On the Connectivity and Multihop Delay of Ad Hoc Cognitive Radio Networks

Wei Ren, Qing Zhao, Ananthram Swami

Abstract

We analyze the multihop delay of ad hoc cognitive radio networks, where the transmission delay of each hop consists of the propagation delay and the waiting time for the availability of the communication channel (*i.e.*, the occurrence of a spectrum opportunity at this hop). Using theories and techniques from continuum percolation and ergodicity, we establish the scaling law of the minimum multihop delay with respect to the source-destination distance in cognitive radio networks. We show the starkly different scaling behavior of the multihop delay in *connected* networks as compared to networks that are only *intermittently connected* due to scarcity of spectrum opportunities.

Index Terms

Cognitive radio network, multihop delay, connectivity, intermittent connectivity, continuum percolation, ergodic theory.

I. Introduction

The basic idea of opportunistic spectrum access is to achieve spectrum efficiency and interoperability through a hierarchical access structure with primary and secondary users [1]. Secondary users, equipped with cognitive radios [2] capable of sensing and learning the communication environment, identify and exploit instantaneous and local spectrum opportunities without causing unacceptable interference to primary users [1].

Using theories and techniques from continuum percolation and ergodicity, we analytically characterize the connectivity and multihop delay of the secondary network. Specifically, we

[0]This work was supported in part by the Army Research Laboratory under Grant DAAD19-01-C-0062, by the Army Research Office under Grant W911NF-08-1-0467, and by the National Science Foundation under Grants ECS-0622200 and CCF-0830685.

consider a Poisson distributed secondary network[1] overlaid with a Poisson distributed primary network in an infinite two-dimensional Euclidean space[2]. Due to the hierarchical structure of spectrum sharing, the transmission delay of each hop in the secondary network consists of two components: the propagation delay and the waiting time for the availability of the communication channel (*i.e.*, the occurrence of a spectrum opportunity at this hop).

A. Main Results

First, we analytically characterize the connectivity of the secondary network which depends on not only the topology of the secondary network but also the transmitting and receiving activities of the privileged primary network. The connectivity of the secondary network is thus determined by two critical parameters: the density λ_S of the secondary users and the density λ_{PT} of the primary transmitters representing the traffic load of the primary network. As illustrated in Fig. 1, we show that according to the connectivity of the secondary network, the $(\lambda_S, \lambda_{PT})$ plane can be partitioned into three regions: disconnected, connected, and intermittently connected, which are all interpreted in the percolation sense. The secondary network is disconnected if there does not exist almost surely (a.s.) an infinite connected component formed by topological links connecting two secondary users within each other's transmission range, and the secondary network is connected if there exists a.s. an infinite connected component formed by communication links, where communication links are those topological links experiencing the spectrum opportunities. Since the set of communication links is a subset of topological links, we define the intermittent connectivity for a secondary network that is not connected as the a.s. existence of an infinite connected component consisting of topological links. The above three concepts are detailed in Sec. III.

Second, we establish the scaling law of the minimum multihop delay in the secondary network with respect to the source-destination distance. When the secondary network is disconnected, there exist only finite topologically connected components. If we randomly choose two secondary users, then they belong to two different topologically connected component a.s., which implies

[1]The notions of cognitive radio networks and secondary networks are used interchangeably in this paper.

[2]This infinite network model is equivalent in distribution to the limit of a sequence of finite networks with a fixed density as the area of the network increases to infinity, *i.e.*, the so-called *extended network* [3].

