On the Throughput Scaling of Cognitive Radio Ad Hoc Networks

Chengzhi Li and Huaiyu Dai
Department of Electrical and Computer Engineering
North Carolina State University, Raleigh, NC
email: {cli3, hdai}@ncsu.edu

Abstract—Due to the emergence of Cognitive Radio, a special type of heterogeneous networks attracts increasing interest recently, in which a secondary network composed of cognitive users shares the same resources opportunistically with a primary network of licensed users. Network throughput in this setting is of essential importance. Some pioneer works in this area showed that this type of heterogeneous networks performs as well as two stand-alone networks, under the dense network model where the size of a network grows with the node density in a fixed area. A key assumption behind this conclusion is that the density of the secondary network is higher than that of the primary one in the order sense, which essentially decouples the two overlaid networks, as the secondary network dominates asymptotically. In this paper we endeavor to investigate this problem with a weaker condition that the dimensions of the two overlaid networks are on the same order, and consider the extended network model where the size of a network scales with the area with the node density fixed. Surprisingly, our analysis shows that this weaker (and arguably more practical) condition does not degrade either network throughput in terms of scaling law. Our result further reveals the potentials of CR technology in real applications.

I. INTRODUCTION

The conflict between increasing demands for bandwidth and scarcity of spectrum in wireless communication strongly propels the study of Cognitive Radio (CR) technology in recent years [1]. This technology aims at providing a flexible way of spectrum management, permitting CR (secondary) users to temporally access spectrum that is not currently used by legacy (primary) users. In many cases, it is preferable that the operation of the secondary network is transparent to the primary network. Due to its secondary role, the CR network should prevent any unacceptable interference to the primary network, while tolerate the interference from the primary transmissions.

The capacity of a single ad hoc network has been extensively explored since the seminal work of Gupta and Kumar [2]. The capacity study of CR ad-hoc networks is more challenging in nature due to coexistence of multiple (typically two) networks. Here is a brief overview of some important works in this area. The cognitive channel, an interference channel composed of one S-D pair from each of the two networks, was studied in [3, 4] from an information-theoretic perspective. The throughput of a one-hop cognitive network was investigated in [5]. Recently the throughput of multi-hop CR networks was considered in [6], where the authors showed that an \( n \)-node primary network and an \( m \)-node secondary network, while coexisting, can achieve the per-node throughput \( \Theta(\sqrt{\frac{1}{m \log m}}) \) and \( \Theta(\sqrt{\frac{1}{n \log n}}) \), respectively.

The same results were achieved in [7] with a more practical assumption, requiring only the knowledge of the locations of primary transmitters, rather than the locations of all the primary nodes as in [6]. Note that the same results could be trivially achieved if cooperations between two networks are permitted. Therefore, as [6, 7] we preclude such an option in our study.

In literature, both the dense network model and extended network model are widely used in the study of wireless ad hoc networks. In the dense network model, the number of nodes grows with node density in a fixed area, while in the extended network model, the number of nodes grows with the area with a fixed density. In many scenarios, the (scaling law) results obtained in these two models coincide, but it is felt that the extended network model is more realistic, where one is free from the concern of the near field effects of electromagnetic propagation [8]. Both [6] and [7] focus on the dense network model and assume that \( m = n^\beta \ (\beta > 1) \), i.e., the density of the secondary network is higher than that of the primary one in the order sense. Under this assumption, the secondary network dominates asymptotically, and the influence of the primary network on the secondary network becomes increasingly negligible. In addition, this assumption is only valid in the dense network model and does not hold in the extended network model. These facts motivate us to study a more practical scenario where \( m/n = \text{constant} \) in this paper, i.e., the dimensions of the two overlaid networks are on the same order. We mainly consider an extended network and show that even with this weaker condition the throughput of both networks can be further boosted to \( \Theta(\sqrt{\frac{1}{m \log m}}) \) and \( \Theta(\frac{1}{\sqrt{m \log m}}) \), respectively, with the help of intelligent design of “highway” systems originally proposed in [9]. Different from [9] where the network is tessellated into independent cells, cells in the tesselated CR network are dependent, which complicates the analysis.

