amount of data a SU can transmit in time t , then for $t > \frac{2 \log m}{\|v\| m}$ and m is sufficiently large, we have:

$$\Pr \left[N(t) > \frac{1}{300} W t \log m \right] > 1 - \frac{1}{(\log m)^2}. \quad (24)$$

Theorem 7. According to the network protocols and combine Lemma 9-11, then we get the achievable per-node throughput of primary networks as:

$$\lambda_p(n) = \begin{cases} \Theta\left(\frac{1}{\log n}\right) & \alpha = 3/2 \\ \Theta\left(\frac{1}{\log n}\right) & 1 < \alpha < 3/2 \\ \Theta\left(\frac{1}{\log n}\right) & \alpha = 1 \\ \Theta\left(\frac{1}{\log n}\right) & 0 \leq \alpha < 1 \end{cases} \quad (25)$$

Proof: Consider a particular primary packet Z_p in its handoff region, which has a deadline at a distance of d away. To guarantee that the primary packets can be successfully handed off *w.h.p.*, the amount of data which have a deadline earlier than that of Z_p should be no larger than the amount of data a SU can transmit in time t , which is at least $\Theta(Wt \log m)$ with probability of no less than $1 - \frac{1}{(\log m)^2}$.

On the other hand, using the Markov's Inequality, we get that the probability that more than $\Theta\left(\frac{nt \lambda_p(n) (\log m) (\log n)}{m}\right)$ amount of data have a deadline earlier than that of Z_p is at most $\frac{1}{\log m}$. Finally, if the inequality $\frac{nt \lambda_p(n) (\log m) (\log n)}{m} \leq Wt \log m$ holds, the probability of successful handoff is at least $1 - \frac{1}{\log m} - \frac{1}{(\log m)^2}$. Thus we obtain the throughput as $\frac{W \min(m, n)}{n \log n}$. Then, the near-optimal throughput $\Theta\left(\frac{1}{\log n}\right)$ can be achieved if $\beta \geq 1$. \square

5.4 Delay Analysis

Based on the primary and secondary protocols, the total time consumption of delivering a primary packet

from source to destination can be basically divided into two parts: 1) **The flight time** T_{fly} which the packet costs from source to destination; 2) **The waiting time** T_{wait} which the packet costs at the primary source node to catch a proper SU. Here we neglect the transmission delay due to use of fluid model.

5.4.1 The Flight Time

We use the following lemma to show how long each primary packet will at most take to travel from the source to destination.

Lemma 13. According to the routing process and Lemma 8, the total flight time each primary packet spends traveling from source to destination can be bounded as follow:

$$T_{fly} = \begin{cases} \Theta(\log n / \sqrt{n}) & \alpha = 3/2 \\ \Theta(n^{1-\alpha}) & 1 < \alpha < 3/2 \\ \Theta(1/\log n) & \alpha = 1 \\ \Theta(1) & 0 \leq \alpha < 1 \end{cases} \quad (26)$$

Proof: As stated in the secondary protocol, after each handoff, the distance between a primary packet and its destination decreases by at least a factor of $\sqrt{3}$. In addition, Lemma 8 gives the bound of number of handoffs for each primary packet as $\log n$. Given the maximum primary S-D distance D which is bounded as $E(\|X_i - Y_i\|)$. Then the total time of traveling from source to destination can be denoted as $\frac{2D(1-3^{-\frac{\log n}{2}})}{v(\sqrt{3}-1)}$. Substituting $E(\|X_i - Y_i\|)$ and $\log n$ for D and N^* , respectively. We complete the proof. \square

5.4.2 The Waiting Time

We assume that the transmission from S to M_0 is always successful due to the TDMA scheduling scheme and the setting of preservation region. Here arise the question: how long does a primary source node have to wait until M_0 enters $a_p(n)$ (the primary cell of S).

Lemma 14. Considering a particular primary source node S , we exploit T_{wait} to denote the time which S waits until a proper SU M_0 first enters $a_p(n)$. Then we have:

$$\Pr \left[T_{wait} > \frac{4\sqrt{\log n}}{v\sqrt{n}^\beta} \right] < e^{-\log n \sqrt{n}^{\beta-1}} \quad (27)$$

Proof: The proof is similar to Lemma 4 in [16], so we omit it for simplification. \square

