Scaling Laws of Cognitive Ad Hoc Networks over General Primary Network Models

Cheng Wang, Changjun Jiang, Shaojie Tang, and Xiang-Yang Li, *Senior Member*, *IEEE*

Abstract—We study the capacity scaling laws for the *cognitive network* that consists of the *primary hybrid network* (PhN) and *secondary ad hoc network* (SaN). PhN is further comprised of an ad hoc network and a base station-based (BS-based) network. SaN and PhN are overlapping in the same deployment region, operate on the same spectrum, but are independent with each other in terms of communication requirements. The primary users (PUs), i.e., the ad hoc nodes in PhN, have the priority to access the spectrum. The secondary users (SUs), i.e., the ad hoc nodes in SaN, are equipped with cognitive radios, and have the functionalities to sense the idle spectrum and obtain the necessary information of primary nodes in PhN. We assume that PhN adopts one out of three classical types of strategies, i.e., *pure ad hoc strategy*, *BS-based strategy*, and *hybrid strategy*. We aim to directly derive multicast capacity for SaN to unify the unicast and broadcast capacities under two basic principles: 1) The throughput for PhN *cannot* be undermined in order sense due to the presence of SaN. 2) The protocol adopted by PhN does *not* alter in the interest of SaN, anyway. Depending on which type of strategy is adopted in PhN, we design the optimal-throughput strategy for SaN. We show that there exists a threshold of the density of SUs according to the density of PUs beyond which it can be proven that: 1) when PhN adopts the pure ad hoc strategy or hybrid strategy, SaN can achieve the multicast capacity of the same order as it is stand-alone; 2) when PhN adopts the BS-based strategy, SaN can asymptotically achieve the multicast capacity of the same order as if PhN were absent, if some specific conditions in terms of relations among the numbers of SUs, PUs, the destinations of each multicast session in SaN, and BSs in PhN hold.

Index Terms—Cognitive networks, hybrid networks, multicast capacity, random networks, percolation theory

1 Introduction

Nowadays, wireless networks are regulated by fixed spectrum assignment policy. The limited available spectrum coexists with the inefficiency in the spectrum usage, [1], [2]. To cope with this problem, dynamic spectrum access with cognitive radio has recently been investigated, which is a novel paradigm, called *cognitive network*, that improves the spectrum utilization by allowing secondary users (SUs) to exploit the existing wireless spectrum opportunistically without having a negative impact on PUs, i.e., licensed users.

In this paper, we focus on scaling laws of multicast capacity for cognitive networks. We construct the *cognitive network* as a superposition of two independent networks, called *primary network* and *secondary network*, that operate at the same time, space and frequency. The SUs are assumed to be equipped with cognitive radios and have the functionalities to sense the idle spectrum and obtain the necessary information of PUs [1], [2], [3]. We assume the primary network to be a *hybrid network*, denoted by PhN, consisting of base stations (BSs) and ad hoc nodes (PUs) [4], [5]. We assume the secondary network as an ad hoc network, denoted by SaN. To match the reality of spectrum consumption better, we assume that the network model has a *Pyramid structure*. That is, the number of PUs, which are licensed to access to the spectrum at any time, is relatively less than the number of SUs, which can opportunistically access to the spectrum.

Our model has three *novel points* relative to the existing works:

1. Since multicast capacity can be regarded as the general result of unicast and broadcast capacities [6], [7], [8], [9], we directly study the multicast capacity for cognitive networks to enhance the generality of this study.
2. Since pure ad hoc networks and BS-based networks (static cellular networks) can be regarded as the special case of hybrid networks in terms of the number of BSs [5], [10], we consider the model where the primary network is a hybrid network, which further enhances the generality of our model.
3. We use the *improved generalized physical model* [11], [12] that can capture the nature of wireless channel better than other classic interference models, such as *protocol model*, *physical model* [13], and the *generalized physical model* [14].
4. We focus on the model where PhN is an *extended scaling network* [14], [15], [16] while SaN is a *dense*

- *C. Wang and C. Jiang are with the Department of Computer Science and Engineering, Tongji University, and with the Key Laboratory of Embedded System and Service Computing, Ministry of Education, Building of Electronics and Information Engineering, NO. 4800, Caoan Road, Shanghai 201804, China. E-mail: 3chengwang@gmail.com, cjjiang@tongji.edu.cn.*
- *S. Tang is with the Department of Computer and Information Science, Temple University, 1805 N. Broad Street, Philadelphia, PA 19122. E-mail: stang7@iit.edu.*
- *X.-Y. Li is with the Tsinghua National Laboratory for Information Science and Technology (TNLIST), the Department of Computer Science and Engineering, Tongji University, and the Department of Computer Science, Illinois Institute of Technology, 10, West 31st Street, Chicago, IL 60616. E-mail: xiangyang.li@gmail.com.*

Manuscript received 2 Jan. 2012; revised 20 May 2012; accepted 11 July 2012; published online 20 July 2012.
Recommended for acceptance by E. Leonardi.
For information on obtaining reprints of this article, please send e-mail to: tpds@computer.org, and reference IEEECS Log Number TPDS-2012-01-0001.
Digital Object Identifier no. 10.1109/TPDS.2012.219.

scaling network [14], [15], rather than the model considered in most existing works where the primary and secondary networks are both dense scaling, [17], [18], [19], [20].

The diversity of scaling patterns of the primary and secondary networks expands the technical challenge and strengthens the theoretical contribution of this paper.

We intend to derive the multicast capacity for SaN under two *basic principles*: 1) The order of throughput for PhN must *not* be undermined by the presence of SaN. 2) The protocol adopted by PhN will *not* alter anyway due to the presence of SaN. These two basic principles are coincident with the abstract of practical techniques of cognitive networks. We first derive the upper bounds of multicast capacity for a single network isomorphic to SaN, called *single* SaN. Obviously, we can use such upper bounds as those for SaN whatever strategy is adopted by PhN, because PhN and SaN always have negative influence (interference) on each other under the noncooperative communication scheme as long as they share the same spectrum at the same time. To compute such upper bounds, we directly exploit the *homogeneity* property and *randomness* property of network topology [21].

Our main work is to design multicast strategies for SaN under two principles mentioned above, by which the multicast throughput, i.e., the lower bounds of multicast capacity, for SaN can be achieved of the optimal order matching the upper bounds. We design two types of multicast strategies for SaN. In the first type of strategy, we devise the hierarchical multicast routing based on the *highway* system consisting of the *first-class highways* (FHs) and *second-class highways* (SHs) [8], and we use a hierarchical TDMA scheme to schedule those highways. In the other type, we build the routing based on the highway system only comprised of SHs, to avoid the bottleneck on the accessing path into *highways* for some cases [14]. By integrating these two strategies together, we obtain the achievable multicast throughput as the lower bounds of multicast capacity for SaN.

Protecting the capacity for PhN from decreasing in order sense is the precondition in the design of any strategy for SaN. Our solution is to set a *preservation area* (PA) for every node in PhN. As an important characteristic different from other related works such as [17], [18], [19], [20], [22], we allow a PA to be *dynamic* according to the state of the corresponding primary node. Benefitting from the dynamics of PAs, SUs can access opportunistically into the spectrum from both time domain and space domain. While, static PAs used in [17], [18], [19], [20], [22] make some SUs never be served. In our solution, an intuitive view is that: When a link in PhN is scheduled, the receiver can receive data at a rate of the same order as in the scenario where PhN monopolizes the spectrum, as long as all active transmitters in SaN are out of a large enough PA of this receiver; similarly, when a link in SaN is scheduled, the receiver can receive data at the same rate (in order sense) as that for a single SaN, as long as this receiver is out of all PAs of the active transmitters in PhN.

Two technical challenges in our design are listed as follows:

What is the optimal size of PA with respect to the capacity for both SaN and PhN? As discussed above, the larger PAs are better for protecting the throughput for PhN. Meanwhile, too large PAs will result in a decrease in throughput for SaN. In other words, there is a tradeoff between the throughputs for PhN and SaN in terms of the size of PAs. Furthermore, it is easy to understand that the design of multicast strategies for SaN depends on the specific strategy adopted by PhN. As a hybrid network, PhN could generally adopt three broad categories of multicast strategies, according to [4], [5], [10], [23]. The first one is the classical BS-based strategy under which communications between any users are relayed by some specific BSs. The second one is the *pure ad hoc strategy*, i.e., the multihop scheme without any BS-supported. The third one is the *hybrid strategy*, i.e., the multihop scheme with BS-supported. According to these three strategies adopted by PhN, we define the appropriate PAs for each PU and BS, and call them *A-Type* PA and *B-Type* PA, respectively. Specifically, under pure ad hoc strategy, the B-Type PAs are never *active*; under BS-based strategy, the B-Type PAs are always *active*; and under hybrid strategy, both A-Type PAs and B-Type PAs might be *active* in a certain time.

How to build the highways, including FHs and SHs? Different from the traditional *highways* in [8], [14], [16], the construction of highways in SaN is more complicated because it is involved with the blocking of some active PAs. For FHs, we design a detouring scheme under which every FH detours the PAs, and we can prove that the produced FHs have the large enough density and capacity to support the relay of data in SaN. For SHs, we design a hierarchical TDMA scheduling scheme by which sufficient amount of SHs can be scheduled in a constant scheduling period, and all SUs have the opportunity to be served via accessing to the SHs, except when PhN adopts BS-based strategy.