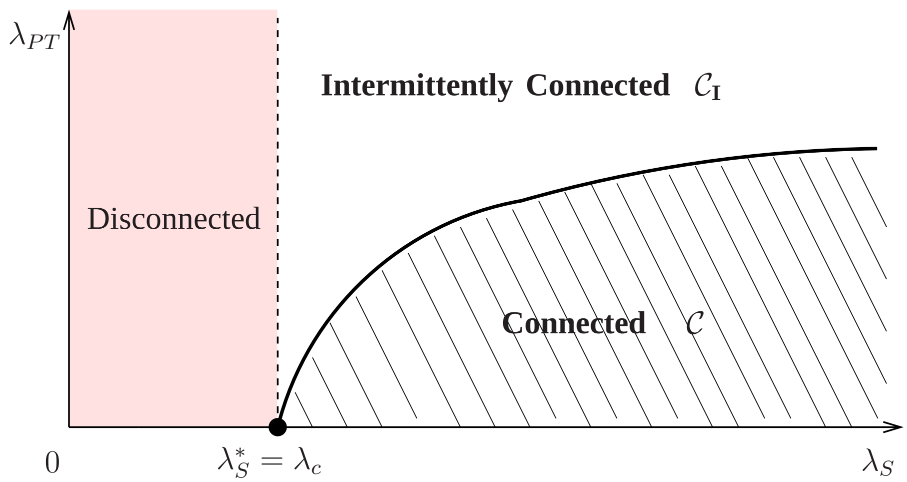

Fig. 1. The shaded region is the set of $(\lambda_S, \lambda_{PT})$ under which the secondary network is connected (*i.e.*, the connectivity region $\mathcal{C}$), the white region is the set of $(\lambda_S, \lambda_{PT})$ under which the secondary network is intermittently connected (*i.e.*, the intermittent connectivity region $\mathcal{C}_I$), and the colored region is the set of $(\lambda_S, \lambda_{PT})$ under which the secondary network is disconnected. The critical density λ_S^* of the secondary users is defined as the infimum density of the secondary users to ensure connectivity under a *positive* density of the primary transmitters, which is equal to the critical density λ_c of a homogeneous network.

that they are not reachable from each other. We thus focus on the multihop delay between two secondary users in the infinite topologically connected component when the secondary network is either connected or intermittently connected.

To highlight the impact of the waiting time for spectrum opportunities on multihop delay, we first study the scaling law of the minimum multihop delay assuming that the propagation delay is negligible. We show that the scaling law of the minimum multihop delay with respect to the source-destination distance has two distinct regimes, corresponding to whether the secondary network is connected or intermittently connected. Specifically, let μ be the source, ν the destination, $t(\mu,\nu)$ the minimum multihop delay from μ to ν, and $d(\mu,\nu)$ the distance between μ and ν, then we show that, a.s.

$$\lim_{d(\mu,\nu)\to\infty} \frac{t(\mu,\nu)}{d(\mu,\nu)} \begin{cases} = 0, & \text{if connected;} \\ > 0, & \text{if intermittently connected.} \end{cases}$$

When the secondary network is connected, a much stronger statement is actually shown, that is,

$$\lim_{d(\mu,\nu)\to\infty} \frac{t(\mu,\nu)}{g(d(\mu,\nu))} = 0 \text{ a.s.},$$

where $g(d(\mu,\nu))$ is any monotonically increasing function of $d(\mu,\nu)$ satisfying $\lim\limits_{d(\mu,\nu)\to\infty} g(d(\mu,\nu)) = \infty$. It implies that the minimum multihop delay $t(\mu,\nu)$ is asymptotically independent of the distance $d(\mu, \nu)$ as $d(\mu, \nu) \to \infty$. Thus when the propagation delay is negligible, a connected cognitive radio (CR) network behaves almost the same as a homogeneous ad hoc network, in the sense that the waiting time for the spectrum opportunities does not affect the scaling law of the multihop delay with respect to the source-destination distance.

The above scaling law of the multihop delay may be illustrated with an analogy of traveling from a place μ to another place ν, where the waiting time for the spectrum opportunities is likened to the waiting time for traffic lights. Suppose that we can move fast enough such that (s.t.) the driving time on the road is negligible. When the secondary network is connected, there exists an infinite connected component consisting of communication links a.s. which can be regarded as a highway without traffic lights between μ and ν. Given that both μ and ν are within a finite distance to the highway (independent of the distance between μ and ν), the traveling time from μ to ν, which is exactly the waiting time for traffic lights before entering the highway and after leaving the highway, is independent of the distance between μ and ν. When the secondary network is intermittently connected, there does not exist an infinite connected component formed by communication links a.s., *i.e.*, such a highway between μ and ν can not be found. Then we have to use local paths and wait for traffic lights from time to time, leading to the linear scaling of the traveling time with respect to the distance between μ and ν.