Now, we directly show the total delay results below,

Theorem 8. According to Lemma 13 and Lemma 14, the delay of primary networks with a cooperative mobile secondary networks is calculated as follow:

$$D_{pm} = T_{fly} + T_{wait} = \begin{cases} \Theta(\log n / \sqrt{n}) & \alpha = 3/2 \\ \Theta(n^{1-\alpha}) & 1 < \alpha < 3/2 \\ \Theta(1/\log n) & \alpha = 1 \\ \Theta(1) & 0 \leq \alpha < 1. \end{cases} \quad (28)$$

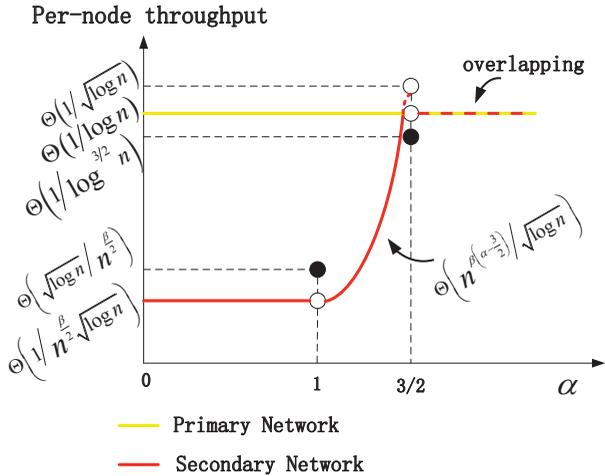


Fig. 4: Per-node throughput of primary networks vs secondary networks in HCRN with cooperative static SUs.

6 DISCUSSION

6.1 HCRN with Cooperative Static SUs

In this section, we will discuss how the heterogeneous distribution of S-D and the cooperation between primary and secondary users affect the throughput and delay of HCRN with cooperative static SUs. Specifically, α measures the S-D distance. The different ranges of α result in different S-D distances and thus affecting the network performance. We can easily find that by introducing the cooperation between primary and secondary users, the per-node throughput of primary networks can be improved to be $\Theta(1/\log n)$ compared with $\Theta(1/\sqrt{n \log n})$ as being a stand-alone network. Further, the number of secondary users required to help improving the PUs throughput can be sharply reduced due to the heterogeneity issue.

6.1.1 Near Optimal Performance of Primary Networks

As defined in the protocols, two-thirds of a secondary time frame are allocated to serve the primary users with the time-sharing strategy. Intuitively, the per-node throughput of primary networks should have varied in accordance with that of secondary networks as the S-D distance changes. However, from Fig. 4 we will discover that the primary network can achieve the near optimal per-node throughput of $\Theta(\frac{1}{\log n})$, which is independent of α . The reason is that the SUs are utilized to help PUs transmit data. By setting appropriate value of N , which is defined in Equation (4), the throughput of $\Theta(\frac{1}{\log n})$ can always be achieved in whatever ranges of α . It is notable that when $\alpha > \frac{3}{2}$ or $\alpha = \frac{3}{2}$, the primary network not only can achieve the near optimal per-node throughput but also can obtain the near optimal delay of $\Theta(\log n)$ and $\Theta(\sqrt{\log n})$ respectively, which is demonstrated in Table 2.

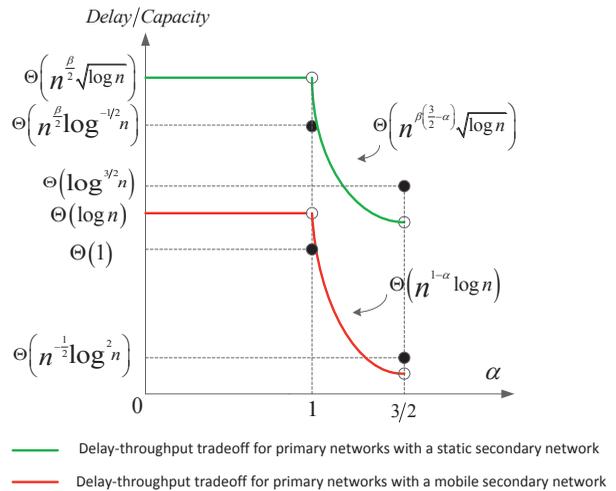


Fig. 5: A comparison of delay-throughput tradeoff

6.1.2 Economical Deployment of HCRN

Table 2 shows the throughput and delay of primary networks and we also put the required number of SUs on the table. It is easy to find that as the S-D distance tends to be shorter, the number of SUs required to help PUs becomes smaller, where $\delta, \delta', \delta'', \varepsilon, \varepsilon'$ can be chosen as arbitrary small positive values. Noting that when $0 \leq \alpha < 1$, the S-D distance is bounded by $\Theta(1)$ and the number of SUs is $n^{2+\varepsilon'}$ which turns to be the same result as that in [12]. So our result is more general compared with that in [12]. More remarkable, as the number of SUs will vary due to the heterogeneous S-D distribution, we can appropriately choose the number of SUs thus making the deployment of HCRN economical in practical scenarios.