As the final result, combining the upper bounds and lower bounds, we show that: 1) When PhN adopts the pure ad hoc strategy or hybrid strategy, the per-session multicast capacity (PMC) for SaN is of order $\Theta(\frac{1}{\sqrt{mm_d}})$ when

$$m_d = O\left(\frac{m}{(\log m)^3}\right),$$

and is of order $\Theta(\frac{1}{m})$ when $m_d = \Omega(\frac{m}{\log m})$, where m is the total number of SUs and m_d is the number of destinations of each multicast session in SaN. 2) When PhN adopts the BS-based strategy, an infinitesimal fraction of SUs cannot be served. The PMC for SaN is *asymptotically* of the same order as in Case 1.

The rest of the paper is organized as follows: In Section 2, we introduce the system model. In Section 3, we present main results. We propose the upper bounds of multicast capacity for the secondary network in Section 4, and derive the achievable throughput as the lower bounds of multicast capacity by designing the specific multicast strategies in Section 5. In Section 6, we review the related works. In Section 7, we draw some conclusions and future perspective.

NOTATIONS: In the paper, we adopt the following notations:

- $x \to \infty$ denotes that variable x takes value to infinity.
- For a discrete set $\mathcal{U}$, $|\mathcal{U}|$ represents its cardinality.
- For a continuous region $\mathcal{R}$, let $\|\mathcal{R}\|$ denote its area.

- For a 2-dimension line segment $\mathcal{L} = uv$, $|\mathcal{L}|$ represents its euclidean length. For a tree $\mathcal{T}$, denote its total euclidean edge lengths by $\|\mathcal{T}\|$.
- For event $\mathcal{E}$, denote the probability of $\mathcal{E}$ by $\Pr(\mathcal{E})$.
- To simplify the description, let the expression $\phi(n) : [\phi_0(n), \phi_1(n)]$ represent that $\phi(n) = \Omega(\phi_0(n))$ and $\phi(n) = O(\phi_1(n))$.

2 System Model

2.1 Network Deployment

The network model has a two-layer structure over a square region $\mathcal{R}(n) = [0, \sqrt{n}]^2$. The first layer is the PhN consisting of $\Theta(n)$ primary users (PUs, primary ad hoc nodes) and $b(n)$ BSs. In PhN, PUs are placed according to a Poisson point process (p.p.p.) of unit intensity over region $\mathcal{R}$; the region $\mathcal{R}$ is partitioned into $b(n)$ square subregions of side length $\sqrt{n/b(n)}$; one BS is located at the center of each subregion. Assume that BSs are connected via the high-bandwidth wired links that are certainly not the bottlenecks throughout the whole routing. The second layer is the *secondary ad hoc network* (SaN) consisting of $\Theta(m)$ secondary users (SUs, secondary ad hoc nodes). In SaN, SUs are distributed according to a p.p.p. of intensity $\frac{m}{n}$ over the region $\mathcal{R}$. We randomly choose n_s (or m_s) nodes from all PUs (or SUs) as the sources of multicast sessions in PhN (or SaN), and for each PU v^p (or SU v^s), pick uniformly at random n_d PUs (or m_d SUs) as the destinations. From Chebychev's inequality (Lemma B.3 in Appendix B, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TPDS.2012.219), we can assume that the numbers of PUs and SUs are n and m, respectively, as in [14], [24], which does not change our results in order sense. The following are our basic assumptions that are coincident with the abstract of practical techniques of cognitive networks.

Assumption 1. *PhN operates as if SaN were absent. That is, PhN does not alter its protocol due to SaN anyway.*

Assumption 2. *SaN knows the locations of nodes in PhN and the protocols adopted by PhN.*

2.2 Network Scaling Model

In the research of scaling laws of network capacity, in terms of the scaling method of network, there are two typical models: the *extended networks*, in which the total area is fixed and the density of nodes increases, and *dense networks*, in which the density of nodes is fixed and the total area increases [14], [15], [25]. Indeed, the *extended networks* and *dense networks* are the representative cases of *extended scaling model* and *dense scaling model*, respectively, [11], [12], [26]. To determine which type of network scaling model SaN belongs to, we dig out the main difference between these two scaling models. Due to limited space, we move the detailed analysis to Appendix B.2, available in the online supplemental material. According to the setting in Section 2.1, PhN obviously belongs to the *extended network model*; the *critical side length* of SaN is of order $\Theta(\sqrt{\log m \cdot \frac{n}{m}})$. From Definition B.5 in Appendix B, available in the online supplemental material, when $n = o(m/\log m)$, SaN is *dense scaling*, which is the case considered in this paper. This case can be formulated into the following assumption.

Assumption 3. *PhN and SaN are overlapped into a layered network with a* Pyramid structure. *Specifically,* $n = o(\frac{m}{\log m})$.

The case where SaN is also *extended scaling* may be treated as a future work.

2.3 Communication Model

We adopt an improved generalized physical model [11], [12], which captures the nature of wireless channels better than the *protocol model* and *physical model* [13], and can avoid the pathological throughput degradation for dense scaling networks that happens under the *generalized physical model*.

Let $\mathcal{V}(\tau)$ denote the set of transmitters in time slot τ. Then, during any time slot $\tau : v_i \in \mathcal{V}(\tau)$, v_i and v_j can communicate via a direct link, over a channel with bandwidth B, of rate

$$R(v_i, v_j; \tau) = \min\left\{R_0, B\log\left(1 + \frac{S(v_i, v_j; \tau)}{N_0 + I(v_i, v_j; \tau)}\right)\right\}, \quad (1)$$

where $R_0 > 0$ is a predefined constant, the constant $N_0 > 0$ is the ambient noise, $S(v_i, v_j; \tau)$ is the strength of the signal initiated by v_i at the receiver v_j, and $I(v_i, v_j; \tau)$ is the sum interference on v_j produced by all nodes belonging to the set $\mathcal{V}(\tau) - \{v_i\}$. The wireless propagation channel typically includes path loss with distance, shadowing and fading effects. In this paper, we assume that the channel gain depends only on the distance between a transmitter and a receiver, ignore shadowing and fading, and define the channel power gain function as $\ell(v_i, v_j) = d_{ij}^{-\alpha}$, where $d_{ij} = d(v_i, v_j) = \|v_i v_j\|$ is the euclidean distance between two nodes v_i and v_j, $\alpha > 2$ denotes the power attenuation exponent, [11], [12]. Based on this, $S(v_i, v_j; \tau)$ and $I(v_i, v_j; \tau)$ are defined as:

$$S(v_i, v_j; \tau) = P(v_i; \tau) \cdot \ell(v_i, v_j), I(v_i, v_j; \tau) = \sum_{v_k \in \mathcal{V}(\tau) - \{v_i\}} P(v_k; \tau) \cdot \ell(v_k, v_j),$$

where $P(v_l; \tau)$ denotes the transmitting power of $v_l \in \mathcal{V}(\tau)$ in time slot τ.

For no intercommunication occurs between the two networks, we have: For a link $v_i \to v_j$ in PhN, denote the sum of interference on v_j produced by all nodes in time slot τ by $I_p(v_i, v_j; \tau)$, then it holds that

$$I_p(v_i, v_j; \tau) = I_{pp}(v_i, v_j; \tau) + I_{sp}(v_i, v_j; \tau), \quad (2)$$

where $I_{pp}(v_i, v_j; \tau)$, or $I_{sp}(v_i, v_j; \tau)$, denotes the sum of interference on v_j produced by all nodes in $\mathcal{V}(\tau) \cap \mathcal{V}_p - \{v_i\}$, or in $\mathcal{V}(\tau) \cap \mathcal{V}_s$, where $\mathcal{V}_p$ and $\mathcal{V}_s$ denote all nodes in PhN or SaN, respectively. For a link $v_i \to v_j$ in SaN, denote the sum of interference on v_j produced by all nodes in time slot τ by $I_s(v_i, v_j; \tau)$, then it holds that

$$I_s(v_i, v_j; \tau) = I_{ps}(v_i, v_j; \tau) + I_{ss}(v_i, v_j; \tau), \quad (3)$$

where $I_{ps}(v_i, v_j; \tau)$, or $I_{ss}(v_i, v_j; \tau)$, denotes the sum of interference on v_j produced by all transmitters in $\mathcal{V}(\tau) \cap \mathcal{V}_p$, or in $\mathcal{V}(\tau) \cap \mathcal{V}_s - \{v_i\}$.

We assume that all PUs and BSs transmit with the constant wireless transmission power P. This setting is the same as in [4], [5]. Note that the work can be further extended into the case that BSs with wireless transmission power $P \cdot \beta(n)$, where $\beta(n) = \Omega(1)$. For SaN, we assume that each SU, say v_i, in time slot τ, transmits with a specific transmission power $P(v_i,\tau) \in (0, P_0]$, if it is scheduled in τ, where $P_0 > 0$ is the maximum of the transmission power.

2.4 Capacity Definition

Since the priority of the licensed users, i.e., the PUs, must be guaranteed, there might be some nonlicensed users, i.e., SUs, that cannot be served. For example, when PhN adopts the *BS-based strategy* as presented in [5], all BSs operate simultaneously all the time. Therefore, in the peripheral of BSs, there will be some SUs that cannot access into the spectrum from neither time domain nor space domain. Since we assume that PhN and SaN operate on the same spectrum, there is no idle frequency via which SaN can access into the spectrum. Hence, we need to generalize the formal definition of multicast capacity proposed in [6]. Consequently, we define *asymptotic multicast capacity* that is similarly defined in [19]. Please see the definitions of *asymptotic aggregated multicast capacity* (Asymp-AMC) and *asymptotic per-session multicast capacity* (Asymp-PMC) in Appendix B.1, available in the online supplemental material. We adopt the same method of constructing multicast sessions as in [19], and assume that all nodes (users) in SaN act as the sources, i.e., $m_s = m$, where m is the total number of nodes and m_s is the number of sources.

