This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Capacity Scaling of General Cognitive Networks

Wentao Huang and Xinbing Wang, *Member, IEEE*

***Abstract*—There has been recent interest within the networking research community to understand how performance scales in cognitive networks with overlapping n primary nodes and m secondary nodes. Two important metrics, i.e., throughput and delay, are studied in this paper. We first propose a simple and extendable decision model, i.e., the hybrid protocol model, for the secondary nodes to exploit spatial gap among primary transmissions for frequency reuse. Then, a framework for general cognitive networks is established based on the hybrid protocol model to analyze the occurrence of transmission opportunities for secondary nodes. We show that if the primary network operates in a generalized TDMA fashion, or employs a routing scheme such that traffic flows choose relays independently, then the hybrid protocol model suffices to guide the secondary network to achieve the same throughput and delay scaling as a standalone network without harming the performance of the primary network, as long as the secondary transmission range is smaller than the primary range in order. Our approach is general in the sense that we only make a few weak assumptions on both networks, and therefore it obtains a wide variety of results. We show secondary networks can obtain the same order of throughput and delay as standalone networks when primary networks are classic static networks, networks with random walk mobility, hybrid networks, multicast networks, CSMA networks, networks with general mobility, or clustered networks. Our work presents a relatively complete picture of the performance scaling of cognitive networks and provides fundamental insight on the design of them.**

***Index Terms*— Capacity, cognitive.**

I. INTRODUCTION

THE ELECTROMAGNETIC radio spectrum is a natural resource, the use of which by transmitters and receivers is licensed by governments. Today, as wireless applications demand ever more bandwidth, efficient usage of spectrum is becoming necessary. However, recent measurement [2] observed a severe underutilization of the licensed spectrum, implying the nonoptimality of the current scheme of spectra management. As a remedy, the Federal Communications Commission (FCC) has recently recommended [2], [3] more flexibility in spectrum assignment so that new regulations would allow for devices that are able to sense and adapt to their spectral environment, such as *cognitive radios*, to become *secondary* or *cognitive users*. Cognitive users could opportunistically access the spectrum originally licensed to *primary users* in a manner in which their transmissions will not affect the performance of primary users. Primary users have a higher priority to the spectrum; they may be legacy devices and may not cooperate with secondary users. The overlapping *primary network* and *secondary network* together form the *cognitive network*.

This paper focuses on the performance scaling analysis of cognitive networks with an increasing number of n primary users and m secondary users. The fundamental scaling laws in ad hoc networks has attracted tremendous interest in the networking community for long. This track of research is initiated by Gupta and Kumar, whose landmark work [4] showed that generally the per-node throughput capacity of a wireless ad hoc network with n users only scales as $O(1/\sqrt{n})$.[1] Following works have covered a wide variety of ad hoc networks with different features, such as mobile ad hoc networks (MANETs) [5], hybrid networks [6], [7], multicast networks [8], [9], hierarchically cooperative networks [10], clustered networks [11], [12], etc. Performance metrics other than capacity are also studied, among which delay and its optimal tradeoff with throughput are of critical importance [13], [14].

As with most related works, under the Gaussian channel model, Jeon *et al.* [15] considered the capacity scaling of a cognitive network where the number of secondary users, m, is larger than n in order. Under a similar assumption, Yin *et al.* [16] developed the throughput–delay tradeoff of both primary and secondary networks, and Wang *et al.* [17] studied the cases of multicast traffic pattern. Interestingly, all these works showed that both primary and secondary networks can achieve similar or same performance bounds as they are standalone networks.

All previous works on cognitive networks [15]–[17] considered some particular scenarios. Typically, they first assumed some particular primary networks with specific scheduling and routing protocols, then proposed the communication schemes for secondary users accordingly, and lastly showed such schemes suffice to achieve the same performance bounds as standalone networks. However, a key principle of cognitive networks is that primary users are spectrum license holders and may operate at their own will without considering secondary nodes. Therefore, though assuming a specific primary network can simplify the problem, the results will heavily depend on the communication schemes of the primary network, which is often unmanageable.

[1]Recall that: 1) $f(n) = O(g(n))$ means that there exists a constant c and integer N such that $f(n) \leq cg(n)$ for $n > N$; 2) $f(n) = o(g(n))$ means that $\lim_{n\to\infty} f(n)/g(n) = 0$; 3) $f(n) = \Omega(g(n))$ means that $g(n) = O(f(n))$; 4) $f(n) = \omega(g(n))$ means that $g(n) = o(f(n))$; 5) $f(n) = \Theta(g(n))$ means that $f(n = O(g(n)))$ and $g(n) = O(f(n))$.

Manuscript received August 24, 2011; revised December 05, 2011; accepted December 10, 2011; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor A. Capone. This work was supported by National Fundamental Research Grant 2011CB302701, NSF China under Grant 60832005, the China Ministry of Education New Century Excellent Talent under Grant NCET-10-0580, the China Ministry of Education Fok Ying Tung Fund under Grant 122002, a Qualcomm Research Grant, the Shanghai Basic Research Key Project under Grant 11JC1405100, and National Key Project of China under Grant 2010ZX03003-001-01. An earlier version of this paper appeared in the Proceedings of the IEEE International Conference on Computer Communications (INFOCOM), Shanghai, China, April 10–15, 2011.

The authors are with Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: yelohuang@sjtu.edu.cn; xwang8@sjtu.edu.cn).

Digital Object Identifier 10.1109/TNET.2011.2180400

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

That motivates us to study a general cognitive network in this paper. Our major contributions are threefold. First, we characterize the regime that cognitive networks can achieve the same order of throughput and delay scaling as standalone networks. Second, we propose a simple decision model for secondary users to identify transmission opportunities and, based on it, establish a framework with which schemes of standalone networks can be readily extended to secondary networks. Third, we apply the framework to various specific scenarios and show that secondary networks can obtain the same order of throughput and delay as standalone networks when primary networks are classic static networks, networks with random walk mobility, hybrid networks, multicast networks, CSMA networks, networks with general mobility, or clustered networks.

In particular, when all of the following three conditions hold, it is sufficient for a general cognitive network to achieve the same throughput and delay bounds as standalone networks.

A1) The cognitive network is subject to the physical interference model. The primary network operates at a signal-to-interference-plus-noise ratio (SINR) level larger than the threshold for successful reception by some small allowance.

A2) The primary network is scheduled in a generalized round-robin TDMA manner, or traffic flows of the primary network choose relays independently for routing.

A3) Scheduling schemes of the secondary network follow $r_{\max}(m) = o(R_{\min}(n))$ and $r_{\max}^{\gamma-2} = o(R_{\min}^{\gamma}/R_{\max}^{2})$ with high probability, where $R(n)$ and $r(m)$ are the transmission ranges of primary and secondary networks, and γ is the path loss exponent.

Intuitively, condition A1 ensures that primary transmission links are neither too dense nor too vulnerable so that there exist opportunities for secondary users. Such opportunities will frequently appear, as a consequence of A2. The first equation of A3 is the generalization of the condition $m = \omega(n)$ in related works, while the second equation is more technical. It characterizes the case that the scheduling of primary networks is somewhat "homogeneous" such that there exists a simple rule for opportunity decision. A1 is based on the physical interference model, and similar results hold under the Gaussian channel model.

We note this paper is not merely a generalization of results from previous works. Our work shows the fact that cognitive networks, and especially secondary networks, can achieve the same throughput and delay scaling as standalone networks is mainly determined by the underlying interference model, and it only weakly relies on the specific settings such as scheduling and routing protocols of primary networks. Such insight is fundamental and implies that for quite general cases, "cognitive" will not be a handicap to performance scaling.

This paper is organized as follows. In Section II, we introduce system models and formalize the operation rules of cognitive networks, and Section III presents an overview of our solution. We propose the hybrid protocol model and establish its physical feasibility in Section IV. Section V identifies the conditions under which the secondary network will have plenty of transmission opportunities if scheduled according to the hybrid protocol model. We apply the general results to several specific cognitive networks in Section VI, and Section VII concludes the paper.

TABLE I
IMPORTANT NOTATIONS

Notation	Definition
X_i	position of primary user i
Y_j	position of secondary user j
$\mathscr{P}(\alpha, \beta)$	feasible family of physical model
$\mathscr{D}(\alpha + \epsilon)$	feasible family of primary scheduler
$\mathscr{Q}_p(\Delta_p), \mathscr{Q}_s(\Delta_s)$	feasible family of protocol model
$\mathscr{H}(\Delta_p, \Delta_{ps}, \Delta_{sp}, \Delta_s)$	feasible family of hybrid protocol model
$\mathcal{S}^{(p)}$	set of active primary links
$\mathcal{S}^{(s)}$	set of active secondary links
$\mathcal{S}$	$\mathcal{S}^{(p)} \cup \mathcal{S}^{(s)}$
R_i	Tx range of active link $(X_i, X_{\mathrm{Rx}(i)})$
r_j	Tx range of active link $(Y_j, Y_{\mathrm{Rx}(j)})$
P	Tx power of primary network
P_j	Tx power of link $(Y_j, Y_{\mathrm{Rx}(j)})$

II. System Model

Throughout this paper, we denote the probability of event E as $\Pr(E)$ and say E happens with high probability (*w.h.p.*) if $\lim_{n\to\infty} \Pr(E) = 1$. By convention, $\{c_i\}$ denotes positive constants, and $\{C_j(n)\}$ parameters dependent on n. Other notations are defined in Table I.

A. Network Topology

We define the network extension $\mathcal{O}$ to be a unit square. Two kinds of nodes, i.e., the *primary nodes* and the *secondary nodes*, overlap in $\mathcal{O}$. They share the same time, space, and frequency dimensions. Unless further specifications are made, by convention we assume n primary nodes are independently and identically distributed (i.i.d.) in $\mathcal{O}$ according to uniform distribution, and so do the m secondary users. Notice that different from previous related works that require $n = o(m)$, i.e., the primary user density is asymptotically smaller than the secondary user density, we allow the relation between m and n to be arbitrary. Their positions are $\{X_i\}_{i=1}^{n}$ and $\{Y_j\}_{j=1}^{m}$. $\forall i, j, X_i, Y_j \in \mathcal{O}$. At times we may denote a node by its position, i.e., we refer to primary node i and secondary node j as X_i and Y_j, and let $|X_i - Y_j|$ be the distance between them. Two types of nodes form their respective networks, the *primary network* and the *secondary network*. In each network, nodes are randomly grouped into source–destination (S–D) pairs, such that every node is both source and destination, with traffic rate λ. Equivalently, we can describe the traffic pattern in matrix form $\lambda\Lambda$, where $\Lambda = [\lambda_{\mathrm{sd}}]$ is a random permutation matrix[2] with $\lambda_{\mathrm{sd}} \in \{0, 1\}$. Note that we do not consider cross-network traffic. We use index p and s to distinguish quantities between primary nodes and secondary nodes when needed, for example, λ_{p} and λ_{s}.

B. Communication Model

We assume all nodes share a wireless channel with bandwidth W b/s. Assume that path loss exponent is $\gamma > 2$, then the normalized channel gain between a transmitter at location Z_{T} and a receiver at Z_{R} is $G(|Z_{\mathrm{T}} - Z_{\mathrm{R}}|) = |Z_{\mathrm{T}} - Z_{\mathrm{R}}|^{-\gamma}$.

[2] $\Lambda = [\lambda_{\mathrm{sd}}]$ is a permutation matrix if $\forall s, d, \lambda_{\mathrm{sd}} \in \{0, 1\}; \forall d, \sum_{\mathrm{s}} \lambda_{\mathrm{sd}} = 1; \forall s, \sum_{\mathrm{d}} \lambda_{\mathrm{sd}} = 1$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Moreover, wireless transmission may be subject to failures or collisions caused by noise or interference. To judge whether a direct wireless link is feasible, we have the following physical model, whose well-known prototype is proposed in [4]

Physical Model: Let $\{X_i; i \in \mathcal{T}^{(\mathrm{p})}\}$ and $\{Y_j; j \in \mathcal{T}^{(\mathrm{s})}\}$ be the subsets of nodes simultaneously transmitting at some time instant. Let P be the uniform power level of primary network, and P_j be the power chosen by secondary node Y_j, for $j \in \mathcal{T}^{(\mathrm{s})}$. Then, for the primary network, the transmission from node X_i is successfully received by node X_j if

$$\frac{PG(|X_i-X_j|)}{N+\sum_{k\in\mathcal{T}^{(\mathrm{p})}\setminus\{i\}}PG(|X_k-X_j|)+\sum_{l\in\mathcal{T}^{(\mathrm{s})}}P_lG(|Y_l-X_j|)}\geq\alpha \tag{1}$$

where N is ambient noise and constant α characterizes the minimum SINR necessary for successful receptions for primary nodes. For the secondary network, the transmission from node Y_i is successfully received by node Y_j if

$$\frac{P_iG\left(|Y_i-Y_j|\right)}{N+\sum\limits_{k\in\mathcal{T}^{(\mathrm{s})}\setminus\{i\}}P_kG(|Y_k-Y_j|)+\sum\limits_{l\in\mathcal{T}^{(\mathrm{p})}}PG(|X_l-Y_j|)}\geq\beta$$

where constant β is the minimum required SINR for secondary network. Note that we allow secondary users to have more flexible power control ability. This is in accord with the design principle of cognitive radios.