The remainder of this paper is organized as follows. The system model and problem formulation are given in Section II. The primary and secondary protocols are presented in Section...
III, together with our main results. The throughput analysis on the secondary network is given in Section IV. And Section V concludes the paper.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider two time-synchronized overlaid extended networks. Suppose that a primary network is deployed in a square $S_n$ with dimension $[0, \sqrt{n}] \times [0, \sqrt{n}]$, according to a Poisson Point Process (PPP) with unit density, i.e., $\lambda_p = 1$, and the secondary network is distributed in the same square, according to a PPP with a constant density $\lambda_s = m/n$, where $m > n$, sharing the same resources with the primary network. It is easy to check, according to the properties of PPP, that the number of primary and secondary nodes lie in $((1 - \epsilon)n, (1 + \epsilon)n)$ and $((1 - \epsilon)m, (1 + \epsilon)m)\forall \epsilon > 0$, with high probability (w.h.p.\textsuperscript{1}). Therefore, their ratio is approximately equal to $n/m$.

Each primary node is paired with another one uniformly at random to form a source-destination pair so that each node is the destination of exactly one source. The secondary source-destination pairs are randomly grouped similarly. The primary network performs as if it stands alone, while the secondary network accesses the spectrum opportunistically to prevent unacceptance interference to the primary network.

Identical transmission power $P_p$ for all the primary nodes, and $P_s$ for all the secondary nodes, are assumed. For simplicity, we only consider path loss for the physical channels (as in the majority literature), i.e., the power attenuation function is given by:

$$l(d) = \min\{1, d^{-\alpha}\},$$

where $d$ is the Euclidean distance between a transmitter and a receiver, and $\alpha > 2$ is the path loss exponent.

The transmission rate $R$ from a transmitter $X_i$ to its corresponding receiver $X_{D(i)}$ is a continuous function of the Signal to Interference plus Noise Ratio (SINR) at $X_{D(i)}$, i.e.,

$$R(X_i, X_{D(i)}) = \log(1 + SINR).$$

Denote by $\{ X_{p,k}; k \in T_1 \}$ the subset of concurrent primary transmitters, and $\{ X_{s,k}; k \in T_2 \}$ the subset of concurrent secondary transmitters. For a primary transmitter $X_{p,i}$ the SINR at its receiver $X_{p,D(i)}$ is given by:

$$SINR_{p} = \frac{P_p l(\| X_{p,i} - X_{p,D(i)} \|)}{N_0 + I_p + I_{sp}},$$

where $N_0$ is the noise power at the receiver; $I_p$ is the interference power from all the other primary transmitters to the receiver $X_{p,D(i)}$, given by:

$$I_p = P_p \sum_{k \in T_1, k \neq i} l(\| X_{p,k} - X_{p,D(i)} \|),$$

and $I_{sp}$ is the interference power from all the secondary transmitters to the receiver $X_{p,D(i)}$, given by:

$$I_{sp} = P_s \sum_{k \in T_2} l(\| X_{s,k} - X_{p,D(i)} \|).$$

\textsuperscript{1}with probability approaching 1 as the the number of nodes in a network goes to infinity.

The SINR at a secondary receiver is defined similarly as:

$$SINR_s = \frac{P_s l(\| X_{s,i} - X_{s,D(i)} \|)}{N_0 + I_s + I_{ps}},$$

where $I_s$ is the interference power from all the other secondary transmitters to the receiver $X_{s,D(i)}$, and $I_{ps}$ is the interference power from all the primary transmitters to the receiver $X_{s,D(i)}$, defined similarly as Eq. (4) and (5).

We study the throughput of both networks, which is the average number of bits per second that all source nodes can simultaneously transmit to their destinations w.h.p. Formally the throughput of a network of size $n$ is defined as:

**Definition 1:** The throughput per S-D pair $T_{Pi}(n)$ in a network of size $n$ under some scheduling scheme $\Pi$ is defined as the maximal quantity satisfying

$$\Pr \left( \min_{i} \lim_{t \rightarrow \infty} \inf \frac{1}{t} B_{i,\Pi}(t) \geq T_{Pi}(n) \right) \rightarrow 1,$$

as $n \rightarrow \infty$, where $B_{i,\Pi}(t)$ is the number of bits that S-D pair $i$ can transfer in $t$ time slots.

Note that the above definition of throughput is an asymptotical property, therefore we require the number of nodes in both networks go to infinity (with fixed ratio) in our analysis.

III. NETWORK PROTOCOLS AND MAIN RESULTS

In this section we introduce the protocols for both the primary and secondary network, followed by their achieved throughput.