3 Main Results

Now, we present the upper bounds and lower bounds of multicast capacity for the SaN; finally, combining the lower bounds and upper bounds, we obtain the multicast capacity.

3.1 Upper Bounds of Multicast Capacity for SaN

We first derive the upper bounds of multicast capacity for SaN as if the PhN were absent. Straightforwardly, such results are also the upper bounds for SaN when PhN works.

Theorem 1. *The PMC for SaN is at most of order*

$$\bar{\Lambda}^{\mathrm{P}} = \begin{cases} O\left(\frac{1}{\sqrt{m_d m}}\right) & \text{when} \quad m_d : \left[1, \frac{m}{(\log m)^2}\right] \\ O\left(\frac{1}{m_d \cdot \log m}\right) & \text{when} \quad m_d : \left[\frac{m}{(\log m)^2}, \frac{m}{\log m}\right] \\ O\left(\frac{1}{m}\right) & \text{when} \quad m_d : \left[\frac{m}{\log m}, m\right] \end{cases}.$$

The aggregated multicast capacity for SaN is at most of order $m \cdot \bar{\Lambda}^{\mathrm{P}}$. Here, m_d denotes the number of destinations of each multicast session in SaN.

The result in Theorem 1 always holds regardless of what strategy PhN adopts.

3.2 Lower Bounds of Multicast Capacity for SaN

We derive the lower bounds of multicast capacity by designing the strategies for SaN corresponding to three classical types of strategies adopted in PhN, [4], [5], [10].

3.2.1 When PhN Adopts Pure Ad hoc Strategy

In this case, all SUs involved in all multicast sessions can be served.

Theorem 2. *The achievable PMT for SaN is of order*

$$\Lambda^{\mathrm{P}} = \begin{cases} \Omega\left(\frac{1}{\sqrt{m_d m}}\right) & \text{when} \quad m_d : \left[1, \frac{m}{(\log m)^3}\right] \\ \Omega\left(\frac{1}{m_d (\log m)^{\frac{3}{2}}}\right) & \text{when} \quad m_d : \left[\frac{m}{(\log m)^3}, \frac{m}{(\log m)^2}\right] \\ \Omega\left(\frac{1}{\sqrt{m m_d \log m}}\right) & \text{when} \quad m_d : \left[\frac{m}{(\log m)^2}, \frac{m}{\log m}\right] \\ \Omega\left(\frac{1}{m}\right) & \text{when} \quad m_d : \left[\frac{m}{\log m}, m\right]. \end{cases} \tag{4}$$

The achievable AMT for SAN is of order $\Theta(m \cdot \Lambda^{\mathrm{P}})$.

3.2.2 When PhN Adopts BS-Based Strategy

In this case, some SUs are covered by the *B-Type* PAs that are always *active*, then they cannot be served. Under our strategy for SaN, we can ensure that there are at least $\rho_s(m) \cdot m$ multicast sessions of SaN whose $\rho_d(m) \cdot m_d$ destinations can be served, where $\rho_s(m) \to 1, \rho_d(m) \to 1$, as $m \to \infty$.

Theorem 3. *Under two cases, i.e., 1) $m_d = \Omega(\log n)$, or 2) $m_d = O(\log n)$ and $b(n) = o(\frac{m}{m_d \cdot \log m})$, the* asymp-achievable *PMT for SaN is of order Λ^{P}, where Λ^{P} is defined in (4). The* asymp-achievable *AMT for SaN is of order $m \cdot \Lambda^{\mathrm{P}}$.*

3.2.3 When PhN Adopts Hybrid Strategy

In this case, we set SaN to be idle when the downlinks and uplinks involved with the BSs are scheduled in PhN, and we schedule SaN in the other phases. Under this strategy, we get the throughput for SaN of the same order as in Theorem 2.

3.3 Multicast Capacity for SaN

Combining the upper bound on multicast capacity for SaN, described in Theorem 1, and the lower bound in Section 3.2, we can obtain

Theorem 4. *When PhN adopts the pure ad hoc strategy or the hybrid strategy, the PMC for SaN is of order*

$$C^{\mathrm{P}} = \begin{cases} \Theta\left(\frac{1}{\sqrt{m_d m}}\right) & \text{when} \quad m_d : \left[1, \frac{m}{(\log m)^3}\right] \\ \Theta\left(\frac{1}{m}\right) & \text{when} \quad m_d : \left[\frac{m}{\log m}, m\right]. \end{cases} \tag{5}$$

The aggregated multicast capacity for SaN is of order $m \cdot C^{\mathrm{P}}$. When PhN adopts the BS-based strategy, Asymp-PMC and Asymp-AMC for SaN are of order C^{P} and $m \cdot C^{\mathrm{P}}$, respectively.

From Theorem 4, we know that there still exists a gap between the upper bounds and lower bounds for the case that $m_d : [m/(\log m)^3, m/\log m]$. How to close the gap may be left for future work.

4 Upper Bounds for SaN

In this section, we compute the upper bounds on multicast capacity for SaN when PhN is absent. Intuitively, such upper bounds are possibly too loose. However, in Section 5, we show that such bounds can be achieved indeed. Thus, we focus on the single SaN, where nodes are distributed into the square region $[0, \sqrt{n}]^2$ according to a p.p.p. of density $\frac{m}{n}$ with $n = o(\frac{m}{\log m})$, and we begin to prove Theorem 1. First, we have

Lemma 1. *The PMC for SaN is at most of order*

$$O\left(\max\left\{\frac{1}{\sqrt{m \cdot m_d}}, \frac{\log m}{m}\right\}\right).$$

Proof. We can use the *homogeneity* property of the network to prove this result. A useful analysis tool, called *transmission arena* [21], is exploited. Please see the detailed proof in Appendix C.1, available in the online supplemental material. □

Lemma 2. *The PMC for SaN is at most of order*

$$\begin{cases} O\left(\frac{1}{m_d \log m}\right) & \text{when} \quad m_d = O\left(\frac{m}{\log m}\right) \\ O\left(\frac{1}{m}\right) & \text{when} \quad m_d = \Omega\left(\frac{m}{\log m}\right). \end{cases}$$

Proof. This result can be proved according to the randomness positions of the nodes in the network. Randomness property produces some relatively isolated clusters of nodes. These clusters can act as a bottleneck on the multicast capacity. Please see the detailed proof in Appendix C.2, available in the online supplemental material. □

Combining Lemma 1 and Lemma 2, we obtain Theorem 1.

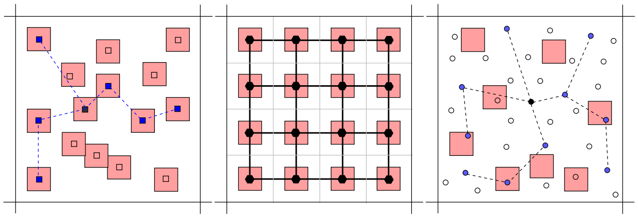
5 Lower Bounds for SaN

Generally, the lower bounds of the capacity can be obtained by designing the specific multicast strategy. Denote a class of the multicast strategies by $\mathbb{S}$, consisting of routing scheme $\mathbb{S}^r$ and transmission scheduling $\mathbb{S}^t$. The routing scheme $\mathbb{S}^r$ might have a hierarchical structure consisting of ς phases that correspond to *subrouting schemes* $\mathbb{S}^{r_1}$, $\mathbb{S}^{r_2}, \ldots, \mathbb{S}^{r_\varsigma}$, accordingly, the transmission scheduling $\mathbb{S}^t$ consists of ς phases, i.e., $\mathbb{S}^{t_1}, \mathbb{S}^{t_2}, \ldots, \mathbb{S}^{t_\varsigma}$, where $\varsigma \geq 1$ is a constant and it means that the routing scheme $\mathbb{S}^r$ is nonhierarchical when $\varsigma = 1$.

5.1 Overview of Multicast Strategy

Due to the overwhelming priority to access the spectrum, PhN operates as if SaN were absent. Denote the protocol for PhN by $\mathbb{S}$ consisting of routing scheme $\mathbb{S}^r$ and transmission scheduling $\mathbb{S}^t$. We will design the multicast strategy for SaN according to the specific strategy adopted by PhN. Assume that every subphase of $\mathbb{S}^t$ operates under an independent TDMA scheme. Denote the scheduling periods of those TDMA schemes by K_j^2, where $K_j \geq 3$ and $j = 1, 2, \ldots, \varsigma$.