We call a couple of nodes a link if they form a transmitter–receiver pair, e.g., (X_i, X_j). Given an interference model, in general there is a number of subsets of links that can be active simultaneously. We call such subsets of links *feasible states*, and define the set of all feasible states as *feasible family*. We use $\mathscr{P}(\alpha,\beta)$ to denote the feasible family of the physical model.

C. Operation Rules

The essential differences between cognitive networks and normal ad hoc networks are the operation rules. Though primary and secondary users overlap and share the channel, they are different essentially because of their behavior. In principle, primary nodes are spectrum license holders and have the priority to access the channel. It is followed by two important implications. First, primary nodes may operate at their own will without considering secondary nodes. They may be legacy devices running on legacy protocols, which are fixed and unmanageable. Therefore, the assumptions made about primary networks should be as few and general as possible. Moreover, the secondary network, which is opportunistic in nature, should control its interference to the primary network and prevent deteriorating the performance of primary users. The challenge is that the primary scheduler may not alter its protocol due to the existence of the secondary network, and its decision model could be different from the physical model (1), i.e., the interference term from the secondary network in the denominator is not available. However, in order to leave some margin for secondary nodes, it is necessary for the decision model to operate at an SINR larger than α by an allowance ϵ.

Operartion Rule 1: Decision model for the primary network: The primary scheduler considers the transmission from X_i to X_j to be feasible if

$$\frac{PG\left(|X_i-X_j|\right)}{N+\sum_{\substack{k\in\mathcal{T}^{(\mathrm{p})}\\k\neq i}}PG\left(|X_k-X_j|\right)}\geq\alpha+\epsilon.$$

The feasible family of the primary decision model is denoted as $\mathscr{D}(\alpha+\epsilon)$.

Then, as the operation rule, secondary nodes should guarantee that the feasible state under the decision model $\mathscr{D}$ above should be indeed feasible under the physical model.

Operation Rule 2: Decision model for the secondary network: Let $\mathcal{S}^{(\mathrm{p})}$ and $\mathcal{S}^{(\mathrm{s})}$ be the sets of active primary links and active secondary links. If $\mathcal{S}^{(\mathrm{p})}\in\mathscr{D}(\alpha+\epsilon)$, then $\mathcal{S}^{(\mathrm{p})}\cup\mathcal{S}^{(\mathrm{s})}\in\mathscr{P}(\alpha,\beta)$, *w.h.p.*

Note that compared to most existing related literatures where the concept of user priority is usually scheme- or network-specific, Operation Rules 1 and 2 formally define the principle of cognitive behaviors in a general sense.

D. Capacity Definition

Definition 1: Feasible throughput: Per-node throughput $g(n)$ of the primary network is said to be feasible if there exists a spatial and temporal scheme for scheduling transmissions, such that by operating the primary network in a multihop fashion and buffering at intermediate nodes when awaiting transmission opportunities, every primary source can send $g(n)$ b/s to its destination on average.

Definition 2: Asymptotic per-node capacity $\lambda_{\mathrm{p}}(n)$ of the primary network is said to be $\Theta(g(n))$ if there exist two positive constants c and c' such that

$$\begin{cases}\lim_{n\to\infty}\Pr\left\{\lambda_{\mathrm{p}}(n)=cg(n)\text{ is feasible}\right\}=1\\ \lim_{n\to\infty}\Pr\left\{\lambda_{\mathrm{p}}(n)=c'g(n)\text{ is feasible}\right\}<1.\end{cases}$$

Similarly, we can define the asymptotic per-node capacity $\lambda_{\mathrm{s}}(m)$ for the secondary network.

III. Overview of Idea and Solution

Our system model begins with a very classical setup, where the network topology, node communication capabilities, and performance metrics fall in the same framework that is most commonly deployed in related works on asymptotic analysis of wireless networks. An extensive body of literature [4]–[14] has investigated various specific networks under this framework. For that matter, the key issue that we aim to address in this paper is how the cognitive principles, i.e., Operation Rules 1 and 2, may impact network performance, especially with respect to the abundant insights already gained in previous works on asymptotic network capacity and delay.

Clearly this is a nontrivial problem: Operation Rules 1 and 2 have introduced fundamental heterogeneities into the network in the sense that nodes now have different levels of priority. Such heterogeneities are exactly the most essential idea of how cognitive networks operate.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

However, though the two operation rules are ideal definitions for cognitive principles, they are not convenient from the perspective of analysis and practice because, despite their simple forms, they actually involve numerous underlying details such as the whole network topology, transmission power, and aggregate interference in judging the eligibility of even a single link. Therefore, we introduce the *hybrid protocol model*, which loosely speaking is a subset of Operation Rules 1 and 2 in the sense that it is a somewhat "stricter" criterion. The hybrid protocol model is significantly simpler to analyze because it only relies on the geometry of node positions and conceals other details. To establish the correspondence between the operation rules and the hybrid protocol model is the main mission of Section IV, where we design the protocol parameters and the underlying power assignment schemes. Then, we may use the hybrid protocol model instead of the operation rules in later analysis of network performance, only at the cost of losing a marginal portion of secondary transmission opportunities due to the slight inequivalence between the two sets of criteria.

The throughput and delay in a network are dependent on the specific scheduling and routing schemes. In Section V, we consider whether there is a class of scheduling or routing schemes to which the cognitive behaviors are benign. In other words, this implies that under the hybrid protocol model, both the primary and secondary networks can achieve the same order of performance as if they are separate without mutual interference. This is especially important for the secondary users because it indicates that though they are inferior in priority, their performance is still guaranteed.

In particular, we identify two classes of primary network communication schemes, one on scheduling and the other on routing, that satisfy this property. The first class of primary networks is scheduled in a generalized cell-partitioned TDMA manner. In this case, due to the geometric property of the hybrid protocol model, every secondary user may associate itself with a certain primary cell at an appropriate distance away, such that whenever the cell is scheduled to be active, the secondary user may transmit without causing (receiving resp.) destructive interference to (from resp.) the primary network. In the second class of networks, every primary traffic flow (i.e., a source destination pair of the primary network) shall make the routing decision independently of other flows. The main idea is that every primary transmission will not only intuitively mute the secondary users nearby, but will also create certain "gap" such that the secondary users in these regions may be "riggered." Because the traffic is somewhat independently distributed (relayed) in the primary network, if a secondary link is muted for a long time *w.h.p.*, indicating the primary traffic nearby is intense, then this link will also be triggered for a considerable time *w.h.p.*.

Lastly, because these two classes of schemes include most of those proposed or discussed in literature, numerous results therein can be easily extended to the settings of cognitive networks, where Operation Rules 1 and 2 apply. These specific examples are discussed in Section VI.

IV. Identifying Opportunities: The Hybrid Protocol Model

In this section, we consider the problem of how to schedule links in the cognitive network under interference constraint. Recall from operation rules that primary nodes are unmanageable, so in fact the key issue is the schedule strategy for the secondary network. Specifically, we will face two challenges: first, how to ensure that secondary transmissions are harmless to the primary network; second, how to establish a secondary link given uncontrollable interference from the primary network. Our goal is to design a practical decision criteria for the secondary users to address these two seemingly contradictory challenges at the same time. Intuitively, that is to say we should find simpler rules for secondary nodes to hunt and exploit opportunities in the network subject to the operation rules.

A. Hybrid Protocol Model

Since we assume the primary network to be a general network that operates according to decision model $\mathscr{D}(\alpha+\epsilon)$, it is our starting point. $\mathscr{D}$ is of physical concern and cares about the aggregate interference and SINR, but the following lemma relates it to a simpler pairwise model. This alternative model is known as the *protocol model* in literature and often plays the role as interference model. However, here we use it as a tool to characterize the relative position of active primary nodes.

Definition 3: Protocol Model for primary network: A transmission from X_i to X_j is feasible if

$$|X_k - X_j| \geq (1+\Delta_{\mathrm{p}})|X_i - X_j| \qquad \forall k \in \mathcal{T}^{(\mathrm{p})}$$

where Δ_{p} defines the guard zone for the primary network. The corresponding feasible family is noted as $\mathscr{Q}_{\mathrm{p}}(\Delta_{\mathrm{p}})$. Likewise, we define protocol model $\mathscr{Q}_{\mathrm{s}}(\Delta_{\mathrm{s}})$ for the secondary network.

First we need to consider the relation of inclusion between different feasible families.

Lemma 1: If $\mathcal{S}^{(\mathrm{p})} \in \mathscr{D}(\alpha+\epsilon)$ and $\Delta_{\mathrm{p}} \leq (\alpha+\epsilon)^{1/\gamma} - 1$, then $\mathcal{S}^{(\mathrm{p})} \in \mathscr{Q}_{\mathrm{p}}(\Delta_{\mathrm{p}})$.

Proof: Let $\mathcal{S}^{(\mathrm{p})} \in \mathscr{D}$, $\forall (X_i, X_j) \in \mathcal{S}^{(\mathrm{p})}$ and $\forall k \neq i, k \in \mathcal{T}^{(\mathrm{p})}$, holds

$$\frac{|X_i - X_j|^{-\gamma}}{|X_k - X_j|^{-\gamma}} \geq \frac{P|X_i - X_j|^{-\gamma}}{N + P\sum_{\substack{l \in \mathcal{T}^{(\mathrm{p})} \\ l \neq i}} |X_l - X_j|^{-\gamma}} \geq \alpha + \epsilon$$

therefore $|X_k - X_j| \geq (\alpha+\epsilon)^{1/\gamma}|X_i - X_j|$, set $\Delta_{\mathrm{p}} \leq (\alpha+\epsilon)^{1/\gamma} - 1$, then $\mathcal{S}^{(\mathrm{p})} \in \mathscr{Q}_{\mathrm{p}}$. ■

Since $\mathscr{Q}_{\mathrm{p}} \supseteq \mathscr{D}$, i.e., any feasible output of the primary scheduler is also feasible under $\mathscr{Q}_{\mathrm{p}}$ as long as $\Delta_{\mathrm{p}} \leq (\alpha+\epsilon)^{1/\gamma} - 1$, and considering the simplicity of the protocol model, it motivates us to define a new *hybrid protocol model* $\mathscr{H}$ based on $\mathscr{Q}_{\mathrm{p}}$ and $\mathscr{Q}_{\mathrm{s}}$, to be the decision model for the secondary network.

Definition 4: The Hybrid Protocol Model with feasible family $\mathscr{H}(\Delta_{\mathrm{p}}, \Delta_{\mathrm{ps}}, \Delta_{\mathrm{sp}}, \Delta_{\mathrm{s}})$: $\forall \mathcal{S} \in \mathscr{H}$, let $\mathcal{S}^{(\mathrm{p})} = \{(X_i, X_j) \in \mathcal{S}\}$ and $\mathcal{S}^{(\mathrm{s})} = \{(Y_i, Y_j) \in \mathcal{S}\}$, then $\mathcal{S}^{(\mathrm{p})} \in \mathscr{Q}_{\mathrm{p}}(\Delta_{\mathrm{p}})$, $\mathcal{S}^{(\mathrm{s})} \in \mathscr{Q}_{\mathrm{s}}(\Delta_{\mathrm{s}})$. Furthermore, $\forall (X_i, X_j) \in \mathcal{S}^{(\mathrm{p})}$

$$|Y_k - X_j| \geq (1+\Delta_{\mathrm{sp}})|X_i - X_j| \qquad \forall k \in \mathcal{T}^{(\mathrm{s})} \tag{2}$$

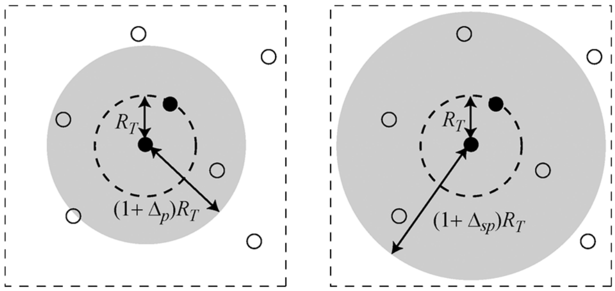


Fig. 1. Examples of the hybrid protocol model: Given an active primary link, the left plot shows the guard zone regarding primary interferers, and the right plot for secondary interferers.