We also study the impact of the propagation delay on multihop delay. When the propagation delay τ is nonnegligible, we show that the minimum multihop delay scales linearly with the source-destination distance in both connected and intermittently connected regimes, but with different rates for the linear scaling. In particular, the limiting behavior of the rate as $\tau \to 0$ is distinct in the two regimes, *i.e.*, a.s.

$$\lim_{\tau \to 0} \lim_{d(\mu,\nu)\to\infty} \frac{t(\mu,\nu)}{d(\mu,\nu)} \begin{cases} = 0, & \text{if connected;} \\ > 0, & \text{if intermittently connected.} \end{cases}$$

It indicates that when the propagation delay is sufficiently small, the scaling rate of the multihop delay for a connected network is much smaller than the one for a intermittently connected network.

B. Related Work

As a fundamental issue for the feasibility and efficiency of large-scale wireless networks, the scaling law has raised increasing interest in the research community since the seminal work of

P. Gupta and P. R. Kumar [4]. The capacity scaling law of CR networks has been analyzed in [5–7]. In [5], the authors also derive the capacity-delay tradeoff as the density of users within a unit square tends to infinity (the so-called *dense network*) for a routing and scheduling algorithm proposed by them, which is shown to achieve the optimal capacity-delay tradeoff for homogeneous ad hoc networks. To our best knowledge, the scaling law of the multihop delay with respect to the source-destination distance in a CR network has not been characterized analytically or experimentally in the literature.

The scaling law of the multihop delay in homogeneous ad hoc networks has been well studied in [8–16]. As the number of users in the network increases to infinity, the multihop delay for a specific routing algorithm is analyzed in [8–10], and the capacity-delay tradeoff is revealed under a given network and mobility model in [11–13]. Based on continuum percolation theory, the scaling law of the multihop delay with respect to the source-destination distance is established in [14–16]. This work shares similarity with [16], which considers homogeneous ad hoc networks with dynamical on-off links and shows that the scaling of the minimum multihop delay behaves distinctly in two regimes, depending on whether the network is percolated. A major difference of this work from [16] is that the states of the links in the secondary network are correlated instead of independent, which complicates the multihop delay analysis.

II. Network Model

We consider a Poisson distributed secondary network overlaid with a Poisson distributed primary network in an infinite two dimensional Euclidean space. The primary network adopts a synchronized slotted structure with a slot length T_S. The realizations of active primary transmitters vary from slot to slot and are assumed to be i.i.d. across slots[3]. Thus T_S can be considered as the time constant of the spectrum opportunities which are determined by the transmitting and receiving activities of the primary users. Without loss of generality, we assume that $T_S = 1$.

At the beginning of each slot, the primary transmitters are distributed according to a two-dimensional Poisson point process X_{PT} with density λ_{PT}. To each primary transmitter, its receiver is uniformly distributed within its transmission range R_p. Here we have assumed that all

[3]The different realizations of active primary transmitters in different slots can be caused by the mobility of these users or changes in the traffic pattern or both.

the primary transmitters use the same transmission power and the transmitted signals undergo an isotropic path loss. Based on the displacement theorem [17, Chapter 5], it is easy to see that the primary receivers form another two-dimensional Poisson point process X_{PR} with density λ_{PT}. Note that the two Poisson processes X_{PT} and X_{PR} are correlated.

The secondary users are distributed according to a two-dimensional Poisson point process X_S with density λ_S, independent of X_{PT} and X_{PR}. The locations of the secondary users are static over time. Based on the scaling argument [18, Chapter 2], we can set the transmission range r_p of the secondary users to 1 without loss of generality.

III. Connectivity vs. Intermittent Connectivity

A secondary network is disconnected if there does not exist an infinite connected component formed by topological links, where a topological link exists between two secondary users if they are within the transmission range of each other. Notice that this condition for the existence of a topological link is equivalent to the one for the existence of a communication link in homogeneous ad hoc networks. As discussed in [18, chapter 3], the connectivity of homogeneous networks, which is defined as the a.s. existence of an infinite connected component, is uniquely determined by its density. Thus, the secondary network with density λ_S is disconnected if and only if $\lambda_S \leq \lambda_c$, where λ_c is the critical density of homogeneous networks.