A. Primary protocol

The primary network adopts the highway system proposed in [9]. Rooted from percolation theory, the highway system is composed of multiple horizontal and vertical paths, and every primary node in the plane can access at least one horizontal and vertical path through one hop (see Fig. 1). The primary protocol is summarized below:

- Tessellation: the square $S_n$ is divided into $\sqrt{n} \times \sqrt{n}$ horizontal corridors, each with dimension $\sqrt{n} \times 2ap \log \frac{\sqrt{n}}{2ap}$, where $ap$ is some constant, independent of $n$. The corridors are then tessellated into diamond cells with side length $ap$ (see Fig. 2); a highway within each corridor, composed of $\Theta(\log \sqrt{n})$ horizontal paths, is built according to [9], using percolation theory (note that only one path is shown in Fig. 1 for simplicity). Similarly we divide the square $S_n$ into $\sqrt{n} \times \sqrt{n}$ vertical corridors, each with dimension $\sqrt{2ap} \log \frac{\sqrt{n}}{2ap} \times \sqrt{n} \log \sqrt{n}$ and $\Theta(\log \sqrt{n})$ vertical paths.

- Routing: three phases are involved in the routing scheme:
  - access phase: the source drives its packet to one of the multiple horizontal paths in the horizontal corridor it is located, through one single hop;
  - express relay: the packet traverses horizontally and then vertically on the highway through multiple hops;

...
Fig. 1: The highway system

- delivery phase: the packet is finally delivered to the destination by a node on a vertical path, through one single hop.

- Transmission power: the transmission power of a primary transmitter is $P_p = P a_p^2$, where $P$ is a constant $^2$, for all three routing stages.

- Scheduling: each time slot is further divided into three sub-slots. For the primary network, the three sub-slots correspond to the three routing phases, shown in Fig. 3. The $\Theta(\log n)$-TDMA scheme is used during the first and third routing phase and the 9-TDMA scheme is adopted during the highway transmission.

B. Secondary protocol

The design of the secondary protocol is challenging. It is required that the secondary network keep its interference to the primary network at an acceptable level, which may limit its operation and performance greatly. Inspired by the highway system design in [9], we propose a non-trivial secondary protocol, which performs as well as the highway system in a stand-alone network, while satisfies the above requirement.

The secondary protocol is summarized as follows:

- Tessellation: the square $S_n$ is divided into $\sqrt{n} / (2a_s \log n / (2a_s))$ horizontal corridors, each with dimension $\sqrt{\frac{n}{2a_s \log \frac{n}{2a_s}}}$ where $a_s = c_s \sqrt{\frac{n}{m}}$ for some constant $c_s$. The corridors are then tessellated into diamond cells with side length$^3 a_s$ (see Fig. (4)). Correspondingly we divide the square into $\sqrt{\frac{n}{2a_s \log \frac{n}{2a_s}}}$ vertical corridors, each with dimension $\sqrt{2a_s \log \frac{n}{2a_s}} \times \sqrt{n}$. A highway system is constructed in each corridor and it is shown in Theorem 3 that there are $\Theta(\log m)$ horizontal (vertical) disjoint paths in each of the horizontal (vertical) corridors.

$^2$The same $P$ is used in the secondary protocol to maintain the power ratio between two networks.

$^3$a_s is a constant since $n/m$ is.

- A preservation region (see Fig.5) is set around each primary node, which is defined as a square composed of 9 secondary cells. The purpose of preservation regions is to limit the interference from the secondary transmitters to primary receivers. All the secondary nodes located in the preservation regions must keep “silent”, i.e., they can not serve as transmitters or relays.

- Routing: the routing scheme of the secondary network is similar to that of the primary network, i.e., the routing is composed of three phases: access phase, express relay and delivery phase. Each secondary transmission is repeated for 3 times to avoid excess interference from the primary transmission.

- Transmission power: the transmission power of a secondary transmitter is $P_s = P a_s^2$ for all three routing stages. Therefore the power ratio between the primary and secondary network is $(a_p/a_s)^n$.

- Scheduling: the secondary network shares the same time frame structure with the primary network. The time slot is also divided into three sub-slots (Fig. 3). During the first and third sub-slot all the secondary nodes keep silent. The second sub-slot is further divided into three mini-slots, which correspond to three routing phases of the secondary network. The $\Theta(\log m)$-TDMA scheme is used during the first and third routing phase and the 9-TDMA scheme is adopted during the highway transmission.

According to the primary and secondary protocol above, our
The main results are summarized below.

**Theorem 1:** Under the primary and secondary protocol given in this section, the primary network achieves per-node throughput $\Theta(1/\sqrt{n})$, in the presence of the secondary network.