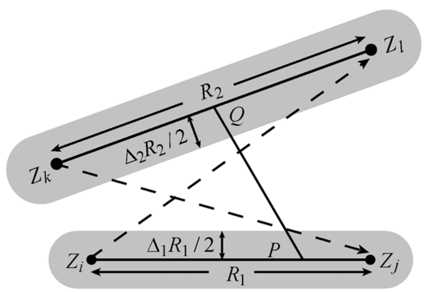


Fig. 2. Disjoint regions of two active transmissions.

and $\forall (Y_i, Y_j) \in \mathcal{S}^{(\mathrm{s})}$

$$|X_k - Y_j| \geq (1+\Delta_{\mathrm{ps}})|Y_i - Y_j| \qquad \forall k \in \mathcal{T}^{(\mathrm{p})} \tag{3}$$

where $\Delta_{\mathrm{sp}}, \Delta_{\mathrm{ps}}$ define internetwork guard zones (Fig. 1).

The hybrid protocol model only depends on pairwise distance between transmitters and receivers. Such simplicity will facilitate our analysis in the next section. Moreover, it is compatible with the classic protocol interference model. Thus, rich communication schemes and results based on the protocol model can be easily extended to cognitive networks, as will be shown in Section V.

In the following, we should prove that if $\mathscr{H}$ is used as a decision model for secondary nodes, it will comply with Operation Rule 2. This involves correctly tuning the parameters Δ_{p}, Δ_{ps}, Δ_{sp}, Δ_{s}, and $\{P_j\}_{j\in\mathcal{T}^{(\mathrm{s})}}$.

B. Interference at Primary Nodes

We first address the challenge that primary transmissions should not be interrupted by secondary nodes. The main task is to bound the interference from the secondary network. We start with a useful property of the hybrid protocol model.

Lemma 2: Given arbitrary $Z_i, Z_j, Z_k, Z_l \in \mathcal{O}$, if (Z_i, Z_j), (Z_k, Z_l) are active links (primary or secondary), and $|Z_k - Z_j| \geq (1+\Delta_1)|Z_i - Z_j|$, $|Z_i - Z_l| \geq (1+\Delta_2)|Z_k - Z_l|$, then the $\Delta_1|Z_i - Z_j|/2$ neighborhood of line segment Z_iZ_j and the $\Delta_2|Z_k - Z_l|/2$ neighborhood of line segment Z_lZ_k are disjoint (See Fig. 2 for an example).

Proof: Let E and F be two arbitrary points on line segment Z_iZ_j and Z_kZ_l, by triangle inequality

$$\begin{aligned}|Z_i - E| + |E-F| + |F-Z_l| + |Z_k - F| + |F-E| + |E-Z_j| \\ \geq |Z_i - Z_l| + |Z_k - Z_j|\end{aligned}$$

Since Z_i, E, Z_j are collinear, substituting lemma condition

$$\begin{aligned}|Z_i - Z_j| + |Z_k - Z_l| + 2|E-F| &\geq |Z_i - Z_l| + |Z_k - Z_j| \\ &\geq (1+\Delta_2)|Z_k - Z_l| + (1+\Delta_1)|Z_i - Z_j|\end{aligned}$$

therefore $|E-F| \geq (\Delta_1/2)|Z_i - Z_j| + (\Delta_2/2)|Z_k - Z_l|$. Now we prove the lemma by contradiction. Suppose the two neighborhoods overlap, then there exist points P on Z_iZ_j and Q on Z_kZ_l and Z such that $|Z-P| < (\Delta_1/2)|Z_i - Z_j|$ and $|Z-Q| < (\Delta_2/2)|Z_k - Z_l|$. Then, $|P-Q| < |Z-P| + |Z-Q| < (\Delta_1/2)|Z_i - Z_j| + (\Delta_2/2)|Z_k - Z_l|$, which is a contradiction.

Corollary 1: Under the hybrid protocol model, we have the following.

- If (X_i, X_j) and (X_k, X_l) are active primary links, the $\Delta_{\mathrm{p}}|X_i - X_j|/2$ neighborhood of line segment X_iX_j and $\Delta_{\mathrm{p}}|X_k - X_l|/2$ neighborhood of X_kX_l are disjoint.
- If (Y_i, Y_j) and (Y_k, Y_l) are active secondary links, the $\Delta_{\mathrm{s}}|Y_i - Y_j|/2$ neighborhood of line segment Y_iY_j and $\Delta_{\mathrm{s}}|Y_k - Y_l|/2$ neighborhood of Y_kY_l are disjoint.
- If (X_i, X_j) is an active primary link and (Y_k, Y_l) is an active secondary link, the $\Delta_{\mathrm{sp}}|X_i - X_j|/2$ neighborhood of line segment X_iX_j and $\Delta_{\mathrm{ps}}|Y_k - Y_l|/2$ neighborhood of Y_kY_l are disjoint.

For active link $(X_i, X_{\mathrm{Rx}(i)})$ and $(Y_j, Y_{\mathrm{Rx}(j)})$, where function Rx indicates the index of receiver, let $R_i = |X_i - X_{\mathrm{Rx}(i)}|$ and $r_j = |Y_j - Y_{\mathrm{Rx}(j)}|$. Notice that R_i in general is a function of n and r_j is a function of m. We say the secondary network adopts *power assignment scheme* $\mathcal{A}(C)$ if for $i \in \mathcal{T}^{(\mathrm{s})}$, $P_i = Cr_i^2 P$. The quadratic power assignment facilitates our effort to upper-bound aggregate interference by converting it to an integral over the network area.

Theorem 1: Under power assignment $\mathcal{A}(C_1)$ and the hybrid protocol model, if $\Delta_{\mathrm{ps}} > \Delta_{\mathrm{s}}$, then for any active primary link $(X_i, X_{\mathrm{Rx}(i)})$, the interference suffered by $X_{\mathrm{Rx}(i)}$ from the secondary network is upper-bounded by $C_2 R_i^{2-\gamma} P$, for some $C_2 = \Theta(C_1)$.

Proof: Let $B(X, r)$ be the disk centered at X with radius r. Then, all $B(Y_j, \Delta_{\mathrm{s}} r_j/2), j \in \mathcal{T}^{(\mathrm{s})}$ should be mutually disjoint according to Corollary 1. As well, $B(Y_j, \Delta_{\mathrm{ps}} r_j/2), j \in \mathcal{T}^{(\mathrm{s})}$ are disjoint with $B(X_{\mathrm{Rx}(i)}, \Delta_{\mathrm{sp}} R_i/2)$. Since $\Delta_{\mathrm{ps}} > \Delta_{\mathrm{s}}$, then all $B(Y_j, \Delta_{\mathrm{s}} r_j/2), j \in \mathcal{T}^{(\mathrm{s})}$, $B(X_{\mathrm{Rx}(i)}, \Delta_{\mathrm{sp}} R_i/2)$ are pairwise disjoint. Denote $D_{ij} = B(X_{\mathrm{Rx}(i)}, |X_{\mathrm{Rx}(i)} - Y_j|) \cap B(Y_j, \Delta_{\mathrm{s}} r_j/2)$, it is clear that all D_{ij} are disjoint (see Fig. 3). Denote by E, F the two points where $B(X_{\mathrm{Rx}(i)}, |X_{\mathrm{Rx}(i)} - Y_j|)$ intersects $B(Y_j, \Delta_{\mathrm{s}} r_j/2)$. It is clear that $\angle F Y_j X_{\mathrm{Rx}(i)} = \angle E Y_j X_{\mathrm{Rx}(i)} \geq \pi/3$ because $|X_{\mathrm{Rx}(i)} - Y_j| > \Delta_{\mathrm{s}} r_j/2$. Thus, the area of D_{ij} is at least one third[3] of $B(Y_j, \Delta_{\mathrm{s}} r_j/2)$. Let $I_{\mathrm{sp}}(i)$

[3]More precisely, D_{ij} is the asymptotic lens generated by the intersection of two disks. Let $\sigma = |2X_{\mathrm{Rx}(i)} - Y_j|/\Delta_{\mathrm{s}} r_j$, then $|D_{ij}|/|B(Y_j, \Delta_{\mathrm{s}} r_j/2)| = (\cos^{-1}(1/2\sigma) + \sigma^2 \cos((2\sigma^2 - 1)/2\sigma^2) - (1/2)\sqrt{(2\sigma-1)(2\sigma+1)})/\pi$.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

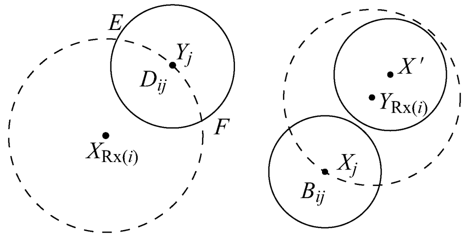


Fig. 3. Analyzing the interference. Left plot shows an example for D_{ij}, and right plot for B_{ij}.

denote the interference at receiver $X_{\mathrm{Rx}(i)}$ from the secondary network, and $\mathrm{d}A$ be the area element

$$\begin{aligned}
I_{\mathrm{sp}}(i) &= \sum_{j\in\mathcal{T}^{(\mathrm{s})}} \frac{P_j}{|Y_j - X_{\mathrm{Rx}(i)}|^{\gamma}} = \sum_{j\in\mathcal{T}^{(\mathrm{s})}} \frac{C_1 r_j^2 P}{|Y_j - X_{\mathrm{Rx}(i)}|^{\gamma}}\\
&= \sum_{j\in\mathcal{T}^{(\mathrm{s})}} \frac{4C_1P}{\pi\Delta_{\mathrm{s}}^2}\int_{B(Y_j,\Delta_{\mathrm{s}} r_j/2)} \frac{\mathrm{d}A}{|Y_j - X_{\mathrm{Rx}(i)}|^{\gamma}}\\
&\quad \left(\text{because} \int_{B(Y_j,\Delta_{\mathrm{s}} r_j/2)} \mathrm{d}A = \pi\Delta_{\mathrm{s}}^2 r_j^2/4\right)\\
&\le \sum_{j\in\mathcal{T}^{(\mathrm{s})}} \frac{12C_1P}{\pi\Delta_{\mathrm{s}}^2}\int_{D_{ij}} \frac{\mathrm{d}A}{|Y_j - X_{\mathrm{Rx}(i)}|^{\gamma}}\\
&\le \sum_{j\in\mathcal{T}^{(\mathrm{s})}} \frac{12C_1P}{\pi\Delta_{\mathrm{s}}^2}\int_{D_{ij}} \frac{\mathrm{d}A}{|X - X_{\mathrm{Rx}(i)}|^{\gamma}}\\
&\quad (\text{where } X \text{ is the position vector of } \mathrm{d}A,\\
&\qquad \text{and} |X - X_{\mathrm{Rx}(i)}| < |Y_j - X_{\mathrm{Rx}(i)}| \text{ for } X \in D_{ij})\\
&= \frac{12C_1P}{\pi\Delta_{\mathrm{s}}^2}\int_{\cup_{j\in\mathcal{T}^{(\mathrm{s})}} D_{ij}} \frac{\mathrm{d}A}{|X - X_{\mathrm{Rx}(i)}|^{\gamma}}.
\end{aligned}$$

Since $(\cup_{j\in\mathcal{T}^{(\mathrm{s})}} D_{ij}) \cap B(X_{\mathrm{Rx}(i)}, \Delta_{\mathrm{sp}} R_i/2) = \emptyset$, we have

$$\begin{aligned}
I_{\mathrm{sp}}(i) &\le \frac{12C_1P}{\pi\Delta_{\mathrm{s}}^2} \int\limits_{|X - X_{\mathrm{Rx}(i)}| \ge \Delta_{\mathrm{sp}} R_i/2} \frac{\mathrm{d}A}{|X - X_{\mathrm{Rx}(i)}|^{\gamma}}\\
&= \frac{12C_1P}{\pi\Delta_{\mathrm{s}}^2} \int\limits_{\Delta_{\mathrm{sp}} R_i/2}^{\infty} \frac{2\pi r \mathrm{d}r}{r^{\gamma}}\\
&= \frac{24C_1P}{\Delta_{\mathrm{s}}^2(\gamma-2)}\left(\frac{2}{\Delta_{\mathrm{sp}} R_i}\right)^{\gamma-2} = C_2 P R_i^{2-\gamma}.
\end{aligned}$$

C. *Interference at Secondary Nodes*

Now we focus on the interference at secondary nodes. The main challenge is to bound the uncontrollable interference from the primary network. Let $R_{\max}(n) = \max R_i(n)$, $R_{\min}(n) = \min R_i(n)$, and $r_{\max}(m) = \max r_j(m)$.