6.2 HCRN with Cooperative Mobile SUs

In this section, we investigate the impact of heterogeneous S-D distribution on the throughput and delay of primary networks under the circumstance where the SUs are assumed to be pure relays and mobile. It is worth noting that even we assume the SUs provide relay service and simultaneously producing data themselves, the results would not change in order sense. From Theorem 7, we can see that the near optimal throughput $\Theta(\frac{1}{\log n})$ is achieved in any case of α . Specifically, in the scenario of HCRN with cooperative static SUs, the throughput of secondary networks is severely constrained by the heterogeneous S-D distribution, which alters with α . Thus we should use various number of cooperative SUs in accordance with different ranges of α , for achieving the near-optimal throughput performance. However in the scenario of HCRN with cooperative mobile SUs, due to the introduction of mobility into secondary networks, we only need the condition of $\beta > 1$ for achieving the near-optimal throughput.

TABLE 2: Throughput and delay of PUs and the number of SUs in HCRN with cooperative static SUs

Ranges of α and β	S-D distance	Throughput	Delay	Number of SUs
$\alpha > \frac{3}{2}, \beta > 1$	$1/\sqrt{n}$	$\Theta\left(\frac{1}{\log n}\right)$	$\Theta(\log n)$	$\Theta(n^{1+\delta})$
$\alpha = \frac{3}{2}, \beta > 1$	$\log n/\sqrt{n}$	$\Theta\left(\frac{1}{\log n}\right)$	$\Theta(\sqrt{\log n})$	$\Theta(n^{1+\delta'})$
$1 < \alpha < \frac{3}{2}, \beta \geq 1/(\alpha - \frac{1}{2})$	$n^{1-\alpha}$	$\Theta\left(\frac{1}{\log n}\right)$	$\Theta(n^{\beta(\frac{3}{2}-\alpha)}/\sqrt{\log n})$	$\Theta(n^{\frac{1}{\alpha-\frac{1}{2}}+\delta''})$
$\alpha = 1, \beta \geq 2$	$1/\log n$	$\Theta\left(\frac{1}{\log n}\right)$	$\Theta(n^{\frac{\beta}{2}}/\log^{\frac{3}{2}} n)$	$\Theta(n^{2+\varepsilon})$
$0 \leq \alpha < 1, \beta \geq 2$	1	$\Theta\left(\frac{1}{\log n}\right)$	$\Theta(n^{\frac{\beta}{2}}/\sqrt{\log n})$	$\Theta(n^{2+\varepsilon'})$

On the other side, the delay performance of primary networks is affected by the heterogeneous distribution of S-D and is further enhanced due to the mobility of cooperative SUs. According to Theorem 8, the delay of primary networks are decreasing as α becomes larger. Particularly, the constant delay $\Theta(1)$ and near-optimal delay $\Theta\left(\frac{1}{\log n}\right)$ is achieved for $0 \leq \alpha < 1$ and $\alpha = 1$, respectively. And for the case of $1 < \alpha < \frac{3}{2}$ and $\alpha = \frac{3}{2}$, the delay is further decreased as the distance between S-D tends to be even shorter. In Fig. 5, we present the delay-throughput tradeoff of primary networks in HCRN with both cooperative static and mobile SUs, respectively. Due to the heterogeneity of S-D distribution, both the two tradeoffs differs in different ranges of α . Moreover, when $0 < \alpha < 1$, the primary network can achieve the near optimal delay-throughput tradeoff and the tradeoff can be even better when increasing the heterogeneous extent.

7 CONCLUSION

In this paper, we study the throughput and delay of heterogeneous cognitive radio networks (HCRN). We utilize the rank based model to depict the heterogeneity of S-D distribution and the S-D distance varies with respect to α . By assuming that the SUs can provide relay service for PUs, we put forward the corresponding transmission protocols of HCRN with cooperative static and mobile SUs, respectively. In particular, we show that the primary network can achieve the near optimal per-node throughput with the aid of either static or mobile SUs, which is free of the heterogeneous extent of S-D distribution. Meanwhile, providing PUs help can significantly enhance the transmission of primary networks, but will not degrade the performance of static SUs in order sense. We also prove that the delay of primary networks with either static or mobile SUs changes with α , and it can achieve a better result when considering the mobility of SUs. Furthermore, we demonstrate that the number of static SUs required to help PUs can be sharply reduced due to the heterogeneity of HCRN. This result indicates that the heterogeneity can be utilized to make the network deployment much more economical.