We first construct the specific PAs for each PU and each BS in PhN, and call them *A-Type* PA and *B-Type PA*, respectively. Please see the illustration in Figs. 1a and 1b. Then, at slot $\tau(\tau_j, j)$, $1 \leq j \leq \varsigma$, and $1 \leq \tau_j \leq K_j^2$, which represents the τ_jth time slot in a scheduling period in phase j, we set the status of PAs of the nodes scheduled in $\tau(\tau_j, j)$



(a) PU Layer (b) BS Layer (c) SU Layer

Fig. 1. PhN consists of *PU layer* and *BS layer*; SaN has only one layer, i.e., *SU layer*. (a) The black small square is the source of a given multicast session. The bigger shaded squares are *A-Type PAs*. (b) The small black hexagons are the BSs that are placed in the center positions of subregions of area $\frac{n}{b(n)}$. The shaded squares around BSs are *B-Type PAs*. (c) Dashed lines denote the *EST* of a given multicast session.

as *active*. Thus, in any time slot τ, the region $\mathcal{R}(n) = [0, \sqrt{n}]^2$ is partitioned into two regions: the *occupied region*$\mathcal{O}(\tau)$, which is a region covered by all *active* PAs in time slot τ, and the *vacant region*$\mathcal{V}(\tau)$, which is the complement of region $\mathcal{O}(\tau)$. Accordingly, we denote the set of the SUs covered by the *occupied region* $\mathcal{O}(\tau)$ or surrounded by the *active* PAs in time slot τ, as the set $\mathcal{P}(\tau)$. Finally, for a given multicast session $\mathcal{M}_{\mathcal{S},i}$ with the source node $v_{\mathcal{S},i}$, when $v_{\mathcal{S},i} \in \bigcap_{\tau(\tau_j,j)} \mathcal{P}(\tau)$, the multicast session $\mathcal{M}_{\mathcal{S},i}$ will be ignored; otherwise, by using the algorithm in [6], we construct the euclidean spanning tree (EST), denoted by $EST(\tilde{\mathcal{U}}_{\mathcal{S},i})$, based on the set $\tilde{\mathcal{U}}_{\mathcal{S},i}$, where $\mathcal{U}_{\mathcal{S},i}$ is the *spanning set* and $\tilde{\mathcal{U}}_{\mathcal{S},i} = \mathcal{U}_{\mathcal{S},i} - \bigcap_{\tau(\tau_j,j)} \mathcal{P}(\tau)$. Then, like the multicast routing designed in [6], our routing for SaN is guided by the spanning tree $EST(\tilde{\mathcal{U}}_{\mathcal{S},i})$. The communication of each link in $EST(\tilde{\mathcal{U}}_{\mathcal{S},i})$ is routed via the highway system similar to that in [8], if applicable. However, intuitively, the routing paths might be broken by the *active* PAs in some (or even all) time slots. Thus, how to deal with such intractability? Is it possible that the optimal throughput for SaN can be achieved? Here, the called optimal order of throughput is the upper bounds of multicast capacity in the single network isomorphic to SaN. Then, given a specific protocol in PhN, there are *three questions* in the design of multicast strategies for SaN.

Question 1. How to construct and schedule the FHs and SHs such that, in any time slot when SaN is scheduled, no link along the highways comes across the active PAs?

Question 2. How large is the density of the highway system in SaN, including the FHs and SHs, if exists?

Question 3. How to ensure our multicast strategy to serve the SUs (or multicast sessions) as much as possible?

Obviously, the status of PAs and the method of constructing the highway system in SaN are determined by the strategy adopted by PhN. Thus, all of these three questions should be answered depending on the protocol of PhN. According to the existing works [4], [5], when the TDMA transmission scheduling scheme is adopted in PhN, the strategy for hybrid network can be classified into three types, i.e., *pure ad hoc strategy*, *BS-based strategy*, and *hybrid strategy*. Next, we introduce concisely these strategies, and answer the three questions above according to the specific protocol of PhN.

5.2 When PhN Adopts Pure Ad Hoc Strategy

In PhN, under the pure ad hoc strategy, since no BS is used, all *B-Type* PAs are always *inactive*. For an *A-Type* PA, its

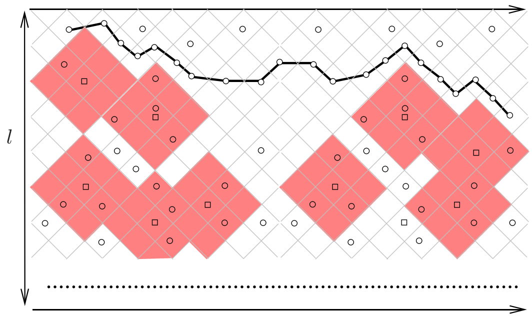


Fig. 2. The cells are of side length $\bar{c} = c \cdot \sqrt{\frac{n}{m}}$. The slab is of side length $l = (\kappa \cdot \log h - \epsilon_h) \cdot \sqrt{2}\bar{c}$. The shaded regions are the *A-Type* PAs. The small square nodes at the center of *A-Type* PAs represent the PUs, and the small circle nodes represent the SUs. Those nonshaded cells containing at least one SU are called *nonprotected open*.

status (*active* or *inactive*) is determined by the routing and transmission scheduling adopted by PhN. To achieve the optimal order of throughput for PhN, we assume that the multicast strategy in [8] is adopted in PhN. The strategy is divided into two phases.

Denote the routing scheme in the first phase in PhN by $\mathbb{S}^{r_1}$ and denote the transmission scheduling in PhN as $\mathbb{S}^{t_1}$. In this phase, the strategy is designed based on the scheme lattice (Definition B.3 in Appendix B, available in the online supplemental material) $\mathbb{L}(\sqrt{n}, c, \frac{\pi}{4})$, where $c > 0$ is a constant defined in [8]. The routing is constructed based on the FHs consisting of the short links of constant length; and those short links are scheduled by a TDMA scheme. Assume that the constant scheduling period is K_1^2 ([8, $K_1 = 3$]).

Denote the routing scheme in the second phase in PhN by $\mathbb{S}^{r_2}$ and denote the transmission scheduling as $\mathbb{S}^{t_2}$. In this phase, the strategy is designed based on the scheme lattice $\mathbb{L}(\sqrt{n}, \sigma\sqrt{\log n} - \epsilon_n, 0)$, where σ is a constant defined in [8, Lemma 4] and ϵ_n is an adjusting constant to ensure the value of $\sqrt{n}/(\sigma\sqrt{\log n} - \epsilon_n)$ to be an integer; the routing is constructed based on the SHs consisting of the links of length of order $\Theta(\sqrt{\log n})$; and those links are also scheduled by a TDMA scheme of constant period K_2^2 ([8, $K_2 = 4$]). An important method is called the *parallel transmission scheduling* under which $\Theta(\log n)$ links initiating from each active cell are simultaneously scheduled.

As in PhN, the highway system in SaN also consists of two levels highways: FHs and SHs. Next, we first introduce them from the situation where PhN is not considered, and then extend them to the real situation in which the priority of PhN is inviolable.

5.2.1 Highways for SaN Absent of PhN

When the PhN is ignored, the highway system can be constructed by the similar method in [8]. The FHs are indeed the highways constructed in [14]. The SHs that are built without using percolation theory [14].

Existence and density of FHs. The FHs are constructed and scheduled based on the *scheme lattice* $\mathbb{L}(\sqrt{n}, c\sqrt{\frac{n}{m}}, \frac{\pi}{4})$ as illustrated in Fig. 2. Since the distribution of SUs follows a Poisson with mean c^2 (derived by the intensity $\frac{m}{n}$ times the area of the cell $c^2 \cdot \frac{n}{m}$), the cell in $\mathbb{L}(\sqrt{n}, c\sqrt{\frac{n}{m}}, \frac{\pi}{4})$ has the same *open* probability as the cell of the lattice in [14, Fig. 2], i.e., $p \equiv 1 - e^{-c^2}$. Let $h = \frac{\sqrt{m}}{\sqrt{2}\cdot c}$. From Lemma B.1 in Appendix B, available in the online supplemental material, by choosing a large enough c, there are, *w.h.p.*, $\Omega(h)$ paths crossing the network from left to right. These paths can be grouped into disjoint sets consisting of $\delta \log h$ paths, with each group crossing a rectangle slab of size $\sqrt{n} \times (\kappa \log h - \epsilon_h) \cdot \sqrt{2 \cdot \frac{n}{m}} \cdot c$, where $\kappa > 0$, δ is small enough, and ϵ_h is vanishingly small so that the side length of each rectangle is an integer.

Therefore, we can divide such a rectangle into $\delta \log h$ slices of size $\sqrt{n} \times \varpi(n,m)$, where $\varpi(n,m) = \Theta(\sqrt{\frac{n}{m}})$. Denote the jth slice in the ith slab by $s_h(i,j)$, where $1 \le i \le \frac{h}{\kappa \cdot \log h - \epsilon_h}$ and $1 \le j \le \delta \cdot \log h$. Then, we allocate the relay burden of nodes in $s_h(i,j)$ to a specific FH, denoted by $\hbar_h(i,j)$ representing the jth horizontal FH in the ith slab. Similarly, for the vertical case, we can explain the corresponding $s_v(i,j)$ and $\hbar_v(i,j)$, and define the mapping between them.

Existence and density of SHs. The SHs are constructed and scheduled based on the scheme lattice

$$\mathbb{L}\left(\sqrt{n}, \sigma\sqrt{\log m \cdot \frac{n}{m}} - \epsilon_m, 0\right).$$

For the *dense scaling* network model, the parallel transmission scheduling does not work [8]. Then, in SaN, having no parallel SHs like in PhN [5], there exists only one SH in each column (or row). Denote each column as $s'_v(i)$, where

$$1 \le i \le \frac{\sqrt{n}}{\sigma \cdot \sqrt{\log m \cdot \frac{n}{m}} - \epsilon_m}.$$

Based on this, we allocate the relay burden of nodes in $s'_v(i)$ to a specific SH, denoted by $\hbar'_v(i)$ representing the SH contained in the ith column. Similarly, for the vertical case, we can explain the corresponding $s'_h(i)$ and $\hbar'_h(i)$, and define the mapping between them. Remark that we can use a TDMA with the constant period K_2^2, to schedule the SHs. We will provide the detailed analysis in Section 5.2.4.

5.2.2 Highways for SaN Present of PhN

Consequently, we construct the highway system for SaN based on percolation theory [14], ensuring that no highway in SaN crosses the *active* PAs in any time slot.