Theorem 2: Under power assignment $\mathcal{A}(C_1)$ and the hybrid protocol model, for any active link $(Y_i, Y_{\mathrm{Rx}(i)})$, the interference at $Y_{\mathrm{Rx}(i)}$ from the primary network is upper-bounded by $c_3 P R_{\min}(n)^{-\gamma}$, for some constant c_3.

Proof: Denote by $I_{\mathrm{ps}}(i) = \sum_{j\in\mathcal{T}^{(\mathrm{p})}} (P/|X_j - Y_{\mathrm{Rx}(i)}|^{\gamma})$ the interference at $Y_{\mathrm{Rx}(i)}$ from the primary network. Pick X' as the interfering primary transmitter closest to $Y_{\mathrm{Rx}(i)}$. From Corollary 1, the distance between any primary transmitter and $Y_{\mathrm{Rx}(i)}$ should be larger than $\Delta_{\mathrm{sp}} R_{\min}/2 + \Delta_{\mathrm{ps}} r_i/2$; the distance between any two primary transmitters is larger than $\Delta_{\mathrm{p}} R_{\min}$.

Now consider the case that $\Delta_{\mathrm{sp}} < \Delta_{\mathrm{p}}$. (Note that if $\Delta_{\mathrm{sp}} > \Delta_{\mathrm{p}}$, $I_{\mathrm{ps}}(i)$ will be smaller, and the upper bound still holds.) Then, all $X_j, j \in \mathcal{T}^{(\mathrm{p})}$ is at least $\Delta_{\mathrm{p}} R_{\min}/2$ away from $Y_{\mathrm{Rx}(i)}$ except X'. First, consider the interference contributed by X'

$$\frac{P}{|X' - Y_{\mathrm{Rx}(i)}|^{\gamma}} \le P\left(\frac{\Delta_{\mathrm{sp}}}{2} R_{\min}\right)^{-\gamma}.$$

Next, consider the interference from some other primary transmitter X_j. Let $B_{ij} = B(X_j, \Delta_{\mathrm{p}} R_{\min}/2) \cap B(Y_{\mathrm{Rx}(i)}, |X_j - Y_{\mathrm{Rx}(i)}|)^c$, as shown in Fig. 3, then

$$\begin{aligned}
\frac{P}{\left|X_j - Y_{\mathrm{Rx}(i)}\right|^{\gamma}} &\le \int\limits_{B_{ij}} \frac{P}{|B_{ij}|}\frac{\mathrm{d}A}{\left|X_j - Y_{\mathrm{Rx}(i)}\right|^{\gamma}}\\
&\quad (\text{where } |B_{ij}| \text{ is the area of disk } B_{ij})\\
&\le \int\limits_{B_{ij}} \frac{P}{|B_{\min}|}\frac{\mathrm{d}A}{\left|X_j - Y_{\mathrm{Rx}(i)}\right|^{\gamma}}\\
&\quad \left(\text{where } |B_{\min}| = \min_j |B_{ij}|\right)\\
&\le \int\limits_{B_{ij}} \frac{P}{|B_{\min}|}\frac{\mathrm{d}A}{\left(\frac{1}{2}\left|X - Y_{\mathrm{Rx}(i)}\right|\right)^{\gamma}}\\
&\quad \left(\text{because } \left|X_j - Y_{\mathrm{Rx}(i)}\right| > |\Delta_{\mathrm{p}} R_{\min}/2|\right)\\
&= \frac{2^{\gamma}P}{|B_{\min}|}\int\limits_{B_{ij}} \frac{\mathrm{d}A}{\left|X - Y_{\mathrm{Rx}(i)}\right|^{\gamma}}\\
&\le \frac{2^{\gamma+3}P}{\pi\Delta_{\mathrm{p}}^2 R_{\min}^2}\int\limits_{B_{ij}} \mathrm{d}A \left|X - Y_{\mathrm{Rx}(i)}\right|^{\gamma}.
\end{aligned}$$

To sum up, let $I'_{\mathrm{ps}}(i) = \sum_{j\in\mathcal{T}^{(\mathrm{p})}\setminus\{X'\}} (P/|X_j - Y_{\mathrm{Rx}(i)}|^{\gamma})$

$$\begin{aligned}
I'_{\mathrm{ps}}(i) &\le \frac{2^{\gamma+3}P}{\pi\Delta_{\mathrm{p}}^2 R_{\min}^2} \sum_{j\in\mathcal{T}^{(\mathrm{p})}\setminus\{X'\}} \int\limits_{B_{ij}} \frac{\mathrm{d}A}{\left|X - Y_{\mathrm{Rx}(i)}\right|^{\gamma}}\\
&= \frac{2^{\gamma+3}P}{\pi\Delta_{\mathrm{p}}^2 R_{\min}^2} \int\limits_{\cup_{j\in\mathcal{T}^{(\mathrm{p})}\setminus\{X'\}} B_{ij}} \frac{\mathrm{d}A}{\left|X - Y_{\mathrm{Rx}(i)}\right|^{\gamma}}\\
&\quad (\text{since } \{B_{ij}\} \text{ are disjoint})\\
&\le \frac{2^{\gamma+3}P}{\pi\Delta_{\mathrm{p}}^2 R_{\min}^2} \int\limits_{|X - Y_{\mathrm{Rx}(i)}| > \frac{\Delta_{\mathrm{p}} R_{\min}}{2}} \frac{\mathrm{d}A}{\left|X - Y_{\mathrm{Rx}(i)}\right|^{\gamma}}\\
&= \frac{2^{\gamma+3}P}{\pi\Delta_{\mathrm{p}}^2 R_{\min}^2} \int\limits_{\frac{\Delta_{\mathrm{p}} R_{\min}}{2}}^{\infty} \frac{2\pi r \mathrm{d}r}{r^{\gamma}}.
\end{aligned}$$

Combining the contribution from X'

$$I_{\mathrm{ps}}(i) \le \left(\frac{2^{\gamma}}{\Delta_{\mathrm{sp}}^{\gamma}} + \frac{2^{2\gamma+2}}{(\gamma-2)\Delta_{\mathrm{p}}^{\gamma}}\right) P R_{\min}^{-\gamma}(n).$$

■

We should also take into account the interference between secondary links. Power assignment scheme $\mathcal{A}$ is well designed so that it not only restricts the interference from the secondary network to the primary network, but also that between secondary links, as shown by the next theorem. Its proof is similar to Theorem 1 and is omitted for space concern.

Theorem 3: Under power assignment scheme $\mathcal{A}(C_1)$ and the hybrid protocol model, for any active secondary link $(Y_i, Y_{\mathrm{Rx}(i)})$, the interference at $Y_{\mathrm{Rx}(i)}$ from all other simultaneously active secondary links is upper-bounded by $I_{\mathrm{ss}}(i) \le C_4 \Pr_i^{2-\gamma}(m)$, where $C_4 = (24 \cdot 2^{\gamma-2}/(\gamma-2)\Delta_{\mathrm{s}}^{\gamma})C_1$.

D. Physical Feasibility of the Hybrid Protocol Model

Lastly, we show under appropriate conditions that the hybrid protocol model is indeed physically feasible. We begin with a lemma.

Lemma 3: Given $A, B, C, a, b > 0$, if $(C/A) \le (b/a(a+b))$ and $(A/B) \ge a + b$, then $(A/B + C) \ge a$.

Then, first consider the primary network.

Lemma 4: If $\mathcal{S}^{(\mathrm{p})} \in \mathscr{D}(\alpha + \epsilon)$, $(X_i, X_{\mathrm{Rx}(i)}) \in \mathcal{S}^{(\mathrm{p})}$, and $I_{\mathrm{sp}}(i) \le PR_i^{-\gamma}(n)(\epsilon/\alpha(\alpha+\epsilon))$, then $(X_i, X_{\mathrm{Rx}(i)})$ is feasible under physical model $\mathscr{P}(\alpha, \beta)$. That is to say, $(PR_i^{-\gamma}(n)/N + I_{\mathrm{sp}}(i) + I_{\mathrm{pp}}(i)) \ge \alpha$, where $I_{\mathrm{pp}}(i)$ is interference at $X_{\mathrm{Rx}(i)}$ from other simultaneously active primary links.

Proof: Let $A = PR_i^{-\gamma}, B = N + I_{\mathrm{pp}}(i), C = I_{\mathrm{sp}}(i), a = \alpha, b = \epsilon$. Note that $\mathcal{S}^{(\mathrm{p})} \in \mathscr{D}(\alpha + \epsilon)$ implies $A/B \ge a + b$, then the result holds from Lemma 3. ■

Lemma 5: If $\Delta_{\mathrm{ps}} > \Delta_{\mathrm{s}}$, $\mathcal{S} \in \mathscr{H}$ and $\mathcal{S}^{(\mathrm{p})} \in \mathscr{D}(\alpha + \epsilon)$, then under power assignment $\mathcal{A}(C_1)$ such that $C_2 \le (\epsilon/\alpha(\alpha + \epsilon)R_{\max}^2(n))$, all primary links are feasible under physical model $\mathscr{P}(\alpha, \beta)$.

Proof: For any $(X_i, X_{\mathrm{Rx}(i)}) \in \mathcal{S}^{(\mathrm{p})}$, from Theorem 1, $I_{\mathrm{sp}}(i) \le C_2 R_i^{2-\gamma} P$, thus the signal-to-inteference ratio (SIR) at $X_{\mathrm{Rx}(i)}$ satisfies

$$\frac{PR_i^{-\gamma}}{I_{\mathrm{sp}}(i)} \ge \frac{PR_i^{-\gamma}}{C_2 PR_i^{2-\gamma}} \ge \frac{1}{C_2 R_{\max}^2} \ge \frac{\alpha(\alpha+\epsilon)}{\epsilon}.$$

Then, from Lemma 4, we have the assertion. ■

Now turn to the secondary network. Similarly, it can be shown that Lemma 6 follows from Theorem 2, and Lemma 7 from Theorem 3,

Lemma 6: Under power assignment $\mathcal{A}(C_1)$ with $C_1 \ge (c_3 c_5 R_{\min}^{-\gamma}(n)/r_i^{2-\gamma}(m))$, if $\mathcal{S}^{(\mathrm{p})} \in \mathscr{H}$, then for any $(Y_i, Y_{\mathrm{Rx}(i)}), i \in \mathcal{T}^{(\mathrm{s})}$, it follows that

$$\frac{C_1 Pr_i^{2-\gamma}(m)}{I_{\mathrm{ps}}(i)} \ge c_5. \tag{4}$$

Lemma 7: Under the condition of Lemma 6, if $\Delta_{\mathrm{s}} \ge (48(2^{\gamma-2}/\gamma-2)c_6)^{1/\gamma}$, for any $(Y_i, Y_{\mathrm{Rx}(i)}), i \in \mathcal{T}^{(\mathrm{s})}$, it follows that

$$\frac{C_1 Pr_i^{2-\gamma}(m)}{N + I_{\mathrm{ss}}(i)} \ge c_6.$$

Then, we are ready to prove the final result.

Theorem 4: If $r_{\max}^{\gamma-2}(m) = o(R_{\min}^{\gamma}(n)/R_{\max}^2(n))$, $\Delta_{\mathrm{p}} \le (\alpha+\epsilon)^{1/\gamma} - 1$, and $\Delta_{\mathrm{ps}} \ge \Delta_{\mathrm{s}} \ge (24(2^{\gamma}/\gamma-2)\beta)^{1/\gamma}$, then there exists power assignment $\mathcal{A}(C_1)$, such that for any $\mathcal{S}^{(\mathrm{p})} \in \mathscr{D}(\alpha+\epsilon)$, it follows that $\mathcal{S}^{(\mathrm{p})} \in \mathscr{H}(\Delta_{\mathrm{p}}, \Delta_{\mathrm{ps}}, \Delta_{\mathrm{sp}}, \Delta_{\mathrm{s}})$. If we schedule secondary network transmissions in the way such that $\mathcal{S}^{(\mathrm{p})} \cup \mathcal{S}^{(\mathrm{s})} \in \mathscr{H}$, then $\mathcal{S}^{(\mathrm{p})} \cup \mathcal{S}^{(\mathrm{s})} \in \mathscr{P}(\alpha, \beta)$ holds.