Due to the hierarchical structure of spectrum sharing, besides the density λ_S of the secondary users, the density λ_{PT} of the primary transmitters affects the connectivity of the secondary network. Specifically, in contrast to the case in homogeneous ad hoc networks, the existence of a communication link between two secondary users depends on not only the distance between them (at most r_p) but also the availability of the communication channel (*i.e.*, the presence of a spectrum opportunity). The latter is determined by the transmitting and receiving activities of the primary network as described below.

A. Spectrum Opportunity

As illustrated in Fig. 2, where we consider the disk signal propagation and interference model, there exists an opportunity from μ, the secondary transmitter, to ν, the secondary receiver, if the transmission from μ does not interfere with nearby *primary receivers* in the solid circle, and the reception at ν is not affected by nearby *primary transmitters* in the dashed circle [19]. Referred to as the interference range of the secondary users, the radius r_I of the solid circle at μ depends on the transmission power of μ and the interference tolerance of the primary receivers, whereas the radius R_I of the dashed circle (the interference range of the primary users) depends on the transmission power of the primary users and the interference tolerance of ν.

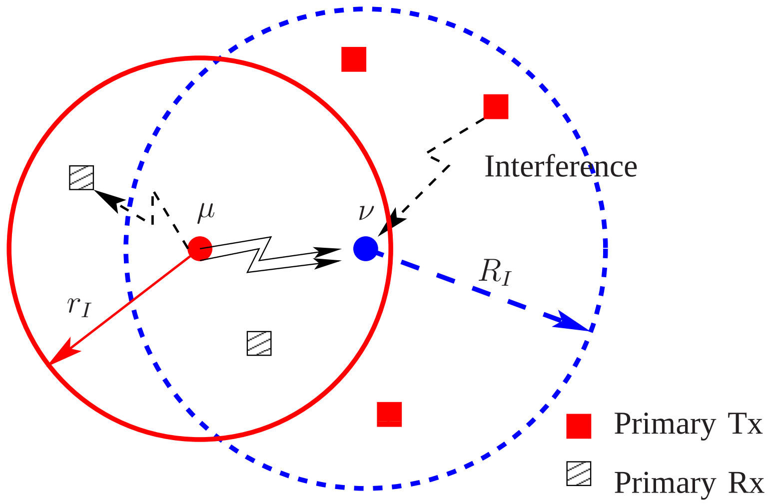

Fig. 2. Definition of spectrum opportunity.

It is clear from the above discussion that spectrum opportunities are *asymmetric*. Specifically, a channel that is an opportunity when μ is the transmitter and ν the receiver may not be an opportunity when ν is the transmitter and μ the receiver. We consider applications with guaranteed delivery where acknowledgements are required to complete communications. Hence, bidirectional spectrum opportunities between μ and ν are needed. As a result, the single-hop transmission delay $t_s(\mu,\nu)$ from μ to ν is the waiting time $t_{sw}(\mu,\nu)$ for the presence of the first bidirectional opportunity plus the propagation delay τ. Assume that $\tau \leq T_S = 1$ s.t. the spectrum opportunity lasts long enough to ensure the success of the transmission and the secondary transmitter intends to transmit the packet at time 0. Let $\mathbb{I}_{(\mu,\nu)}(n)$ be an indicator s.t. $\mathbb{I}_{(\mu,\nu)}(n) = 1$ if an bidirectional opportunity exists between μ and ν during the nth primary slot, and $\mathbb{I}_{(\mu,\nu)}(t) = 0$ otherwise, then we have

$$\begin{aligned} t_s(\mu,\nu) &= t_{sw}(\mu,\nu) + \tau \\ &= \underset{n\in\{0,1,2,\ldots\}}{\arg\min} \{\mathbb{I}_{(\mu,\nu)}(n) = 1\} + \tau. \end{aligned}$$

Due to the i.i.d. distribution of the primary network over slots, $t_{sw}(\mu, \nu)$ is obviously a geometric random variable with parameter p_0, where p_0 is the probability of having a bidirectional opportunity between μ and ν at any given time. As shown in Proposition 1 (see Appendix A), given the parameters of both the primary and the secondary network ($i.e.,$ λ_{PT}, R_p, R_I, r_p, and r_I are fixed), the probability p_0 of the bidirectional opportunity is a function of the distance $d(\mu, \nu)$ between μ and ν. Consequently, the distribution of $t_{sw}(\mu, \nu)$ depends on $d(\mu, \nu)$.