Existence and construction of FHs. The FHs in SaN will be scheduled in the first phase in PhN, i.e., $\mathbb{S}^{r_1}$ or $\mathbb{S}^{t_1}$. In this phase, around each PU, we build its *A-Type* PA as a cluster of nine cells of side length $c\sqrt{\frac{n}{m}}$, as illustrated in Fig. 2. In any time slot τ_1, an *A-Type* PA is *active* or *inactive* depending on whether the central PU is scheduled (including both transmitting and receiving) or not. Recall that the transmission scheduling of FHs in PhN is a TDMA scheme with constant period K_1^2. Then, in any time slot of the scheduling for FHs in PhN, the number of scheduled cell will be $\frac{1}{K_1^2}$ of the number of all cells, i.e., $\frac{n}{c^2}$. That is, in some time slots, the number of scheduled PUs is of order $\Theta(n)$. Hence, it has no impact on our results when the dynamic of the status of *A-Type* PAs in the first phase is ignored, i.e., all *A-Type* PAs are always regarded as *active* in the first phase.

Next, we build the FHs in SaN that coexists with PhN. We first modify the definition of *open* cells [14]. A cell in

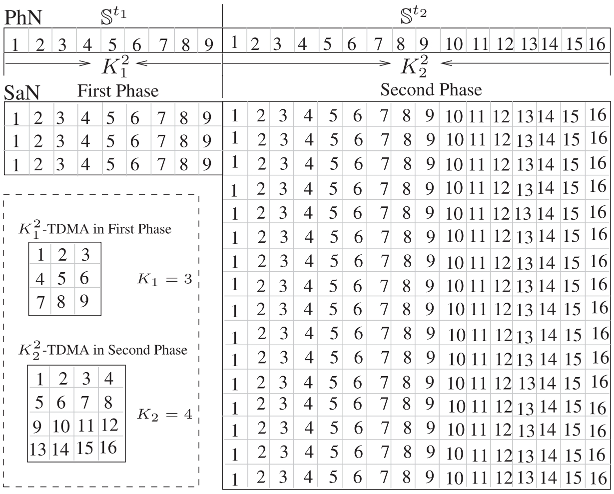


Fig. 3. Illustration of scheduling scheme. For readability, we describe the case that $K_1 = 3$ and $K_2 = 4$. The scheduling for SaN is divided into two phases corresponding to the two phases in PhN. In the first (or second) phase, SaN schedules in sequence $K_1 \times K_1$ (or $K_2 \times K_2$) cells in the *scheme lattice* $\mathbb{L}(\sqrt{n}, c\sqrt{n/m}, \frac{\pi}{4})$ (or $\mathbb{L}(\sqrt{n}, \sigma\sqrt{\log m \cdot n/m} - \epsilon_m, 0)$) during one period of $3K_1^2$ (or $K_2^4 = K_2^2 \times K_2^2$) slots; that is, each cell will be scheduled continuously 3 (or K_2^2) slots. Remark that, during the continuous K_2^2 slot for each cell in $\mathbb{L}(\sqrt{n}, \sigma\sqrt{\log m \cdot n/m} - \epsilon_m, 0)$, the cell is really scheduled only when it is not covered by *active* PAs.

$\mathbb{L}(\sqrt{n}, c\sqrt{\frac{n}{m}}, \frac{\pi}{4})$ is called *nonprotected open* if it is nonempty, i.e., it contains at least one SU, and does not belong to any *A-Type* PAs. Please see the illustration in Fig. 2. Then, we have the following lemma.

Lemma 3. *When $n = o(m)$, a cell in $\mathbb{L}(\sqrt{n}, c\sqrt{\frac{n}{m}}, \frac{\pi}{4})$ is nonprotected open with probability $p_s \to p$ as $n \to \infty$.*

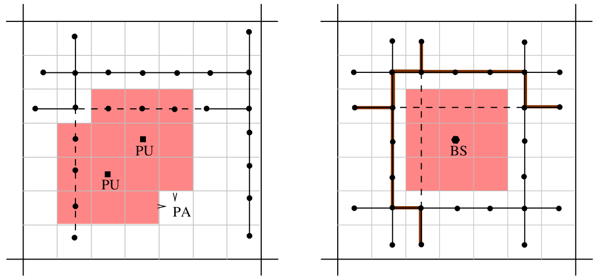
Please see the detailed proof in Appendix C.3, available in the online supplemental material. By Lemma 3, we can prove the existence of FH in SaN, and obtain the same density of FHs in SaN as that in PhN. Thus, we can use the same notations of FHs in the situation absent of PhN, which will be used in Algorithm 1.

Scheduling of FHs in SaN. Let the transmitting power of SUs in the first phase be $P' \cdot (c\sqrt{\frac{n}{m}})^\alpha$, where $P' \in (0, P_0]$ is a constant. Recall that the constant P_0, defined in Section 2.3, is the maximum transmitting power in SaN. Obviously, $P' \cdot (c\sqrt{\frac{n}{m}})^\alpha \in (0, P_0]$. Because the FHs in SaN detour all PAs, the capacity of FHs in PhN can be protected from increasing in terms of order, which is proved in Theorem 5. As illustrated in Fig. 3, in the first phase in SaN, the *scheduling unit* is also the cluster of $K_1 \times K_1$ cells. Unlike in PhN, each cell in a *scheduling unit* is scheduled continually three slots. By this method, it holds that there is at least one out of these three slots during which the nearest distance between the transmitter in PhN to the receiver in SaN is of a constant order.

Scheduling of SHs in SaN. As the new *scheme lattice*

$$\mathbb{L}\left(\sqrt{n}, \sigma\sqrt{\log m \cdot \frac{n}{m}} - \epsilon_m, 0\right)$$

is used, we define the new *A-Type* PA that is a cluster of nine cells in $\mathbb{L}(\sqrt{n}, \sigma\sqrt{\log m \cdot \frac{n}{m}} - \epsilon_m, 0)$ centered at a PU.



(a) SHs Meets *A-Type* PA (b) SHs Meets *B-Type* PA

Fig. 4. SHs built based on $\mathbb{L}(\sqrt{n}, \sigma\sqrt{\log m \cdot \frac{n}{m}} - \epsilon_m, 0)$. (a) When PhN adopts pure ad hoc strategy, the SHs in SaN need not detour the PAs, but wait for their *inactive* status. (b) When PhN adopts BS-based strategy, since the PAs are always *active*, SHs in SaN have to detour all *B-Type* PAs along the SHs adjacent to the PAs. The bold polylines denote the detouring paths.

Please see the illustration in Fig. 4a. Let the transmitting power of SUs in the second phase be

$$P' \cdot \left(\sigma\sqrt{\log m \cdot \frac{n}{m}} - \epsilon_m\right)^\alpha,$$

where $P' \in (0, P_0]$ is a constant. Obviously,

$$P' \cdot \left(\sigma\sqrt{\log m \cdot \frac{n}{m}} - \epsilon_m\right)^\alpha \in (0, P_0].$$

Do the SHs, constructed for SaN absent of PhN, still work now? The following Lemma 4 will answer this question.

Lemma 4. *$\bigcap_{\tau_2=1}^{K_2^2} \mathcal{P}(\tau(\tau_2, 2)) = \emptyset$, where $\mathcal{P}(\tau(\tau_2, 2))$ represents the set of SUs covered or surrounded by the active PAs in the time slot $\tau(\tau_2, 2)$, i.e., the τ_2th scheduling slot of the second phase in PhN.*

Please see the detailed proof in Appendix C.4, available in the online supplemental material. This lemma means that for any SU, there is at least one slot out of the scheduling period of the second phase in PhN, i.e., K_2^2 time slots, in which the SU can be possibly scheduled. Recall that K_2^2 is the constant period of TDMA scheme used for SHs in SaN absent of PhN. Hence, we can use a TDMA scheme with the period of $K_2^4 = K_2^2 \times K_2^2$ to schedule the SHs at least once. Since $K_2^2 \in (0, +\infty)$, we can obtain the same order of capacity for SHs in SaN regardless of the presence of PhN.

5.2.3 Multicast Strategy for SaN

For a given multicast session $\mathcal{M}_{\mathcal{S},i}$ with source $v_{\mathcal{S},i}$ and the *spanning set* $\mathcal{U}_{\mathcal{S},i}$, we first construct the EST $EST(\mathcal{U}_{\mathcal{S},i})$ by the method in [6]. Then, we can build the multicast routing tree based on the highway system and $EST(\mathcal{U}_{\mathcal{S},k})$. More specifically, for each communication pair in $EST(\mathcal{U}_{\mathcal{S},k})$, i.e., an edge, the packets will access to the specific FH via the specific SH. The strategy for SaN is divided into two phases that are synchronous to the two phases in PhN. See the illustration in Fig. 3a. The detailed multicast routing scheme is presented in Algorithm 1. To clarify the description, we first recall the formulation of the highway system:

- $s_h(x, y)$: The yth horizontal slice in the xth horizontal slab in the *scheme lattice* $\mathbb{L}(\sqrt{n}, c\sqrt{\frac{n}{m}}, \frac{\pi}{4})$.

- $s'_h(z)$: The zth row in
$$\mathbb{L}\left(\sqrt{n}, \sigma\sqrt{\log m \cdot \frac{n}{m}} - \epsilon_m, 0\right).$$
- $\hbar_h(x, y)$: The horizontal FH bearing the relay load initiated from the nodes in the slice $s_h(x, y)$.
- $\hbar'_h(z)$: The horizontal SH bearing the relay load initiated from the nodes in the row $s'_h(z)$.

The formulations for the vertical case are similarly defined.