Proof: The first claim follows from Lemma 1. To prove the second claim, first notice every active primary link is physically feasible if the condition of Lemma 5 is verified, i.e.,

$$C_2 = \Theta(C_1) = o\left(1/R_{\max}^2(n)\right). \tag{5}$$

On the other hand, consider the secondary network, if

$$C_1 = \omega\left(R_{\min}^{-\gamma}(n)/r_i^{2-\gamma}(m)\right) \tag{6}$$

then according to Lemma 6, (4) holds for any positive constant c_5. In combination with Lemma 7, it is clear that SINR at any secondary receiver is greater than $(1/2)(\min\{c_5, c_6\}) = \beta$. Since $r_{\max}^{\gamma-2}(m) = o((R_{\min}^{\gamma}(n)/R_{\max}^2(n)))$, we can indeed find C_1 and C_2, such that (5) and (6) hold, proving the theorem. ■

The condition $r_{\max}^{\gamma-2}(m) = o(R_{\min}^{\gamma}(n)/R_{\max}^2(n))$ characterizes the regime that primary links are homogeneous in range. In other words, if this condition fails, it implies that the scheduling of the primary network is excessively heterogeneous in transmission ranges and a simple decision model like $\mathscr{H}$ does not suffice to identify transmission opportunities. Fortunately, this condition usually holds because $R_{\max}(n)$ and $R_{\min}(n)$ typically do not differ much in order and we usually tend to employ a small $r(m)$.

V. Availability of Transmission Opportunities

Section IV addresses the problem of how to identify transmission chances for secondary networks: Given a set $\mathcal{S}^{(\mathrm{p})}$ of simultaneously active primary links, we allow a set $\mathcal{S}^{(\mathrm{s})}$ of simultaneously active secondary links according to hybrid protocol model $\mathscr{H}$. This section, on the other hand, considers the problem that for those secondary links that desire to transmit, how frequently do these opportunities occur. Of particular interest is to compare this result to an identical standalone network. Standalone networks provide trivial performance upper bounds since cognitive secondary networks will suffer from additional transmission constraints imposed by primary networks. To alleviate the performance loss due to such constraints, it is intuitive that one should reduce the range of secondary links, and this fact is indeed verified by the hybrid protocol model and Theorem 4. This section will further show that if $r_{\max}(m) = o(R_{\min}(n))$, then for quite general cases, such performance loss is insignificant and has no impact in order sense. In other words, all secondary links have a constant ratio of time to be *unconstrained* as if they were in a standalone network, as long as they employ a transmission range that is small enough. Moreover, note it is well known that to achieve better scalability, a smaller range is also favorable. This coincidence implies that secondary networks can achieve the optimal scaling performance of a standalone network. Also notice that by stating the secondary users as unconstrained, we mean that the secondary links activated by the scheduler can transmit without constraint subject to the dynamics of primary network activities. However, the secondary scheduler itself should comply with the constraint of $r_{\max}(m) = o(R_{\min}(n))$. We remark that the other regime of $r_{\max}(m) = \omega(R_{\min}(n))$, which is not covered in this paper, is also a very interesting and

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

important scenario to study and still needs further explorations in future works.

Now, we formally introduce the concept of unconstraint and analyze the unconstrained time period in the following. Let s.a. be the abbreviation of standalone:

Definition 5: Given arbitrary $\mathcal{S}^{(\mathrm{s})}_{\mathrm{s.a.}} \in \mathscr{Q}_{\mathrm{s}}(\Delta_{\mathrm{s}})$ and arbitrary $\mathcal{S}^{(\mathrm{p})} \in \mathscr{Q}_{\mathrm{p}}(\Delta_{\mathrm{p}})$, there exists a unique maximal $\mathcal{S}^{(\mathrm{s})} \subset \mathcal{S}^{(\mathrm{s})}_{\mathrm{s.a.}}$ such that $\mathcal{S}^{(\mathrm{p})} \cup \mathcal{S}^{(\mathrm{s})} \in \mathscr{H}(\Delta_{\mathrm{p}}, \Delta_{\mathrm{ps}}, \Delta_{\mathrm{sp}}, \Delta_{\mathrm{p}})$. We say a link $(Y_i, Y_{\mathrm{Rx}(i)}) \in \mathcal{S}^{(\mathrm{s})}_{\mathrm{s.a.}}$ is *unconstrained* if $(Y_i, Y_{\mathrm{Rx}(i)}) \in \mathcal{S}^{(\mathrm{s})}$.

Note the fraction of time that the link is constrained characterizes the performance loss relative to the corresponding standalone network.

A. Cell-Partitioning Round-Robin Mode

We first consider the case that primary networks operate according to a common scheduling paradigm: the cell-partitioning round-robin active scheme. It first spatially tessellates the network into cells, then assigns color to each cell such that cells with the same color, if limiting their transmissions to neighbors, will not interfere with each other. Then, we allow cells with the same color to transmit simultaneously, and let different colors take turns to be active. A simple TDMA scheme will suffice. This very widely employed scheme [4], [7], [13], [14] features a high degree of spatial concurrency and thus frequency reuse. It is deterministic and therefore simple. To the best of our knowledge, all previous works on asymptotic analysis of cognitive networks focused on some particular variants of such a TDMA scheme.

Definition 6: A *network tessellation* is a set of disjoint cells $\{V_i \subset \mathcal{O}\}$. A *round-robin TDMA scheme* is a scheduling scheme that: 1) tessellates the network into cells such that every cell is contained in a disk of radius $\rho(n)$; 2) allows noninterfering cells to be simultaneously active and transmit to neighbor cells, where two cells V_i, V_j are noninterfering if $\sup\{|E-F| : E \in V_i, F \in V_j\} \geq (2+\Delta_{\mathrm{p}})4\rho(n)$; and 3) activates different groups of cells in a round-robin TDMA fashion, and guarantees every cell can be active for at least c_7 fraction of time in one round, for some constant $c_7 > 0$.

The existence of round-robin TDMA schemes is a consequence of the well-known theorem about vertex coloring of graphs. Such a scheme is favorable to secondary networks because it deterministically ensures transmission opportunities not only for every primary cell, but also for every secondary link. We now show for a generic primary scheduling policy that operates in the round-robin fashion that the unconstrained time fraction for any short range secondary link is constant, as a simple consequence of the hybrid protocol model.

Theorem 5: If the primary network operates according to a round-robin TDMA scheme and $\Delta_{\mathrm{p}} > 2$, $\Delta_{\mathrm{sp}} \leq (\Delta_{\mathrm{p}} - 2/2)$, then every secondary link with range $r(m) = o(R_{\min}(n))$ has at least c_7 fraction of time to be unconstrained in one round.

Proof: Consider a generic link $(Y_i, Y_{\mathrm{Rx}(i)})$, pick a point X such that $|X - Y_i| = (4+2\Delta_{\mathrm{p}})\rho(n)$, and denote by V the cell X belongs to. We claim whenever V is scheduled to be active, $(Y_i, Y_{\mathrm{Rx}(i)})$ is unconstrained. To that end, we first verify transmitter Y_i will not upset transmissions in V. Indeed, any point E that belongs to V should lie within distance $2\rho(n)$ from X, thus any point F that belongs to a neighbor cell of V should lie within distance $4\rho(n)$ from X, then

$$\begin{aligned}|Y_i - F| &\geq (4+2\Delta_{\mathrm{p}})\rho(n) - 4\rho(n)\\ &= 2\Delta_{\mathrm{p}}\rho(n) \geq (1+\Delta_{\mathrm{sp}})4\rho(n)\\ &\geq (1+\Delta_{\mathrm{sp}})|E-F|.\end{aligned} \tag{7}$$

Now consider another simultaneous active cell V'. It is clear that any point $X' \in V'$ is at least $(2+\Delta_{\mathrm{p}})4\rho(n)$ away from X, then $|X' - Y_i| \geq (4+2\Delta_{\mathrm{p}})\rho(n) = |X - Y_i|$. Together with (7), condition 2 is verified. Moreover, since $r = o(R_{\min})$, condition 3 is obvious. This completes the proof. ■

We observe that under hybrid protocol model $\mathscr{H}$, $\Delta_{\mathrm{p}} > 2$ is critical to guarantee transmission opportunities for secondary nodes, as shown in Theorem 5. Equivalently, it implies $\alpha + \epsilon \geq 2^{\gamma}$. This is an assumption about primary networks, and we assume it always holds from now on. However, we conjecture this assumption is not fundamental and can be relaxed by introducing a criterion with more flexible form, i.e., allowing Δ_{sp} and Δ_{ps} to be dependent on n. Such decision models may have a better capability of digging into the potential of available gaps, at the cost of complexity. We leave for future work a more in-depth analysis of such cases.

B. Independent Relay Mode

Theorem 5 suffices to provide rich performance scaling results on cognitive networks, for the scenario it considers, i.e., the round-robin TDMA scheme, covers most centralized control networks. However, some other cases are also of interest such as networks which employ distributed CSMA protocol [18]. Exceptions also exist in centralized control networks, such as the protocol proposed in [19], which schedules the network in a more generic and ideal manner without relying on the concept of cells. For that matter, Theorem 5 relies on the scheduling of primary networks, but sometimes it is desired to relax this requirement. In the following, we show that some general assumptions on the routing protocol of primary networks will sufficiently lead to similar conclusions.

Intuitively, according to the hybrid protocol model, on one hand primary transmissions will not be too dense spatially, thus leaving gaps for the secondary network. On the other hand, they also mute nearby secondary links. We shall show every primary link can create some gap and mute some area. More formally, given link $(X_i, X_{\mathrm{Rx}(i)})$ and $(Y_j, Y_{\mathrm{Rx}(j)})$, we say the former *triggers* the latter if $(Y_j, Y_{\mathrm{Rx}(j)})$ is unconstrained as long as $(X_i, X_{\mathrm{Rx}(i)})$ is active and *shades* the latter conversely. Because nodes are i.i.d., whether a primary link will trigger or shade a secondary link is a random event. Then, if traffic is somewhat "independently" distributed (relayed) to primary links, according to the law of large numbers, if a secondary link is shaded for a long time *w.h.p.*, i.e., the primary traffic nearby is intense, it will also be triggered for considerable time.

Lemma 8: Consider link $(X_i, X_{\mathrm{Rx}(i)})$ and $(Y_j, Y_{\mathrm{Rx}(j)})$. If $\Delta_{\mathrm{p}} > 2$, $\Delta_{\mathrm{sp}} \leq (\Delta_{\mathrm{p}} - 2/2)$ and $r_j(m) = o(R_i(n))$, then a sufficient condition that $(X_i, X_{\mathrm{Rx}(i)})$ triggers $(Y_j, Y_{\mathrm{Rx}(j)})$ is that Y_j lies in the ring of points with distance to line segment $X_i X_{\mathrm{Rx}(i)}$ larger than $(1+\Delta_{\mathrm{sp}})R_i(n)$ and less than

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

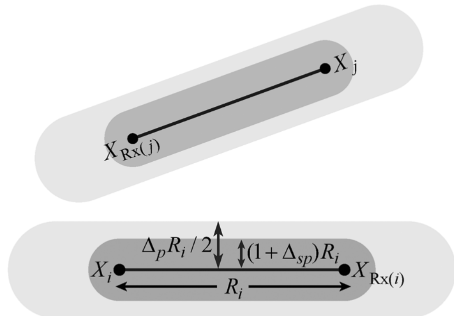


Fig. 4. Active primary link $(X_i, X_{\mathrm{Rx}(i)})$ can shade some area (the dark region) and trigger some area (the outside ring). Note the shading and triggering region of two links are disjoint according to Corollary 1.

$\Delta_{\mathrm{p}} R_i(n)/2$. A necessary condition that $(X_i, X_{\mathrm{Rx}(i)})$ shades $(Y_j, Y_{\mathrm{Rx}(j)})$ is that Y_j lies within the $(1+\Delta_{\mathrm{sp}})R_i(n)$ neighborhood of line segment $X_i X_{\mathrm{Rx}(i)}$, as shown in Fig. 4.

Proof: The necessary condition is obvious. As to the sufficient condition, (3) holds for $r_j = o(R_i)$. It is also clear that $|Y_j - X_{\mathrm{Rx}(i)}| \geq (1+\Delta_{\mathrm{sp}})R_i$. For any other active primary link $(X_k, X_{\mathrm{Rx}(k)}), k \neq i$, its receiver is at least $\Delta_{\mathrm{p}} R_k/2$ away from Y_j due to Corollary 1. Therefore, (2) also holds, proving the lemma. ■

As a consequence, we can term $(X_i, X_{\mathrm{Rx}(i)})$ triggers $(Y_j, Y_{\mathrm{Rx}(j)})$ and $(X_i, X_{\mathrm{Rx}(i)})$ triggers Y_j interchangeably.