REFERENCES

- [1] P. Gupta, P. R. Kumar, "The Capacity of Wireless Networks," in *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388-404, Mar. 2000.
- [2] M. Vu, N. Devroye, M. Sharif, and V. Tarokh, "Scaling laws of cognitive networks," in *Proc. CrownCom*, pp. 2-8, Aug. 2007.
- [3] M. Vu, V. Tarokh, "Scaling laws of single-hop cognitive networks," in *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, pp. 4089-4097, Aug. 2009.
- [4] S.-W. Jeon, N. Devroye, M. Vu, S.-Y. Chung, and V. Tarokh, "Cognitive Networks Achieve Throughput Scaling of a Homogeneous Network," in *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5103-5115, Aug. 2011.
- [5] C. Yin, L. Gao, and S. Cui, "Scaling Laws for Overlaid Wireless Networks: A Cognitive Radio Network versus a Primary Network," in *IEEE/ACM Transactions on Networking*, vol. 18, no. 4, pp. 1317-1329, Aug. 2010.
- [6] N. M. Tichy, M. L. Tushman, and C. Fombrun, "Social Network Analysis for Organizations," in *The Academy of Management Review*, vol. 4, no. 4, pp. 507-519, 1979.
- [7] Y.-Y. Ahn, S. Han, H. Kwak, S. Moon, and H. Jeong, "Analysis of Topological Characteristics of Huge Online Social Networking Services," in *16th International Conference on World Wide Web*, New York, USA, pp. 835-844, 2007.
- [8] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, A. Tomkins, and R. L. Graham, "Geographic Routing in Social Networks," in *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 33, pp. 11623-11628, 2005.
- [9] R. Kumar, D. Liben-Nowell, et.al, "Theoretical Analysis of Geographic Routing in Social Networks," in *MIT-CSAIL-TR-2005-040*, June 2005.
- [10] B. Azimdoost, H. R. Sadjadpour, J. J. Garcia-Luna-Aceves, "The Impact of Social Groups on The Capacity of Wireless Networks," in *Proceedings of IEEE Network Science Workshop(NSW)*, pp. 30-37, 2011.
- [11] L. Fu, W. Huang and X. Wang, "Capacity of wireless social networks," Shanghai Jiao Tong University, Tech. Rep., 2011. [Online]. Available: <http://iwct.sjtu.edu.cn/Personal/xwang8/paper/social-capacity.pdf>
- [12] L. Gao, R. Zhang, C. Yin, and S. Cui, "Throughput and Delay Scaling in Supportive two-tier Networks," in *IEEE Journal on Selected Areas of Communications*, vol. 30, no.2, pp. 415-424, 2012
- [13] R. Jia, J. Zhang, X. Wang, X. Tian, and Q. Zhang, "Scaling Laws for Heterogeneous Cognitive Radio Networks with Cooperative Secondary Users," to appear in *Proceedings of IEEE INFOCOM, Toronto*, Apr, 2014
- [14] Mathias Grossglauser and David Tse, "Mobility increases the capacity of ad-hoc wireless networks", in *Proceedings of IEEE INFOCOM, Anchorage, AK*, Apr, 2001
- [15] M. J. Neely and E. Modiano, "Capacity and Delay Tradeoffs for Ad Hoc Mobile Networks," in *IEEE Trans. Inform. Theory*, Vol. 51, No. 6, pp. 1917-1937, Jun. 2005
- [16] N. Bansal and Z. Liu, "Capacity, Delay and Mobility in wireless ad-hoc networks", in *Proceedings of IEEE INFOCOM, San Francisco, CA*, Apr, 2003
- [17] X. Lin, G. Sharma, R. R. Mazumdar, and N. B. Shroff, "Degenerate delay-capacity tradeoffs in ad-hoc networks