B. Connectivity and Connectivity Region

In a primary slot t, based on the conditions for the existence of a communication link, we can obtain an undirected random graph $\mathcal{G}_H(\lambda_S, \lambda_{PT}, t)$ which represents the connectivity of the secondary network in this slot. As illustrated in Fig. 3, this graph $\mathcal{G}_H(\lambda_S, \lambda_{PT}, t)$ is determined by the three Poisson point processes in slot t: X_S, X_{PT}, and X_{PR}, where X_{PT} and X_{PR} are correlated.

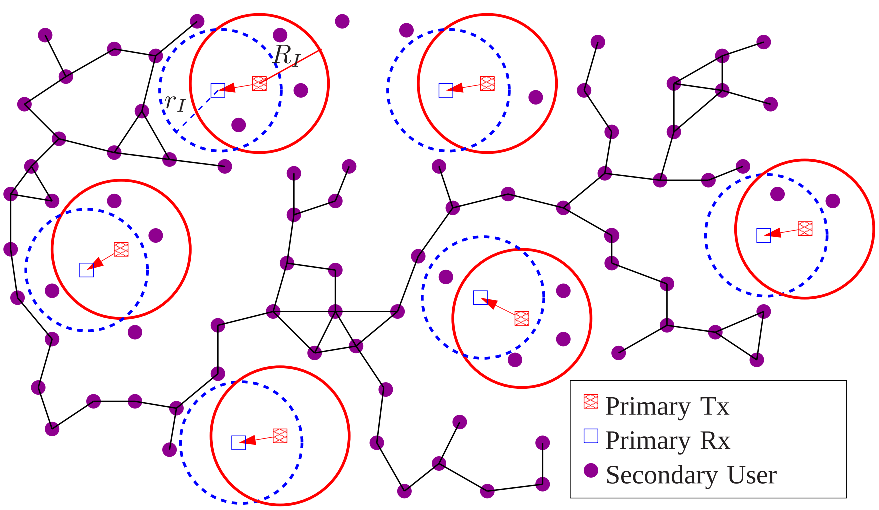

Fig. 3. A realization of the random graph $\mathcal{G}_H(\lambda_S, \lambda_{PT}, t)$ which consists of all the secondary users and all the communication links in the primary slot t (denoted by solid lines). The solid circles denote the interference regions of the primary transmitters within which secondary users can not successfully receive, and the dashed circles denote the required protection regions for the primary receivers within which secondary users should refrain from transmitting.

We define the connectivity of the secondary network as the a.s. existence of an infinite connected component in $\mathcal{G}_H(\lambda_S, \lambda_{PT}, t)$ for all t. Given the transmission power and the interference tolerance of both the primary and the secondary users ($i.e.,$ R_p, R_I, r_p, and r_I are fixed), the connectivity region $\mathcal{C}$ is defined as

$$\mathcal{C} \stackrel{\Delta}{=} \{(\lambda_S,\ \lambda_{PT}):\ \mathcal{G}_H(\lambda_S, \lambda_{PT}, t) \text{ is connected for all } t\}.$$

A detailed analytical characterization of $\mathcal{C}$ is given in [20, 21].

Referred to as the critical density of the secondary users, λ_S^* is the infimum density of the secondary users to ensure connectivity under a positive density of active primary transmitters:

$$\lambda_S^* \stackrel{\Delta}{=} \inf\{\lambda_S : \exists \lambda_{PT} > 0 \text{ s.t. } \mathcal{G}_H(\lambda_S, \lambda_{PT}, t) \text{ is connected for all } t\}.$$

It is shown in [20, 21] that λ_S^* equals the critical density λ_c of a *homogeneous* ad hoc network.