Definition 7: Consider a regular network tessellation of square cells. We assume every source route traffic to its destination along these cells in multihop fashion, such that at every hop, a packet is transmitted to a relay node in a neighbor cell. We say the network routing operates in the *independent relay* mode if, at each hop, flows choose relays randomly and independently among all nodes in the receiving cells.

The regular tessellation of square cells is only a technical assumption for the ease of presentation. A similar result also holds for other topologies. By saying "independent," we do not mean the routes of two flows are independent. In fact, they could be highly related, such as choosing a same sequence of cells to forward. Instead, we only require that two flows independently choose relays for a certain cell. Intuitively, independently relaying implies there are no special designated nodes in the network, and is in accord with the design principles of distributed systems such as ad hoc networks. It is notable that the class of independent relay protocol is quite general and common [18], [19].

The next lemma follows from the well-known connectivity criterion [20], and Lemma 10 is a standard application of the Chernoff bound.

Lemma 9: For an independent relay protocol, to ensure asymptotic connectivity of the overall network, the side length L of square cells is at least $\Theta(\sqrt{\log n/n})$.

Lemma 10: For an independent relay protocol, there exist positive constants c_8 and c_9, such that *w.h.p.* every cell contains more than $c_8 n L^2$ and less than $c_9 n L^2$ primary nodes.

Now we characterize the shading/triggering events with a Bernoulli-like probability model:

Lemma 11: Consider arbitrary neighboring cells V_1, V_2 and link $(Y_j, Y_{\mathrm{Rx}(j)})$, and let X_i and $X_{\mathrm{Rx}(i)}$ be independently and uniformly distributed in V_1 and V_2, respectively. Denote by p the probability that $(X_i, X_{\mathrm{Rx}(i)})$ triggers $(Y_j, Y_{\mathrm{Rx}(j)})$, and q the probability of shading. Then, $\forall$ constant $\delta_1, \delta_2 > 0$, among all primary links from V_1 to V_2, *w.h.p.*, there are at least $p(1-\delta_1)(c_8 n L^2)^2$ links that trigger $(Y_j, Y_{\mathrm{Rx}(j)})$, and at most $q(1+\delta_2)(c_9 n L^2)^2$ links that shade it.

Proof: We only prove the first part of the lemma. From Lemma 10, there are at least $c_8 n L^2$ nodes in each cell, denoted by X_{u_k} and $X_{v_l}, k, l = 1, \ldots, c_8 n L^2$. Thus, we have at least $(c_8 n L^2)^2$ candidate links from V_1 to V_2. Consider this subset of links, define $I_{k,l}$ as

$$I_{k,l} = \begin{cases} 1, & \text{if } (X_{u_k}, X_{v_l}) \text{ triggers } (Y_j, Y_{\mathrm{Rx}}(j)) \\ 0, & \text{otherwise.} \end{cases}$$

Because nodes are i.i.d., $\{I_{k,l}\}$ are identically distributed and with probability p equal to 1. Moreover, $I_{a,b}$ and $I_{k,l}$ are independent if $a \neq k$ and $b \neq l$. Construct I_k as

$$I_k = \begin{cases} \sum_{i=1}^{c_8 n L^2 - \frac{k-1}{2}} I_{i+\frac{k-1}{2}, i}, & k \text{ odd} \\ \sum_{i=1}^{c_8 n L^2 - \frac{k}{2}} I_{i, i+\frac{k}{2}}, & k \text{ even.} \end{cases}$$

Then, I_k is the sum of *i.i.d.* random variables. Applying the law of large numbers, one can easily show that $\forall \delta_3 < 2$, $k < \delta_3 c_8 n L^2$, $\forall \delta_1 > 0$, the following holds *w.h.p.*:

$$I_k > \begin{cases} p(1-\delta_1)\left(c_8 n L^2 - \frac{k-1}{2}\right), & k \text{ odd} \\ p(1-\delta_1)\left(c_8 n L^2 - \frac{k}{2}\right), & k \text{ even.} \end{cases}$$

Lastly, note the relationship of summing $\{I_{k,l}\}$ and $\{I_k\}$

$$\sum_k \sum_l I_{k,l} = \sum_{k=1}^{2c_8 n L^2 - 1} I_k \geq \sum_{k=1}^{\delta_3 c_8 n L^2} I_k$$
$$\geq \left(p(1-\delta_1) - (2-\delta_3)\right)\left(c_8 n L^2\right)^2 \quad \text{(w.h.p.)}.$$

Making δ_3 arbitrarily close to 2, we have the claim. ■

In the next step, we characterize the relation between p and q. The main idea is to couple the triggering and shading events through a continuous transformation in $\mathbb{R}^4$. We first cite a property of Lebesgue measure [21].

Lemma 12: (*Integration by change of variable*) Let $S \subset \mathbb{R}^n$ be an open set, and let $\mathcal{L}$ be a Lebesgue measure on S. Let $T(x) = (y_1(x), \ldots, yn(x)), x = (x_1, \ldots, x_n) \in S$ be a given homeomorphism $T: S \to \mathbb{R}^n$ with the continuous derivatives $(\partial y_i / \partial x_j), i, j = 1, \ldots, n$ on S, and we note with $\tau(T, x) = ((\partial y_i / \partial x_j))$ the nondegenerate Jacobian matrix for all $x \in S$. Then, for any nonnegative Borel function f defined on the open set $T(S)$, we have

$$\int_{T(S)} f(y)\mathrm{d}y = \int_S f(Tx) \cdot |\tau(T, x)|\, \mathrm{d}x$$

where $\mathrm{d}x, \mathrm{d}y$ denote the integration with respect to $\mathcal{L}$.

Theorem 6: Define p, q as in Lemma 11, and under the condition of Lemma 8, there exists constant $c_{10} > 0$, such that $p > c_{10} q$.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

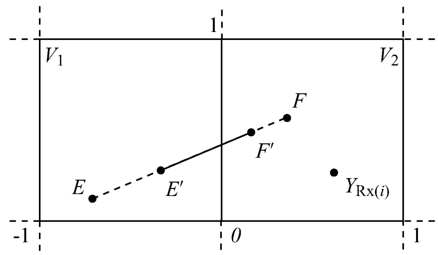


Fig. 5. Transformation T_θ shrinks EF to $E'F'$ such that $|E'F'| = \theta|EF|$ ($\theta = 1/2$ in this figure) and ensures that $\omega' \in \Omega$.

Proof: Without loss of generality, let $V_1 = [-1, 0] \times [0, 1], V_2 = [0, 1] \times [0, 1]$, and $(\Omega, \mathscr{F}, \mathcal{L})$ be the probability space[4] of interest, where $\Omega = V_1 \times V_2$ and $\mathcal{L}$ is the Lebesgue measure restricted on Ω. Given $\omega = (x_1, y_1, x_2, y_2) \in \Omega$, define $T_\theta : \Omega \to \Omega$ as $T_\theta(\omega) = \omega' = (x_1', y_1', x_2', y_2')$

$$\begin{cases} x_1' = -\theta x_1 \\ y_1' = y_1 + (1-\theta)\frac{x_1(y_2-y_1)}{x_1-x_2} \\ x_2' = \theta x_2 \\ y_2' = \theta y_2 + (1-\theta)\left(y_1 + \frac{x_1(y_2-y_1)}{x_1-x_2}\right). \end{cases}$$

Intuitively, let $E = (x_1, y_1)$ and $F = (x_2, y_2)$, then T_θ linearly shrinks line segment EF to $E'F'$ with $\theta < 1$, preserving its geometric topology. See Fig. 5 for an example.

Let $d(\omega)$ be the distance from Y_j to line segment EF and $S_{\mathrm{shd}} = \{\omega \in \Omega : Y_j \text{ is shaded}\}$, $S_{\mathrm{trg}} = \{\omega \in \Omega : Y_j \text{ is triggered}\}$. Assume S_{shd} is not empty and consider any $\omega_0 \in S_{\mathrm{shd}}$ and the set $H(\omega_0) = \{\omega' \in \Omega : \omega' = T_\theta(\omega_0), \theta > 0\}$. Define θ_{cr} and $\theta_{\min}$ such that

$$\frac{d\left(T_{\theta_{\mathrm{cr}}}(\omega_0)\right)}{1+\Delta_{\mathrm{sp}}} = \theta_{\mathrm{cr}}|EF| = |EF|_{\mathrm{cr}}$$
$$\frac{2d\left(T_{\theta_{\min}}(\omega_0)\right)}{\Delta_{\mathrm{p}}} = \theta_{\min}|EF| = |EF|_{\min}.$$

According to Lemma 8, it is clear that θ_{cr} and $\theta_{\min}$ uniquely exist and $\theta_{\mathrm{cr}} > \theta_{\min}$. Moreover, $\forall \omega' \in H(\omega_0)$, $\omega' \in S_{\mathrm{shd}} \Rightarrow |E'F'| > |EF|_{\mathrm{cr}}$; $|EF|_{\min} \le |E'F'| \le |EF|_{\mathrm{cr}} \Rightarrow \omega' \in S_{\mathrm{trg}}$. Therefore, we can introduce a mapping from the set of line segments with length $(|EF|_{\mathrm{cr}}, \sqrt{5})$ to those with length $(|EF|_{\min}, |EF|_{\mathrm{cr}})$, where $\sqrt{5}$ is an upper bound of $|E'F'|$, i.e.,

$$\theta(\omega_0) = \frac{|EF|_{\min} + \frac{|EF| - |EF|_{\mathrm{cr}}}{\sqrt{5} - |EF|_{\mathrm{cr}}}\left(|EF|_{\mathrm{cr}} - |EF|_{\min}\right)}{|EF|}.$$

Then, $\forall \omega \in S_{\mathrm{shd}}, T_{\theta(\omega)}(\omega) \in S_{\mathrm{trg}}$, i.e., we construct a mapping between S_{shd} and S_{trg}, and by invoking Lemma 12, the relation between p and q can be established. To that end, we first simply $\theta(\omega)$. It is obvious on a little thought that $\forall S \subset S_{\mathrm{shd}}$, if $\tilde{\theta} \le \min_{\omega \in S}(|EF|_{\mathrm{cr}} - |EF|_{\min}/\sqrt{5} - |EF|_{\mathrm{cr}})$, then $\mathcal{L}(T_{\theta(\omega)}(S)) \ge \mathcal{L}(T_{\tilde{\theta}}(S))$. Additionally

$$\begin{aligned} \frac{|EF|_{\mathrm{cr}} - |EF|_{\min}}{\sqrt{5} - |EF|_{\mathrm{cr}}} &\ge \frac{\frac{d(T_{\theta_{\mathrm{cr}}}(\omega_0))}{1+\Delta_{\mathrm{sp}}} - \frac{2d\left(T_{\theta_{\min}}(\omega_0)\right)}{\Delta_{\mathrm{p}}}}{\sqrt{5}} \\ &\ge \frac{d_{\mathrm{cr}}}{\sqrt{5}(1+\Delta_{\mathrm{sp}})} - \frac{2d_{\mathrm{cr}}\left(|EF|_{\mathrm{cr}} - |EF_{\min}|\right)}{\sqrt{5}\Delta_{\mathrm{p}}} \\ &\ge \frac{\Delta_{\mathrm{p}} + \Delta_{\mathrm{sp}} - 2}{2\sqrt{5}\Delta_{\mathrm{p}}(1+\Delta_{\mathrm{sp}})} d_{\mathrm{cr}}. \end{aligned}$$

Lastly, observe in the worst case that Y_j lies in V_1 or V_2, and there are some $S_{\mathrm{shd}}' \subset S_{\mathrm{shd}}, \mathcal{L}(S_{\mathrm{shd}}') > 1/8\mathcal{L}(S_{\mathrm{shd}})$, such that $S_{\mathrm{shd}}' \supset \cup_{\omega \in S_{\mathrm{shd}}'} H(\omega)$ and $\forall \omega \in S_{\mathrm{shd}}', d(\omega) > 1/4$. Therefore, $d_{\mathrm{cr}} > 1/4$. Define $\tilde{\theta} \triangleq (\Delta_{\mathrm{p}} + \Delta_{\mathrm{sp}} - 2/8\sqrt{5}\Delta_{\mathrm{p}}(1+\Delta_{\mathrm{sp}}))$, and it follows that

$$\begin{aligned} q = \Pr\{Y_j \text{ is shaded}\} &= \mathcal{L}(S_{\mathrm{shd}}) \le 8\mathcal{L}\left(S_{\mathrm{shd}}'\right) \\ &\le 8\mathcal{L}\left(T_{\tilde{\theta}}\left(S_{\mathrm{shd}}'\right)\right)/\tilde{\theta}^3 \\ &\quad (\text{let } S = S_{\mathrm{shd}}' \text{ and } f \equiv 1 \text{ in Lemma 12} \\ &\quad\quad \text{and because } |\tau(T_\theta, x)| \equiv \theta^3) \\ &\le 8\mathcal{L}\left(T_{\theta(\omega)}\left(S_{\mathrm{shd}}'\right)\right)/\tilde{\theta}^3 \\ &\le \frac{8}{\tilde{\theta}^3}\mathcal{L}(S_{\mathrm{trg}}) = \frac{8}{\tilde{\theta}^3}\Pr\{Y_j \text{ is triggered}\} = \frac{8}{\tilde{\theta}^3}p. \end{aligned}$$

Let $c_{10} = \tilde{\theta}^3/8$, and we complete the proof. ■

Theorem 7: If the primary network employs an independent relay protocol and $\Delta_{\mathrm{p}} > 2$, $\Delta_{\mathrm{sp}} \le (\Delta_{\mathrm{p}} - 2/2)$, then every secondary link with range $r(m) = o(R_{\min}(n))$ has at least on average c_{11} fraction of time to be unconstrained, where constant $c_{11} > 0$.