Algorithm 1. Multicast Routing based on FHs and SHs
Input: The multicast session $\mathcal{M}_{\mathcal{S},k}$ and $EST(\mathcal{U}_{\mathcal{S},k})$.
Output: A multicast routing tree $\mathcal{T}(\mathcal{U}_{\mathcal{S},k})$.
1: **for** each link $u_i \to u_j$ of $EST(\mathcal{U}_{\mathcal{S},k})$ **do**
2: According to the position of u_i and u_j, determine the indexes a_i, b_i, y_i and c_j, d_j, x_j, where $u_i \in s_h(a_i, b_i) \cap s'_v(y_i)$; $u_j \in s_v(c_j, d_j) \cap s'_h(x_j)$.
3: Packets are drained from u_i into the horizontal FH $\hbar_h(a_i, b_i)$ via the vertical SH $\hbar'_v(y_i)$.
4: Packets are carried along the horizontal FH $\hbar_h(a_i, b_i)$.
5: Packets are carried along the vertical FH $\hbar_v(c_j, d_j)$.
6: Packets are delivered from the vertical FH $\hbar_v(c_j, d_j)$ to u_j along the horizontal SH $\hbar'_h(x_j)$.
7: **end for**
8: Considering the resulted routing graph, we merge the same edges (hops), and remove those circles which have no impact on the connectivity of the communications for $EST(\mathcal{U}_{\mathcal{S},k})$. Finally, we obtain the final multicast routing tree $\mathcal{T}(\mathcal{U}_{\mathcal{S},k})$.

5.2.4 Analysis of Multicast Throughput for SaN

Without loss of compatibility to most existing related results, we assume that $m_s = \Theta(m)$, i.e., $|\mathcal{S}| = \Theta(m)$. Please see the corresponding detailed proofs in Appendix C, available in the online supplemental material.

Above all, we should guarantee the priority of PhN in terms of the throughput. Then, we propose the following theorem.

Theorem 5. *By using Algorithm 1 to construct the multicast routing for SaN, denoted by $\mathbb{S}_s^r$, and the transmission scheme described in Fig. 3, denoted by $\mathbb{S}_s^t$, to schedule SaN, the capacity of highways in PhN, including FHs and SHs, can be protected from decreasing in order sense due to SaN.*

Because SaN does not add the load of any highway in PhN, by Theorem 5, we obtain that the presence of SaN has no impact on the order of throughput for PhN, when the strategy for SaN is designed as in Theorem 5.

Next, we study the throughput for SaN. First, we answer Question 3 proposed above. The meaning of $\mathbb{S}_s^r$ and $\mathbb{S}_s^t$ can be found in Theorem 5, so we have

Theorem 6. *Under the multicast routing $\mathbb{S}_s^r$ and transmission scheme $\mathbb{S}_s^t$, all multicast sessions in SaN can be served.*

Henceforth, we start to analyze the multicast throughput for SaN under $\mathbb{S}_s^r$ and $\mathbb{S}_s^t$, by using Theorem B.2 in Appendix B, available in the online supplemental material. We first compute the capacity of the FHs and SHs in SaN.

Theorem 7. *Under the transmission scheduling $\mathbb{S}_s^t$, the capacity of FHs and SHs in SaN can achieve $\Omega(1)$.*

According to Theorem 7, we can obtain Theorem 8.

Theorem 8. *During the first and second phases, when*
$$m_d = O\left(\frac{m}{(\log m)^2}\right),$$
the per-session multicast throughputs for SaN are achieved of $\Omega(\frac{1}{\sqrt{m m_d}})$ and $\Omega(\frac{1}{m_d} \cdot (\log m)^{-\frac{3}{2}})$, respectively.

Like in the single *random extended network* [8], [16], when the number of destinations is beyond some threshold, to be specific, $m_d = \Omega(\frac{m}{(\log m)^2})$, the multicast throughput derived by the multicast routing based on the FHs and SHs cooperatively is not optimal in order sense. For this case, the multicast routing based only on SHs can derive a larger throughput. Next, we describe such a routing scheme in Algorithm 2.

Algorithm 2. Multicast Routing based on Only SHs
Input: The multicast session $\mathcal{M}_{\mathcal{S},k}$ and $\mathrm{EST}(\mathcal{U}_{\mathcal{S},k})$.
Output: A multicast routing tree $\mathcal{T}(\mathcal{U}_{\mathcal{S},k})$.
1: **for** each link $u_i \to u_j$ of $\mathrm{EST}(\mathcal{U}_{\mathcal{S},k})$ **do**
2: According to the position of u_i and u_j, determine the indexes x_i and y_j, where $u_i \in s'_h(x_i)$; $u_j \in s'_v(y_j)$.
3: Packets are drained from u_i into the horizontal SH $\hbar'_h(x_i)$ by a single hop.
4: Packets are carried along the horizontal SH $\hbar'_h(x_i)$.
5: Packets are carried along the vertical SH $\hbar'_v(y_j)$.
6: Packets are delivered from the vertical SH $\hbar'_v(y_j)$ to u_j by a single hop.
7: **end for**
8: Using the similar procedure in Step 8 of Algorithm 1, we can obtain the final multicast routing tree $\mathcal{T}(\mathcal{U}_{\mathcal{S},k})$.

Theorem 9. *By using the multicast routing based only on SHs, i.e., the multicast routing constructed by Algorithm 2, and scheduling only for SHs, the per-session multicast throughput for SaN can be achieved of order*
$$\begin{cases} \Omega\left(\dfrac{1}{\sqrt{m_d m \log m}}\right) & \text{when} \quad m_d = O\left(\dfrac{m}{\log m}\right) \\ \Omega(1/m) & \text{when} \quad m_d = \Omega\left(\dfrac{m}{\log m}\right). \end{cases}$$

Then, according to Theorem B.2 in Appendix B, available in the online supplemental material, combining Theorem 8 and Theorem 9, we can get Theorem 2.

5.3 When PhN Adopts BS-Based Strategy

For this case, PhN adopts the classical BS-based strategy based on the *scheme lattice* $\mathbb{L}(\sqrt{n}, \sqrt{\frac{n}{b}}, 0)$, where $b = b(n)$ is the number of BSs in PhN. Under this strategy, the sources deliver the data to BSs during the Uplink phase and BSs deliver the received data to destinations during the Downlink phase. The communication between any pairs of PUs will be relayed by the BSs. To achieve a better throughput, the BS-based strategy is adopted in PhN only when $b = \Omega(n/\log n)$ [5]. Because BSs are regularly placed, i.e., each BS locates at the center of each cell, all cells can be simultaneously scheduled during both Uplink phase and Downlink phase, and can sustain a rate of $(\frac{b}{n})^{\frac{\alpha}{2}}$, [5]. In each cell, all downlinks and uplinks are scheduled in sequence. All *B-Type* PAs will be always *active*. The SUs contained in such PAs cannot be served. Thus, we study the multicast capacity for SaN under the general definition of multicast capacity, i.e., *asymptotic multicast capacity* (Definition B.2 in Appendix B, available in the online supplemental material). Obviously, the classic definition of capacity [6], [13], can be regarded as a special case of the *asymptotic multicast capacity*. Next, we focus on the Downlink phase in which SaN is scheduled. Whether or not SaN is scheduled in Uplink phase has no impact on the multicast throughput in order sense.

5.3.1 Highway System

SaN still prefers to adopt the multicast strategy based on the FHs and SHs as in the case that all BSs are always *inactive*. Also, the *B-Type* PA for each BS in the first phase is a cluster of nine cells in the scheme lattice $\mathbb{L}(\sqrt{n}, c\sqrt{\frac{n}{m}}, \frac{\pi}{4})$; and in the second phase it is a cluster of nine cells in

$$\mathbb{L}\left(\sqrt{n}, \sigma\sqrt{\log m \cdot \frac{n}{m}} - \epsilon_m, 0\right).$$

Notice that the effectiveness of such a *B-Type* PA relies closely on the fact that SaN is *dense scaling* while PhN is *extended scaling*. Intuitively, such FHs or SHs as in the case absent of *active* BSs might be damaged by the PAs as the number of BSs is increasing. Fortunately, for $b = O(n)$ then $b = o(\frac{m}{\log m})$, and BSs are regularly placed, [5], which guarantees SaN the same density, in order sense, of FHs as in the case absent of BSs. The previous SHs in SaN are still remained. However, when the packets carried along an SH are stopped by a PA, they will detour the PA along the adjacent SHs. Hence, the load of the SHs around the PAs is probably heavier than that of other SHs. See illustration in Fig. 4b. Throughout the routing, the bottleneck in the second phase should be in those SHs with heavy burden. We exploit this fact to analyze the multicast throughput for SaN.

5.3.2 Served Set of SaN

Now, we deal with the question that how many SUs are not served at all. Denote the set of all SUs that are not served by $\mathcal{P}$. Denote the set of all sources in SaN by $\mathcal{S}$. Based on the sets $\mathcal{P}$ and $\mathcal{S}$, we propose a definition of the *served set* of multicast sessions that can be divided into two regimes depending on m_d, i.e., the number of destinations of each multicast session.

Definition 1 (Served set). *The* served set, *denoted by* $\tilde{\mathcal{S}}$, *is a subset of* $\mathcal{S}$. *Define* $\tilde{\mathcal{S}} := \mathcal{S} - \mathcal{S} \cap \mathcal{P}$, *when* $m_d = \omega(\log m)$; *and define* $\tilde{\mathcal{S}} := \{v_{\mathcal{S},i} | \mathcal{U}_{\mathcal{S},i} \cap \mathcal{P} = \emptyset\}$, *when* $m_d = O(\log m)$.