**Theorem 2:** Under the primary and secondary protocol given in this section, with a constant outage probability $4$, the per-node throughput of the secondary network is $\Theta(\frac{1}{\sqrt{m}})$.

**Remark 1:** As will be seen in our analysis below, the outage probability of the secondary network depends on the ratio $m/n$. The higher the ratio, the lower the outage is. The two theorems above imply that, at least in terms of the scaling law, there is no performance loss for either of the two coexisting networks; this is particularly interesting as to the CR network despite its secondary role.

In the interest of space, we mainly focus on the analysis of the secondary network in this paper with all the proof omitted. The reader is referred to [10] for more details.

**IV. THROUGHPUT ANALYSIS OF THE SECONDARY NETWORK**

In this section, we focus on the analysis of the per-node throughput of the secondary network, which is our primary concern. We first show how to construct a highway with consideration of preservation regions. Then we evaluate the one-hop data rate achievable in each of the three routing phases, based on which we obtain the per-node throughput.

**A. Highway construction**

We describe the approach to build the highway in a horizontal corridor, which applies almost verbatim to a vertical corridor as well. As we mentioned, each (horizontal) corridor is partitioned into diamond cells of side length $a_v$. We call such a cell open (see Fig. (5)) if both of the following two events happen:

1) $E_1$ : there is at least one secondary node in the cell;
2) $E_2$ : there is no primary node in either this cell or its 8 neighboring cells; this event differentiates the CR network with a stand-alone network studied in [3].

Otherwise we call it closed. According to the properties of P.P.P., the probabilities of the two events are given by:

$$p_1 \triangleq P(E_1) = 1 - e^{-c_s^2},$$

$$p_2 \triangleq P(E_2) = e^{-\frac{m}{\sqrt{n}}c_s^2}.$$

The outage probability is defined as the percentage of the secondary S-D pairs which can not be served.

Due to the independence of $E_1$ and $E_2$, a cell is open with probability

$$p = p_1p_2 = (1 - e^{-c_s^2})e^{-\frac{m}{\sqrt{n}}c_s^2}.$$ (8)

Some interesting observations on $p$ are in order:

1) The value of $p$ depends on $c_s$ and the secondary density $m/n$, both of which are constants and irrelevant to the scaling law study.
2) If the assumption $m = n^\beta (\beta > 1)$ in [6] and [7] is imposed here the probability of $E_2$ approaches to 1 asymptotically, and the two overlaid networks essentially decouple.
3) The state of a secondary cell (open or closed) is dependent on the states of its neighboring cells, since one primary node can cause 9 secondary cells closed. In contrast, the cell states in a stand-alone network such as the one considered in [9] are independent from each other.

A key step in highway construction is to find (or ensure the existence of) disjoint paths across the network. For this purpose we map each of the tessellated secondary corridor into the bond percolation model on a $r \times \log r$ grid $G_r$, where $r = \lceil \frac{\sqrt{n}}{\sqrt{2}a_s} \rceil = \lceil \frac{\sqrt{m}}{\sqrt{2}a_s} \rceil$. Edges of the grid are composed of horizontal diagonals of some cells and vertical diagonals of the other cells. The grid $G_r$ is given in Fig. 4, where the dashed lines represent the tessellation of the corridor and the solid lines represent the grid. The states (open or closed) of the edges are the same as their corresponding cells, thus they are not independent either. For a horizontal corridor, an L(eft)R(right) open path corresponds to a sequence of connected open edges on the grid, formally defined below.

**Definition 2:** In a 2-D grid $G \triangleq [0,n] \times [0,\log n]$, let $C_v = (x_v, y_v)$ be the coordinates of a vertex $v$. If there exist a series of vertices $v_1, v_2, ..., v_m$, such that there is an open edge connecting two consecutive vertices and $\forall 1 \leq i < m$ $0 \leq x_{v_{i+1}} - x_{v_i} \leq 1$, $x_{v_1} = 1$ and $x_{v_m} = n$, then the path consisting of these edges is called an LR open path. For each open edge, there is at least one secondary node located in its corresponding cell which stays outside preservation regions. Therefore, an LR open path can be mapped back to a routing path in the network. An LR open path and its corresponding routing path are depicted in Fig. 6, where the dark solid line is an LR open path and the dark dashed line is an actual routing path.