Proof: Intuitively, due to the fact that the triggering (success) probability p is at least of the same order as the shading (failure) probability q, as shown in Theorem 6, then if a secondary link is shaded for a significant fraction of time, this indicates that primary transmissions nearby (Bernoulli trials) are intense and the link will also be triggered for a substantial fraction of time with high probability. The formal proof is presented below.

Without loss of generality, consider a time interval of unit length and a particular secondary link $(Y_j, Y_{\mathrm{Rx}(j)})$. We only discuss the case that Y_j is shaded by transmissions from some cell V_1 to V_2 for a least some constant fraction of time, otherwise the proof is trivial. This implies that the shading probability q is lower-bounded by $q_1 = \Theta(1)$, and $C_a\lambda_{\mathrm{p}} = \Theta(1)$, where C_a is the number of flows that choose this route, and $\lambda_{\mathrm{p}} = O(1/\sqrt{n})$ is the per-node throughput of primary network. Then, from Theorem 6 we have the triggering probability $p > c_{10}q_1 = \Theta(1)$.

According to Lemmas 10 and 11, the fraction of candidate links that trigger Y_j is at least $p_1 = pc_8/2c_9$.

Let $\mathbb{I}$ be the logical indicator function, and define $J = \sum_{i=1}^{C_a} \mathbb{I}\{\text{flow } i \text{ chooses a link that triggers } Y_j\}$, then J is sum of i.i.d. Bernoullian random variables with mean $p_2 > p_1$. Denote $\mathbb{E}$ as expectation, and by applying Chernoff bounds, we get

$$\Pr\left\{J < \frac{1}{2}\mathbb{E}[J] = \frac{1}{2}C_a p_2\right\} < e^{-C_a p_2/8}. \tag{8}$$

[4]With abuse of notation, we use Ω to denote a set and ω an element instead of order when no confusion is caused.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Equation (8) indicates Y_j will be triggered for a constant fraction of time. We need to show this fact holds uniformly for all secondary links. To that end, we tessellate the network into $C_b n$ subsquares for some $C_b = \omega(1)$, then it is clear that all secondary transmitters within a same subsquare share the same status of being unconstrained or not. Denote $J_k = \sum_{i=1}^{C_a} \mathbb{I}_{\{\text{flow } i \text{ triggers subsquare } k\}}$, then by the subadditivity of probability measure

$$\begin{aligned}\Pr\left\{\cap_k J_k > \frac{1}{2}\mathbb{E}[J_k]\right\} &\geq 1 - \sum_k \Pr\left\{J_k < \frac{1}{2}\mathbb{E}[J_k]\right\} \\ &\geq 1 - C_b n e^{-C_a p_2/8} \to 1\end{aligned}$$

where the last limit holds for any $C_b = n^{\mu}, \mu \in \mathbb{R}$ due to $C_a p_2 = \Omega(1/\lambda_{\rm p}) = \Omega(\sqrt{n})$. Then *w.h.p.*, every secondary link is triggered for at least $(1/2)C_a p_2 \lambda_{\rm p} = \Theta(1)$ seconds, proving the theorem. ∎

VI. Optimal Performance Scaling

In this section, we present results on throughput and delay scaling of general cognitive networks, as well as a number of corollaries under various specific settings.

Theorem 8: If the primary network operates in the round-robin TDMA or the independent relay fashion, for any protocol interference model based scheme that schedules and routes the secondary network such that $r_{\max}(m) = o(R_{\min}(n))$, $r_{\max}^{\gamma-2}(m) = o(R_{\min}^{\gamma}(n)/R_{\max}^2(n))$ *w.h.p.*, and achieves per-node throughput $\lambda_{\rm s}$ and delay $D_{\rm s}$ in the case that secondary network is standalone, there exists a corresponding scheme that can achieve per-node throughput $\Theta(\lambda_{\rm s})$ and delay $\Theta(D_{\rm s})$ when the primary network is present and Operation Rules 1 and 2 apply.

Proof: First, we hypothesize the secondary network is standalone, and denote by $\mathfrak{c}_{\rm s.a.}^{ij}$ the throughput rate of link (Y_i, Y_j), then $\{\mathfrak{c}_{\rm s.a.}^{ij}\}$ is determined by the scheduling scheme. For example, if we assume slotted time, then a deterministic scheduling scheme is characterized by a sequence $(\mathcal{S}_{\rm s.a.}^t)_{t=1}^T$, $\mathcal{S}_{\rm s.a.}^t \in \mathscr{Q}_{\rm s}$, and

$$\mathfrak{c}_{\rm s.a.}^{ij} = \frac{W}{T}\sum_{t=1}^{T} \mathbb{I}_{(Y_i,Y_j)\in\mathcal{S}_{\rm s.a.}^t}.$$

The network can be mapped to a graph $\mathcal{G}$, where m vertices stand for secondary nodes, and $\{\mathfrak{c}_{\rm s.a.}^{ij}\}$ compose edges. The network traffic is represented as a multicommodity flow instance on $\mathcal{G}$ [22], and the routing scheme is defined by $\{f_{ij}^{\rm sd}\}$, the average fraction of traffic from $Y_{\rm s}$ to $Y_{\rm d}$ that is routed through link (Y_i, Y_j). Because the overall scheme achieves per-node throughput $\lambda_{\rm s}$, it holds that

$$\lambda_{\rm s}\sum_{\rm s}\sum_{\rm d}\lambda_{\rm sd} f_{ij}^{\rm sd} \leq \mathfrak{c}_{\rm s.a.}^{ij} \qquad 1 \leq i, j \leq m.$$

Now, let the primary network joins, and we stick the secondary network to the prior scheme except for only allowing the unconstrained links to be active. According to Theorem 4, such scheduling is physically feasible. Denote the corresponding throughput rate of link (Y_i, Y_j) as $\mathfrak{c}_{\rm c.r.}^{ij}$, and from Theorems 5 and 7, $\mathfrak{c}_{\rm c.r.}^{ij} \geq c_{12}\mathfrak{c}_{\rm s.a.}^{ij}$, where constant $c_{12} = \min\{c_7, c_{11}\}$. Letting $\lambda_{\rm s}' = c_{12}\lambda_{\rm s} = \Theta(\lambda_{\rm s})$, it follows that

$$\lambda_{\rm s}'\sum_{\rm s}\sum_{\rm d}\lambda_{\rm sd} f_{ij}^{\rm sd} \leq \mathfrak{c}_{\rm c.r.}^{ij} \qquad 1 \leq i, j \leq m.$$

Therefore, no edge is overloaded, and throughput $\lambda_{\rm s}'$ is feasible.

As to delay, the definition and calculation of it depend on specific network settings, such as packet size or mobility patterns [13], [14]. However, we note that a general baseline in prior works is that per-hop delay for a packet is at least $\Omega(1)$ (this may include transmission delay, queueing delay and delay incurred by mobility, etc.). In secondary networks, packets will suffer from extra delay because at each hop, or at each time they are transmitted by a link, they must wait until the link is unconstrained. According to Theorems 5 and 7, such delay penalty is upper-bounded by the duration of one round or unit time, which is $\Theta(1)$. Therefore, the order of overall delay is preserved. ∎

With Theorem 8, we can conveniently extend optimal schemes and results of standalone networks to cognitive networks. The optimality is preserved in cognitive networks unless we allow cooperation between primary and secondary nodes, which is beyond the scope of this paper. The following two corollaries are straightforward from [13] and [14]. For clarity, we assume by convention that both networks are static and operate at the same timescale unless further specifications are made.

Corollary 2: The optimal throughput–delay tradeoff is $D_{\rm p} = \Theta(n\lambda_{\rm p})$, $\lambda_{\rm p} \leq \Theta(1/\sqrt{n})$ for the primary network and $D_{\rm s} = \Theta(m\lambda_{\rm s})$, $\Theta(n\lambda_{\rm p}/m) < \lambda_{\rm s} \leq \Theta(1/\sqrt{m})$ for the secondary network, if $\Theta(1/\sqrt{m}) > \Theta(n\lambda_{\rm p}/m)$.

Remark 1: In the primary networks (classic static wireless ad hoc networks), optimal throughput–delay tradeoff can be achieved by [13, Scheme 1], which is a cell-partitioned TDMA scheme with a squarelet of side length $l_{\rm p}(n)$. Then, $D_{\rm p}(n) = \Theta(1/l_{\rm p}(n))$ and $\lambda_{\rm p}(n) = \Theta(1/nl_{\rm p}(n))$. Notice that primary transmission range *w.h.p.* is $l_{\rm p}(n)$. Let the secondary networks also follow [13, Scheme 1], with $l_{\rm s}(n) = o(l_{\rm p}(n))$, and therefore $\lambda_{\rm s}(n) = \Theta(1/ml_{\rm s}(n)) = \Omega(n\lambda_{\rm p}(n)/m)$. On the other hand, $\Theta(1/\sqrt{n})$ and $\Theta(1/\sqrt{m})$ is the maximal achievable throughput in primary and secondary networks, respectively.

Corollary 3: If primary nodes move according to the random walk model, then the optimal throughput–delay tradeoff for the primary network is $D_{\rm p} = \Theta(n\lambda_{\rm p})$ if $\lambda_{\rm p} \leq \Theta(1/\sqrt{n})$, and $D_{\rm p} = \Theta(n\log n)$ if $\Theta(1/\sqrt{n}) < \lambda_{\rm p} \leq \Theta(1)$. The optimal throughput–delay tradeoff for the secondary network is $D_{\rm s} = \Theta(m\lambda_{\rm s})$, $\Theta(n\min(1/\sqrt{n}, \lambda_{\rm p})/m) < \lambda_{\rm s} \leq \Theta(1/\sqrt{m})$, if $\Theta(1/\sqrt{m}) > \Theta(n\min(1/\sqrt{n}, \lambda_{\rm p})/m)$.

Remark 2: Our framework could well accommodate node mobility. Notice that neither the hybrid protocol model nor the unconstraint time analysis rely on mobility pattern. Specifically, if primary users move according to the random walk model, optimal throughput–delay tradeoff in the primary network can be achieved by the cell-partitioned TDMA scheme [13, Scheme 3]. Therefore, the results in Corollary 3 can be obtained following a similar argument as in Remark 1.

We can extend the theorem to other variations of ad hoc networks, such as hybrid networks [7].

Corollary 4: If the primary network is equipped with $k = \Omega(\sqrt{n})$ base stations, the capacity of it is $\lambda_{\rm p} = \Theta(k/n)$, and the optimal throughput–delay tradeoff for the secondary network is $D_{\rm s} = \Theta(m\lambda_{\rm s})$, $\Theta(\sqrt{\lambda_{\rm p} n}/m) < \lambda_{\rm s} \leq \Theta(1/\sqrt{m})$, if $\Theta(1/\sqrt{m}) > \Theta(\sqrt{\lambda_{\rm p} n}/m)$.

Remark 3: Though base stations are wired together, they still need to access the wireless channel to communicate with primary users. In these kinds of hybrid networks, [7] shows that the throughput capacity can be achieved by tessellating the network into hexagons and assigning one base station for each hexagon. The diameter of the hexagon, $\Theta(\sqrt{1/k})$, is the average primary transmission range. Different groups of hexagons are scheduled in a cell-partitioned TDMA fashion.

Now we try to apply the results to networks with multicast traffic pattern.