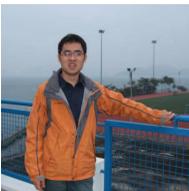
- with Brownian mobility”, in *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2777-2784, June 2006
- [18] K. Lee, Y. Kim, S. Chong, I. Rhee, and Y. Yi, “Delay-capacity tradeoffs for mobile networks with Levy walks and Levy flights”, in *Proceedings of IEEE INFOCOM, Shanghai, Apr, 2011*
- [19] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, “Optimal throughput-daly scaling in wireless networks-part I: the fluid model,” in *IEEE Transactions on Information Theory*, vol. 52, no.6, pp. 2568-2592, June 2006
- [20] M. Franceschetti, O. Dousse, D. N. C. Tse, and P. Thiran, “Closing the Gap in the Capacity of Wireless Networks Via Percolation Theory,” in *IEEE Transactions on Information Theory*, vol. 53, no. 3, pp. 1009-1018, Mar. 2007.
- [21] R. Motwani and P. Raghavan, “Randomized Algorithms.” Cambridge University Press, 1995.



Riheng Jia received his B. E. degree in Electronics and Information Engineering from Huazhong University of Science and Technology, China, in 2012. He is currently pursuing the PHD degree in Electronic Engineering in Shanghai Jiao Tong University. His research of interests are in the area of scaling laws analysis in wireless networks.



Jinbei Zhang received his B. E. degree in Electronic Engineering from Xidian University, Xi'an, China, in 2010, and is currently pursuing the Ph.D. degree in electronic engineering at Shanghai Jiao Tong University, Shanghai, China. His current research interests include network security, capacity scaling law and mobility models in wireless networks.



Feng Yang received his Ph.D degree in Information and Communication from Shanghai JiaoTong University. Since 2008, he has been on the faculty at ShangHai JiaoTong University, where he is an associate professor of electronic information. His research interest include wireless video communication, multihop communication. By now he is the PI of some national projects, including the National High Technology Research and Development Program of China (863 Program),

National Natural Science Foundation of China.



Xiaoying Gan received her Ph. D degrees in Electronic Engineering from Shanghai Jiao Tong University, Shanghai, China in 2006. She is currently with Institute of Wireless Communication Technology, at Department of Electronic Engineering, Shanghai Jiao Tong University (SJTU), where she is an Associate Professor. From 2009 to 2010, she worked as a visiting researcher at California Institute for Telecommunications and Information Technology (Calit2), University of

California San Diego, CA, U.S. Her current research interests include Heterogenous Cellular Network, Cognitive Network, Multi-user Multichannel Access and Dynamic Radio Resource Management.



Xiaohua Tian received his B.E. and M.E. degrees in communication engineering from Northwestern Polytechnical University, Xian, China, in 2003 and 2006, respectively. He received the Ph.D. degree in the Department of Electrical and Computer Engineering (ECE), Illinois Institute of Technology (IIT), Chicago, in Dec. 2010. He is currently an Assistant Professor in Department of Electronic Engineering of Shanghai Jiao Tong University, China. He won the Highest Standards of Academic Achievement 2011 of IIT and Fieldhouse Research Fellowship 2009 of IIT, which is awarded to only one student over IIT each year. His research interests include application-oriented networking, Internet of Things and wireless networks. He serves as the guest editor of International Journal of Sensor Networks, publicity co-chair of WASA 2012. He also serves as the Technical Program Committee member for Wireless Networking Symposium, Ad Hoc and Sensor Networks Symposium of IEEE GLOBECOM 2013, Ad Hoc and Sensor Networks Symposium of IEEE ICC 2013, Wireless Networking Symposium of IEEE GLOBECOM 2012, Communications QoS, Reliability, and Modeling Symposium (CQRM) of GLOBECOM 2011, and WASA 2011.



Pengyuan Du received his B. E. degree in Electronic Engineering from Shanghai Jiao Tong University, China, in 2013. His research interests include the area of asymptotic analysis of capacity in wireless ad hoc networks.



Xinbing Wang (SM'12) received the B.S. degree (with honors) in automation from Shanghai Jiao Tong University, Shanghai, China, in 1998, the M.S. degree in computer science and technology from Tsinghua University, Beijing, China, in 2001, and the Ph.D. degree with a major in electrical and computer engineering and minor in mathematics from North Carolina State University, Raleigh, in 2006. Currently, he is a professor with the Department of Electronic Engineering, Shanghai Jiao Tong University. His research interests include resource allocation and management in mobile and wireless networks, TCP asymptotics analysis, wireless capacity, cross-layer call admission control, asymptotics analysis of hybrid systems, and congestion control over wireless ad hoc and sensor networks. Dr. Wang has been a member of the Technical Program Committees of several conferences including ACM MobiCom 2012, ACM MobiHoc 2012, IEEE INFOCOM 2009-2013.