We show in the next theorem that with dependent edges, there are $\Theta(\log m)$ paths in a horizontal corridor. The same
conclusion holds for a vertical corridor.

**Theorem 3:** With large enough \( p \), there are \( \Theta(\log m) \) disjoint LR open paths in each horizontal corridor. The proof follows the same line of [9]. Caution should be taken to deal with the situation that the states of cells are dependent (due to the introduction of preservation regions).

A horizontal (vertical) corridor is further evenly divided into multiple slabs with size \( \sqrt{n} \times h \) (\( h \times \sqrt{n} \)), where \( h = \sqrt{2a_s \log(\sqrt{n}/(\sqrt{2a_s}))} \) is a constant. The value of \( h \) guarantees that the number of slabs are as many as the open paths in a corridor so that the nodes in each of the slabs share exactly one path. The slabs in a horizontal corridor are shown in Fig. 7, where there are 5 slabs corresponding to 5 routing paths.

All the nodes in a cell share one access point, which is the closest node located on their corresponding horizontal path. And in the access phase, source nodes load data to their access points through a single hop with length at most \( \sqrt{2a_s \log(\sqrt{n}/(\sqrt{2a_s}))} \) (the width of the corridor). The operations in the delivery phase follows a similar way. One hop in a highway has to support the traffic of all the nodes in a slab since all the nodes in a slab drive their data to the same path.

**B. Throughput analysis**

We first calculate the number of nodes in a cell and a slab, and then evaluate the throughput of one hop during each phase.

**Lemma 1:** The number of nodes in a cell is at most \( \log m \), and the number of nodes in a slab is at most \( 2h\sqrt{m/n} \sqrt{m} \), both w.h.p.

The following two lemmas show that one hop on the highway can support a constant rate and the throughput for an access (delivery) link scales as \( \Omega(1/(\log m)^\alpha) \).

**Lemma 2:** On the highway of the secondary network, each secondary cell can support traffic with a constant rate \( K_s \), where \( K_s > 0 \) is independent of \( m \).

**Lemma 3:** The throughput for an access (delivery) link scales as \( \Omega(1/(\log m)^\alpha) \).

During the access and delivery phase, each link supports \( \Omega(1/(\log m)^\alpha) \) data rate (Lemma 3) and \( \log m \) nodes (Lemma 1) share it. Therefore, the per-node throughput during these two phases scales as \( \Omega(1/(\log m)^{\alpha+1}) \). And on the highway one hop can support a constant data rate (Lemma 2) and there are at most \( 2h\sqrt{m/n} \sqrt{m} \) nodes (Lemma 1) sharing the same hop. Therefore the per-node throughput on the highway scales as \( \Theta(1/\sqrt{m}) \). Considering the throughput of all the phases we reach the conclusion of Theorem 2.

**C. Outage analysis**

According to the secondary protocol, outage occurs inevitably since any S-D pairs composed of the secondary nodes in the preservation regions can not be served. Denote by \( m_u \) the number of secondary nodes in all the preservation regions, which is given by \( m_u \leq 2(1+\epsilon) m \sqrt{n} 2a_s m^{\frac{1}{2}} \) w.h.p., due to the fact that the total area of preservation regions is bounded by \( (1+\epsilon) m^{\frac{1}{2}} s^2 \). Therefore, the outage probability \( P_o \) can be calculated as follows:

\[
P_o = \frac{m_u}{\text{the number of secondary nodes}} \leq \frac{2(1+\epsilon) m \sqrt{n} 2a_s m^{\frac{1}{2}}}{(1-\epsilon) m} = \frac{18(1+\epsilon) 2a_s}{1-\epsilon} \triangleq \delta,
\]

where \( \delta < 1 \) is a constant depending on \( n/m \). Therefore, higher secondary density \( m/n \) leads to lower outage probability. Note that if the higher order condition \( m = n^\beta (\beta > 1) \) is applied, the outage probability is vanishing, w.h.p., which coincides with the conclusion in [6].

**V. conclusions**

In this paper, we have studied the throughput of a type of heterogeneous networks consisting of a primary network of size \( n \) and a cognitive radio ad hoc network of size \( m \) under a more practical model. We show that the per-node throughput of the primary network scales as \( \Theta(1/\sqrt{n}) \) and there is indeed no performance loss, in terms of scaling law, with the coexistence of the CR network. The CR network can also achieve per-node throughput \( \Theta(1/\sqrt{m}) \), performing as well as a stand-alone network except suffering from a non-vanishing outage probability.

**References**