For a multicast session $\mathcal{M}_{\mathcal{S},i}$, we define a subset of $\mathcal{U}_{\mathcal{S},i}$ as $\tilde{\mathcal{U}}_{\mathcal{S},i} = \{v_{\mathcal{S},i}\} \cup \mathcal{D}'_{\mathcal{S},i}$, where $\mathcal{D}'_{\mathcal{S},i} = \mathcal{D}_{\mathcal{S},i} - \mathcal{D}_{\mathcal{S},i} \cap \mathcal{P}$. Then, for $\mathcal{M}_{\mathcal{S},i}$, we build its spanning tree $\mathrm{EST}(\tilde{\mathcal{U}}_{\mathcal{S},i})$, and make it as the guideline of multicast routing, correspondingly acting as $EST(\mathcal{U}_{\mathcal{S},i})$ in the case that PhN adopts pure ad hoc strategy. Significantly, we have the following result.

Theorem 10. *As* $n \to \infty$, *it holds that* $|\tilde{\mathcal{S}}| \to |\mathcal{S}|$; *and for each* $v_{\mathcal{S},i} \in \tilde{\mathcal{S}}$, *it holds that uniform w.h.p.,* $|\mathcal{D}'_{\mathcal{S},i}| \to |\mathcal{D}_{\mathcal{S},i}|$.

Then, according to Theorem 10, the throughput derived by the strategy based on the *served set* is *asymp-achievable*.

5.3.3 Guarantee of Priority of PhN

In the first phase of SaN, the sum interference produced by SaN at a receiving PU in PhN, denoted by I_{sp}, is of order $O(1)$, due to the setting of PAs. Then, $I_{sp} = O(N_0)$, where the constant $N_0 > 0$ is the ambient noise. Hence, the presence of SaN does not change the order of the capacity of FHs in PhN. Similarly, we can prove that SaN does not impair the capacity of SHs in PhN.

5.3.4 Asymp-Achievable Multicast Throughput for SaN

We first consider the capacity of FHs and SHs in SaN.

Lemma 5. *In the first phase, the sum interference produced by PhN at a receiving SU is of order* $O(1)$.

Please see the detailed proof in Appendix C.10, available in the online supplemental material. According to the proof of Theorem 7 in Appendix C.7, available in the online supplemental material, during any time slot τ_1 in the first phase when a link $v_i \to v_j$ is scheduled, the interference on v_j produced by SaN itself is bounded by $I_{ss}(v_i, v_j; \tau_1) = O(1)$. Combining with Lemma 5, we get that the capacity of FHs in SaN does not decrease due to PhN, and it is still of order $\Omega(1)$. Using the similar procedure, we prove that the capacity of SHs in SaN is still of order $\Omega(1)$.

Next, we should analyze the load of FHs and SHs, respectively. The former is obviously the same as the load of FHs in SaN when BSs are absent. For the latter, the scheme of detouring increases the load of the SHs adjacent to PAs; but it can be proved that the increment does not change the order of the load of those SHs nonadjacent to the PAs. According to the technical Theorem B.1 in Appendix B, available in the online supplemental material, such questions come down to bounding the area of *sufficient regions* (Definition B.6 in Appendix B, available in the online supplemental material). Thus, it is the time to prove Theorem 3, as one of our main results. Please see the proof in Appendix C.11, available in the online supplemental material.

5.4 When PhN Adopts Hybrid Strategy

In this case, the strategy for PhN can be divided into four phases: FHs phase, SHs phase, Downlink phase, and Uplink phase [4], [5]. Then, we can use a simple *intermission method* to preserve the order of throughput and capacity for SaN as in the case that no BS is used, described in Theorem 2 and Theorem 4. That is, let SaN be idle during Downlink phase and Uplink phase, and schedule SaN in FHs phase and SHs phase as in the case that PhN adopts pure ad hoc strategy. Please see the illustration in Fig. 5.

6 Literature Review

In this section, we mainly review the literature about capacity scaling laws for cognitive networks.

In [3], the primary source destination and cognitive S-D pairs are modeled as an interference channel with

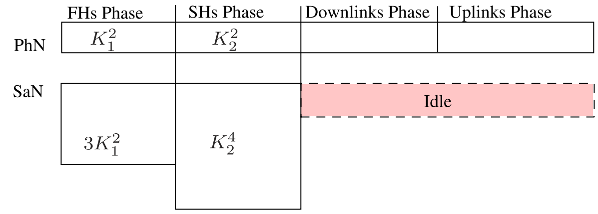


Fig. 5. Scheduling scheme in SaN when PhN adopts the hybrid strategy.

asymmetric side information. In [27], the communication opportunities are modeled as a two-switch channel. Note that both [3] and [27] only considered the *single-user* case in which a single primary and a single cognitive S-D pairs share the same spectrum. Recently, a *single-hop* cognitive network was considered in [28], where multiple secondary S-D pairs transmit in the presence of a single primary S-D pair. They showed that a linear scaling law of the *single-hop* secondary network is obtained when its operation is constrained to guarantee a particular outage constraint for the primary S-D pair. For multihop and multiple users case, Jeon et al. [17], [18] first studied the achievable unicast throughput for cognitive networks. In their cognitive model, the primary network is a random dense SANET or a dense BS-based network [5], and the secondary network is always a random dense SANET; two networks operate on the same space and spectrum. Following the model of [17], [18], Wang et al. [20] studied the multicast throughput for the primary and secondary networks. To ensure the priority of PUs in meanings of the throughput, they defined a new metric called *throughput decrement ratio* (TDR) to measure the ratio of the throughput of PaN in presence of SaN to that of PaN in absence of SaN. Endowing PaN with the right to determine the threshold of the TDR, they [20] devised the multicast strategies for SaN. Both the unicast routing in [17], [18] and multicast routing in [20] are built based on the backbones similar to the SHs in [8], which suggests that the derived throughputs are not optimal under the Gaussian Channel model for most cases. By introducing percolation-based routing [8], [14], Wang et al. [19] improved the multicast throughput for the same cognitive network model as in [17], [18], [20]; they showed that under some conditions, there exist the corresponding strategies to ensure both networks to achieve asymptotically the upper bounds of the capacity as they are stand-alone.

One of the common characteristics in [17], [18], [19], [20] is that the primary and secondary networks in all three models are dense scaling. More importantly, the common problem of three works is that all the strategies in [17], [18], [19], [20] shield the time domain, which makes the routing path always detour the PAs (or preservation regions), although they are sometimes *inactive*. Under those strategies, there are possibly some SUs that can never be served. As an important characteristic different from existing related works, our strategies allow a PA to be *dynamic* according to the state of the corresponding primary node. Thanks to the dynamics of PAs, SUs can access opportunistically into the spectrum from both time domain and space domain.

Huang and Wang [29] studied the throughput and delay scaling of general cognitive networks. They proposed a hybrid protocol model for secondary nodes to identify transmission opportunities. Based on it, they showed that secondary networks can obtain the same optimal performance as stand-alone networks when primary networks are some classical wireless networks. This work presented a fundamental insight on the architectural design of cognitive networks. Recently, Li et al. [30] studied the capacity and delay scaling laws for cognitive radio network with static PUs and heterogeneous mobile SUs coexist in the unit region. Liu et al. [31] investigated the scaling behavior of transmission delay in large scale ad hoc cognitive networks by analyzing the ratio of delay to distance as the distance goes to infinite.

7 Conclusion and Future Work

We study the multicast capacity for cognitive networks that operate under TDMA scheme. The network model consists of a PhN and a SaN. We devise the dynamic PA for each primary node according to the strategy adopted in PhN. Based on PAs, we design the multicast strategy for SaN under which the highway system acts as the multicast backbone. Under the precondition that SaN should have no negative impact on the order of the throughput for PhN, our strategy has the following merits: 1) By our strategy, the optimal throughput for SaN can be (asymptotically) achieved for some cases. 2) Under our strategy, unlike most related works, secondary nodes can access opportunistically into the spectrum from both time domain and space domain. 3) Under our strategy, all SUs can be served except for the case that PhN adopts BS-based strategy.

There are some future directions to be considered:

1. The case where PhN and SaN are both extended scaling may be studied.
2. An interesting and significant issue is to study the network model in which the secondary network is a mobile ad hoc network.
3. As for multicast capacity of stand-alone wireless ad hoc networks, the most challenging issue is to close the remaining gaps between the lower and upper bounds of multicast capacity in some regimes by establishing possibly tighter upper bounds or creating more effective strategies to improve the lower bounds.
4. In our system model, when SHs are scheduled, A-Type PAs are dynamic, then their statuses, i.e., the statuses of those centered primary nodes, are necessary for the corresponding secondary nodes. However, in a practical cognitive network, it is difficult for the secondary nodes to know the locations of primary receiving nodes. A more reasonable assumption is that SUs can locate the primary transmitters, [30], [32]. In some existing literatures, such as [32], some protocols, which only construct the preservation regions around the primary transmitters, are designed by amplifying the sizes of preservation regions. The effectiveness of this tricky technique depends on the fact that the length of each hop under the protocol in the primary network is limited to a certain order, thus, the distance between a secondary node and a primary

receiver can be limited by restricting the distance between this secondary node and the sender of this primary receiver. Intuitively, the logic is based on the triangle inequality. Our system model will be possibly improved by introducing this technique into our strategies.