Corollary 5: If each primary user wishes to send information to k independently chosen destinations, with $k = o(n/\log n)$, then the per-node multicast capacity of the primary network is $\lambda_{\rm p} = \Theta(1/\sqrt{kn\log n}))$. The optimal throughput–delay tradeoff for the secondary network is $D_{\rm s} = \Theta(m\lambda_{\rm s})$, $\Theta(\sqrt{n}/m\sqrt{\log n}) < \lambda_{\rm s} \leq \Theta(1/\sqrt{m})$, if $m = \Omega(n^M)$ for some constant $M > 1$.

Remark 4: Capacity of the primary network can be achieved by a cell-partitioned TDMA scheme [8, Section V]. The primary transmission range *w.h.p.* is $\Theta(\sqrt{\log n/n})$. Notice that our analysis can be directly extended to a multicast traffic pattern since the capacity achieving scheme does not require multireception or broadcasting.

The above corollaries are consequences of centralized TDMA scheduling of primary networks. In the following, we consider two examples of independently relaying. An interesting case is that primary networks make use of distributed random access protocols such as carrier-sensing multiple access (CSMA) [18].

Corollary 6: If the primary network employs independent relay protocol and CSMA protocol, the capacity of primary network is $\Theta(1/\sqrt{n\log n})$. The optimal throughput–delay tradeoff for the secondary network is $D_{\rm s} = \Theta(m\lambda_{\rm s})$, $\Theta(\sqrt{n}/m\sqrt{\log n}) < \lambda_{\rm s} \leq \Theta(1/\sqrt{m})$, if $m = \Omega(n^M)$ for some constant $M > 1$.

Remark 5: Note that the CSMA protocol is one of the most prominent exceptions of the cell-partitioning round-robin paradigm discussed in Section V-A. In such cases, the independent relaying assumption (Section V-B) would lead to similar results. Indeed, by employing the sensing techniques proposed in [18], one can achieve $\lambda_{\rm p} = \Theta(1/\sqrt{n\log n})$ with simple manhattan routing over a network tessellation with unit area $\log n/n$, which follows the independent routing assumption. Hence, the above corollary follows according to Theorem 8.

Now we consider a primary network with general mobility [19]. The next result follows from the mobile version of Theorem 7, which is analogous to the static one.

Corollary 7: If the mobility of primary nodes can be characterized by a stationary spatial distribution function[5] with support of diameter $1/f(n) = \omega(1/\sqrt{n})$, then the capacity of primary network is $\lambda_{\rm p} = \Theta(1/f(n))$. The optimal throughput–delay tradeoff for the secondary network is $D_{\rm s} = \Theta(m\lambda_{\rm s})$, $\Theta(\sqrt{n}/m) < \lambda_{\rm s} \leq \Theta(1/\sqrt{m})$, if $m = \omega(n)$.

[5]Refer to [19] for a rigorous definition.

Remark 6: According to [19], the capacity of the primary network is achieved by employing a generic scheduling scheme that always enables transmissions when the receiver is within the transmission range $c_{13}/\sqrt{n}$ of the transmitter, and all other nodes are outside an interference range that is proportional to the transmission range. Such ideal scheduling cannot be characterized by a generalized TDMA scheme. However, the above results can be obtained by observing that the capacity-achieving routing protocol again follows the independent relaying mode, i.e., manhattan routing over a network tessellation with unit area $1/f^2(n)$.

A. Gaussian Channel Model

The physical interference model is a fixed rate on–off channel model. An alternative, i.e., the Gaussian channel model, generalizes data rate to be continuous in SINR, based on Shannon's capacity formula for the additive Gaussian noise channel.

We now briefly extend our results to this model. Specifically, the communication rate of a primary link (X_i, X_j) is given by

$$W = \log\left(1 + \frac{P|X_i - X_j|^{-\gamma}}{N + I_{\rm pp} + I_{\rm sp}}\right)$$

where $I_{\rm pp}$ and $I_{\rm sp}$ are interference from primary network and secondary network, respectively (cf. Section III). Similarly, we define the channel model for secondary links. We limit our interest to the case that W is bounded between two positive constants, i.e., $W_1 < W < W_2$. We note this is also the most realistic case and suffice to generalize the results of prior works [15], [16]. Let $\lambda_{\rm p}^{\rm s.a.}$ be the per-node throughput of the primary network in the absence of the secondary network.

Operation Rule 3: The secondary scheduling should ensure that $(\lambda_{\rm p}/\lambda_{\rm p}^{\rm s.a.}) \geq 1 - \delta_{\rm loss}$.

Theorem 9: Theorem 8 holds under Gaussian channel model if we substitute Operation Rule 3 for Operation Rules 1 and 2 with $\delta_{\rm loss} \in (0, 1)$.

Proof: We claim that Operation Rules 1 and 2 with appropriate parameters are sufficient conditions for Operation Rule 3. Indeed, setting $\alpha \leq e^{W_1} - 1$ and $\epsilon \leq ((1+\alpha) - (1+\alpha)^{1-\delta_{\rm loss}})^{1/1-\delta_{\rm loss}}$, it is easy to show that

$$\frac{\log\left(1 + \frac{P|X_i - X_{{\rm Rx}(i)}|^{-\gamma}}{N + I_{\rm pp} + I_{\rm sp}}\right)}{\log\left(1 + \frac{P|X_i - X_{{\rm Rx}(i)}|^{-\gamma}}{N + I_{\rm pp}}\right)} \geq \frac{\log(1+\alpha)}{\log(1+\alpha+\epsilon)} \geq 1 - \delta_{\rm loss}$$

for any active link $(X_i, X_{{\rm Rx}(i)})$. On the other hand, to ensure the scheduling of secondary network under the Gaussian channel model is feasible under $\mathscr{P}$, we set $\beta \leq e^{W_1} - 1$. The rest of the theorem is obvious. ■

Based on Theorem 9, we study the class of clustered networks [11], [12].

Corollary 8: If the distribution of the primary users follows shot-noise Cox processes, the primary network can achieve capacity following the scheme proposed in [12]. The asymptotic throughput and delay of the secondary network is the same as standalone networks if the communication scheme satisfies the condition of Theorem 9.

Remark 7: Since the model of shot-noise Cox processes used to model the nonuniform distribution of nodes and the corresponding results are complicated, we refer interested readers to [11] and [12] for details. We emphasize that even in these heterogeneous clustered networks, cell-partitioned TDMA schemes are still highly efficient to achieve maximal spatial concurrency, i.e., though the cell-partitioned scheme is designed to be hierarchical to handle inhomogeneities in the network, it still falls under our framework as a generalized TDMA scheme. Therefore, Corollary 8 is a straightforward consequence of Theorem 9.

Lastly, the results that can be obtained are not limited to the cases listed above. Since our framework only relies on a few general conditions, it is flexible and is able to accommodate various cognitive networks with different specific forms. For instance, one can otherwise let both the networks or only the secondary network be mobile.

VII. Conclusion

This paper studies the throughput and delay scaling of general cognitive networks and characterizes the conditions for them to achieve the same throughput and delay scaling as standalone networks. We propose a hybrid protocol model for secondary nodes to identify transmission opportunities and show that, based on it, communication schemes of standalone networks can be easily extended to secondary networks without harming the performance of primary networks. In particular, we show that secondary networks can obtain the same optimal performance as standalone networks when primary networks are classic static networks, networks with random walk mobility, hybrid networks, multicast networks, CSMA networks, networks with general mobility, or clustered networks. Our work provides fundamental insight on the understanding and design of cognitive networks.

References

[1] W. Huang and X. Wang, "Throughput and delay scaling of general cognitive networks," in *Proc. IEEE INFOCOM*, 2011, pp. 2210–2218.

[2] "Report of the Spectrum Efficiency Working Group," Federal Communications Commission Spectrum Policy Task Force, Washington, DC, Nov. 2002.

[3] "Cognitive radio technologies proceeding," FCC, Washington, DC, ET Docket No. 03-108, 2003 [Online]. Available: http://www.fcc.gov/oet/cognitiveradio/

[4] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[5] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 477–486, Aug. 2002.

[6] A. Agarwal and P. R. Kumar, "Capacity bounds for ad hoc and hybrid wireless networks," *Comput. Commun. Rev.*, vol. 34, no. 3, pp. 71–81, 2004.

[7] B. Liu, Z. Liu, and D. Towsley, "On the capacity of hybrid wireless networks," in *Proc. IEEE INFOCOM*, San Franciso, CA, 2003, vol. 2, pp. 1543–1552.

[8] X.-Y. Li, "Multicast capacity of wireless ad hoc networks," *IEEE/ACM Trans. Netw.*, vol. 17, no. 3, pp. 950–961, Jun. 2009.

[9] X.-Y. Li, Y. Liu, S. Li, and S. Tang, "Multicast capacity of wireless ad hoc networks under Gaussian channel model," *IEEE/ACM Trans. Netw.*, vol. 18, no. 4, pp. 1145–1157, Aug. 2010.

[10] A. Özgür, O. Lévêque, and D. N. C. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3549–3572, Oct. 2007.

[11] G. Alfano, M. Garetto, and E. Leonardi, "Capacity scaling of wireless networks with inhomogeneous node density: Upper bounds," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1147–1157, Sep. 2009.

[12] G. Alfano, M. Garetto, and E. Leonardi, "Capacity scaling of wireless networks with inhomogeneous node density: Lower bounds," in *Proc. IEEE INFOCOM*, 2009, pp. 1890–1898.

[13] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Optimal throughput-delay scaling in wireless networks: Part I: The fluid model," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2568–2592, Jun. 2006.

[14] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Optimal throughput-delay scaling in wireless networks—Part II: Constant-size packets," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5111–5116, Nov. 2006.

[15] S.-W. Jeon, N. Devroye, M. Vu, S.-Y. Chung, and V. Tarokh, "Cognitive networks achieve throughput scaling of a homogeneous network," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5103–5115, Aug. 2011.

[16] C. Yin, L. Gao, and S. Cui, "Scaling laws for overlaid wireless networks: A cognitive radio network versus a primary network," *IEEE/ACM Trans. Netw.*, vol. 18, no. 4, pp. 1317–1329, Aug. 2010.

[17] C. Wang, S. Tang, X.-Y. Li, and C. Jiang, "Multicast capacity of multihop cognitive networks," in *Proc. IEEE MASS*, 2009, pp. 274–283.

[18] C.-K. Chau, M. Chen, and S. C. Liew, "Capacity of large-scale CSMA wireless networks," in *Proc. ACM MobiCom*, Beijing, China, 2009, pp. 97–108.

[19] M. Garetto, P. Giaccone, and E. Leonardi, "Capacity scaling in delay tolerant networks with heterogeneous mobile nodes," in *Proc. ACM MobiHoc*, 2007, pp. 41–50.

[20] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity," in *Proc. 37th IEEE Conf. Decision Control*, 1998, vol. 1, pp. 1106–1110.

[21] E. Hewitt and K. Stromberg, *Real and Abstract Analysis*. New York: Springer, 1975.

[22] Y. Aumann and Y. Rabani, "An o(log k) approximate min-cut max-flow theorem and approximation algorithm," *SIAM J. Comput.*, vol. 27, no. 1, pp. 291–301, 1998.

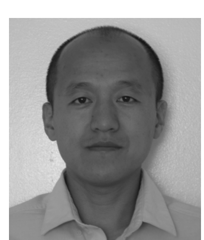
[23] C. Wang, X.-Y. Li, C. Jiang, S. Tang, Y. Liu, and J. Zhao, "Scaling laws on multicast capacity of large scale wireless networks," in *Proc. IEEE INFOCOM*, 2009, pp. 1863–1871.

[24] P. Li and Y. Fang, "The capacity of heterogeneous wireless networks," in *Proc. IEEE INFOCOM*, 2010, pp. 1–9.

[25] P. Li, Y. Fang, and J. Li, "Throughput, delay, and mobility in wireless ad hoc networks," in *Proc. IEEE INFOCOM*, 2010, pp. 1–9.

Wentao Huang received the B.S. degree in communications engineering from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2008, and is currently pursuing the M.S. degree in electronic engineering at Shanghai Jiao Tong University, Shanghai, China.

His current research interests include distributed systems, mobile computing, and network security.



Xinbing Wang (M'06) received the B.S. degree in automation (with honors) 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 faculty member with the Department of Electronic Engineering, Shanghai Jiao Tong University. His research interests include the scaling law of wireless networks and cognitive radio.

Dr. Wang has been an Associate Editor of the IEEE Transactions on Mobile Computing and a member of the Technical Program Committees of several conferences including ACM MobiCom 2012, ACM MobiHoc 2012, and IEEE INFOCOM 2009–2012.