5. In this paper, all our strategies are designed for the model where the primary and secondary networks are both of homogeneous ad hoc node density, and the results are derived based on the percolation theory for Poisson-distributed networks. When network models of inhomogeneous node density, such as those in [11], [12], are considered, the clustering behavior of users will affect the distribution of PAs and possibly invalidate the multicast strategies for the homogeneous model. Since the spatial inhomogeneity appears to be a quite ubiquitous feature of real network models, it is a significant future work to extend our work to the case where either the primary network or secondary network is of inhomogeneous node density.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their constructive comments. The research of authors is partially supported by the National Natural Science Foundation of China (NSFC) under grant No. 61202383, the National Basic Research Program of China (973 Program) under grant No. 2010CB328101, the Program for New Century Excellent Talents in University (NCET) under grant No. NCET-12-0414, the Natural Science Foundation of Shanghai under grant No. 12ZR1451200, the National Science Foundation for Postdoctoral Scientists of China under grant No. 2012M510118, the Research Fund for the Doctoral Program of Higher Education of China (RFDP) under grant No. 20120072120075, the Program for Changjiang Scholars and Innovative Research Team in University, the Shanghai Key Basic Research Project under grant No. 10DJ1400300, and the US National Science Foundation (NSF) CNS-1035894.

REFERENCES

[1] I.F. Akyildiz, W.-Y. Lee, M.C. Vuran, and S. Mohanty, "Next Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey," *Computer Networks*, vol. 50, pp. 2127-2159, 2006.

[2] C. Fortuna and M. Mohorcic, "Trends in the Development of Communication Networks: Cognitive Networks," *Computer Networks*, vol. 53, no. 9, pp. 1354-1376, 2009.

[3] N. Devroye, P. Mitran, and V. Tarokh, "Achievable Rates in Cognitive Radio Channels," *IEEE Trans. Information Theory*, vol. 52, no. 5, pp. 1813-1827, May 2006.

[4] B. Liu, P. Thiran, and D. Towsley, "Capacity of a Wireless Ad Hoc Network with Infrastructure," *Proc. ACM MobiHoc*, 2007.

[5] C. Wang, X.-Y. Li, C. Jiang, S. Tang, and Y. Liu, "Multicast Throughput of Hybrid Wireless Networks Under Gaussian Channel Model," *IEEE Trans. Mobile Computing*, vol. 10, no. 6, pp. 839-852, June 2011.

[6] X.-Y. Li, "Multicast Capacity of Wireless Ad Hoc Networks," *IEEE/ACM Trans. Networking*, vol. 17, no. 3, pp. 950-961, June 2009.

[7] A. Keshavarz-Haddad and R. Riedi, "Multicast Capacity of Large Homogeneous Multihop Wireless Networks," *Proc. IEEE Sixth Int'l Symp. Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks and Workshops (WiOpt)*, 2008.

[8] C. Wang, C. Jiang, X.-Y. Li, S. Tang, Y. He, X. Mao, and Y. Liu, "Scaling Laws of Multicast Capacity for Power-Constrained Wireless Networks Under Gaussian Channel Model," *IEEE Trans. Computers*, vol. 61, no. 5, pp. 713-725, May 2012.

[9] C. Hu, X. Wang, and F. Wu, "Motioncast: On the Capacity and Delay Tradeoffs," *Proc. ACM MobiHoc*, 2009.

[10] X. Mao, X.-Y. Li, and S. Tang, "Multicast Capacity for Hybrid Wireless Networks," *Proc. ACM MobiHoc*, 2008.

[11] G. Alfano, M. Garetto, and E. Leonardi, "Capacity Scaling of Wireless Networks with Inhomogeneous Node Density: Upper Bounds," *IEEE J. Selected Areas in Comm.*, vol. 27, no. 7, pp. 1147-1157, Sept. 2009.

[12] G. Alfano, M. Garetto, E. Leonardi, and V. Martina, "Capacity Scaling of Wireless Networks with Inhomogeneous Node Density: Lower Bounds," *IEEE/ACM Trans. Networking*, vol. 18, no. 5, pp. 1624-1636, Oct. 2010.

[13] P. Gupta and P.R. Kumar, "The Capacity of Wireless Networks," *IEEE Trans. Information Theory*, vol. 46, no. 2, pp. 388-404, Mar. 2000.

[14] M. Franceschetti, O. Dousse, D. Tse, and P. Thiran, "Closing the Gap in the Capacity of Wireless Networks via Percolation Theory," *IEEE Trans. Information Theory*, vol. 53, no. 3, pp. 1009-1018, Mar. 2007.

[15] A. ÖzgÜr, O. LÉvÊque, and D. Tse, "Hierarchical Cooperation Achieves Optimal Capacity Scaling in Ad Hoc Networks," *IEEE Trans. Information Theory*, vol. 53, no. 10, pp. 3549-3572, Oct. 2007.

[16] S. Li, Y. Liu, and X.-Y. Li, "Capacity of Large Scale Wireless Networks under Gaussian Channel Model," *Proc. ACM MobiCom*, 2008.

[17] S.-W. Jeon N. Devroye, M. Vu, S.-Y. Chung, and V. Tarokh, "Cognitive Networks Achieve Throughput Scaling of a Homogeneous Network," *Proc. IEEE Seventh Int'l Symp. Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, 2009.

[18] S.-W. Jeon, N. Devroye, M. Vu, S.-Y. Chung, and V. Tarokh, "Cognitive Networks Achieve Throughput Scaling of a Homogeneous Network," *IEEE Trans. Information Theory*, vol. 57, no. 8, pp. 5103-5115, Aug. 2011.

[19] C. Wang, S. Tang, X.-Y. Li, and C. Jiang, "Multicast Capacity Scaling Laws for Multihop Cognitive Networks," *IEEE Trans. Mobile Computing*, vol. 11, no. 11, pp. 1627-1639, Nov. 2012.

[20] C. Wang, C. Jiang, X.-Y. Li, and Y. Liu, "Multicast Throughput for Large Scale Cognitive Networks," *ACM/Springer Wireless Networks*, vol. 16, no. 7, pp. 1945-1960, 2010.

[21] A. Keshavarz-Haddad and R. Riedi, "Bounds for the Capacity of Wireless Multihop Networks Imposed by Topology and Demand," *Proc. ACM MobiHoc*, 2007.

[22] M. Vu, N. Devroye, and V. Tarokh, "On the Primary Exclusive Regions in Cognitive Networks," *IEEE Trans. Wireless Comm.*, vol. 8, no. 7, pp. 3380-3385, July 2009.

[23] A. Agarwal and P. Kumar, "Capacity Bounds for Ad Hoc and Hybrid Wireless Networks," *ACM SIGCOMM Computer Comm. Rev.*, vol. 34, no. 3, pp. 71-81, 2004.

[24] R. Zheng, "Information Dissemination in Power-Constrained Wireless Networks," *Proc. IEEE INFOCOM*, 2006.

[25] O. Dousse and P. Thiran, "Connectivity vs Capacity in Dense Ad Hoc Networks," *Proc. IEEE INFOCOM*, 2004.

[26] A. ÖzgÜr, R. Johari, D. Tse, and O. Lévque, "Information-Theoretic Operating Regimes of Large Wireless Networks," *IEEE Trans. Information Theory*, vol. 56, no. 1, pp. 427-437, Jan. 2010.

[27] S. Jafar and S. Srinivasa, "Capacity Limits of Cognitive Radio with Distributed and Dynamic Spectral Activity," *IEEE Trans. Computers*, vol. 25, no. 5, pp. 529-537, Apr. 2007.

[28] M. Vu and V. Tarokh, "Scaling Laws of Single-Hop Cognitive Networks," *IEEE Trans. Wireless Comm.*, vol. 8, no. 8, pp. 4089-4097, Aug. 2009.

[29] W. Huang and X. Wang, "Throughput and Delay Scaling of General Cognitive Networks," *Proc. IEEE INFOCOM*, 2011.

[30] Y. Li, X. Wang, X. Tian, and X. Liu, "Scaling Laws for Cognitive Radio Network with Heterogeneous Mobile Secondary Users," *Proc. IEEE INFOCOM*, 2012.

[31] Z. Liu, X. Wang, W. Luan, and S. Lu, "Transmission Delay in Large Scale Ad Hoc Cognitive Radio Networks," *Proc. ACM MobiHoc*, 2012.

[32] C. Yin, L. Gao, and S. Cui, "Scaling Laws for Overlaid Wireless Networks: A Cognitive Radio Network versus a Primary Network," *IEEE/ACM Trans. Networking*, vol. 18, no. 4, pp. 1317-1329, Aug. 2010.

Cheng Wang received the PhD degree in the Department of Computer Science at Tongji University, in 2011. Currently, he is an associate research professor of computer science at Tongji University. His research interests include wireless communications and networking, mobile social networks, and mobile cloud computing.

Changjun Jiang received the PhD degree from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 1995. Currently, he is a professor with the Department of Computer Science and Engineering, Tongji University, Shanghai. He is also a council member of China Automation Federation and Artificial Intelligence Federation, the director of Professional Committee of Petri Net of China Computer Federation, the vice director of Professional Committee of Management Systems of China Automation Federation, and an information area specialist of Shanghai Municipal Government. His current areas of research are concurrent theory, Petri net, and formal verification of software, wireless networks, concurrency processing, and intelligent transportation systems.

Shaojie Tang received the BS degree in radio engineering from Southeast University, China, in 2006, and has been working toward the PhD degree in the Computer Science Department at the Illinois Institute of Technology since 2006. His current research interests include algorithm design and analysis for wireless ad hoc networks, wireless sensor networks, and online social networks.

Xiang-Yang Li (M'99, SM'08) received the bachelor's degree at the Department of Computer Science and the bachelor's degree at the Department of Business Management from Tsinghua University, China, both in 1995, and the MS and PhD degrees in 2000 and 2001 at the Department of Computer Science from the University of Illinois at Urbana-Champaign. He is a professor of computer science at the Illinois Institute of Technology. He serves as an editor of several journals, including *IEEE TPDS* and *Networks*. He also serves in Advisory Board of *Ad Hoc & Sensor Wireless Networks* from 2005, and *IEEE CN* from 2011. He is a senior member of the IEEE.