Let $\theta(\lambda_S, \lambda_{PT})$ denote the probability that an arbitrary secondary user is connected to an infinite connected component[4] in $\mathcal{G}_H(\lambda_S, \lambda_{PT}, t)$, then we have that

$$\theta(\lambda_S, \lambda_{PT}) \begin{cases} > 0, & \text{if } (\lambda_S, \lambda_{PT}) \in \mathcal{C}; \\ = 0, & \text{otherwise.} \end{cases} \tag{1}$$

C. Intermittent Connectivity

By connecting two secondary users which are within the transmission range of each other via a topological link, we derive an undirected random graph $\mathcal{G}_S(\lambda_S)$ which depends only on the Poisson point process X_S of the secondary network. Similarly, we define the connectivity of $\mathcal{G}_S(\lambda_S)$ as the a.s. existence of an infinite connected component in it. It follows from the classic result on homogeneous networks [18, Chapter 3] that $\mathcal{G}_S(\lambda_S)$ is connected if and only if $\lambda_S > \lambda_c$, where λ_c is the critical density of homogeneous networks.

$\mathcal{G}_S(\lambda_S)$ can also be obtained by adding topological links that do not see the opportunities in slot t to the random graph $\mathcal{G}_H(\lambda_S, \lambda_{PT}, t)$. Thus, even if the secondary network is not connected, it is still possible that $\mathcal{G}_S(\lambda_S)$ is connected. On the other hand, given a connected $\mathcal{G}_S(\lambda_S)$, there may not be enough topological links which experience the opportunities *simultaneously* to form an infinite connected component in $\mathcal{G}_H(\lambda_S, \lambda_{PT}, t)$, but for any two secondary users μ and ν in the infinite connected component of $\mathcal{G}_S(\lambda_S)$, packets from μ can reach ν along a path in $\mathcal{G}_S(\lambda_S)$ with a finite multihop delay. In this case, although the transmission of the packets may not be completed in one primary slot, the packets can stay whenever the absence of the spectrum opportunities blocks their transmission and wait for some finite time to get through. The finiteness

[4]Since the distribution of the primary network is i.i.d. over slots, it is easy to see that this probability θ is invariant of time.

of the waiting time is guaranteed by the following two facts: (i) for each topological link in the secondary network, the probability of the spectrum opportunity is strictly positive no matter how large the density of the primary transmitters is (see Proposition 1 in Appendix A); (ii) the spectrum opportunities are time-varying due to the i.i.d. distribution of the primary network across slots. We thus define the intermittent connectivity for a secondary network that is not connected as the connectivity of $\mathcal{G}_S(\lambda_S)$. We also define the intermittent connectivity region $\mathcal{C}_I$ as

$$\mathcal{C}_I \stackrel{\Delta}{=} \{(\lambda_S,\ \lambda_{PT}) \notin \mathcal{C}:\ \ \lambda_S > \lambda_c\}.$$

IV. Multihop Delay

In this section, we analytically characterize the asymptotic behavior of the minimum multihop delay as the source-destination distance tends to infinity. Let $C(\mathcal{G}_S(\lambda_S))$ be the infinite connected component in $\mathcal{G}_S(\lambda_S)$ when $\lambda_S > \lambda_c$, *i.e.*, the secondary network is either connected or intermittently connected. The question we aim to answer here is the scaling law of the minimum multihop delay between two arbitrary users in $C(\mathcal{G}_S(\lambda_S))$ with respect to the distance between them. As shown in the following two theorems which consider the two cases when the propagation delay $\tau = 0$ and $\tau > 0$, the connectivity of the secondary network determines the scaling law of the minimum multihop delay, where the highway provided by the infinite connected component consisting of communication links plays an indispensable role.

Theorem 1: Assume that $\tau = 0$. For any two secondary users $\mu,\ \nu\ \in C(\mathcal{G}_S(\lambda_S))$, where $C(\mathcal{G}_S(\lambda_S))$ is the infinite connected component of $\mathcal{G}_S(\lambda_S)$, let $t(\mu,\nu)$ denote the minimum multihop delay from μ to ν and $d(\mu,\nu)$ the distance between μ and ν, then

T1.1 if $(\lambda_S, \lambda_{PT}) \in \mathcal{C}$,

$$\lim_{d(\mu,\nu)\to\infty} \frac{t(\mu,\nu)}{g(d(\mu,\nu))} = 0 \text{ a.s.},$$

where $g(d)$ is any monotonically increasing function of d with $\lim_{d\to\infty} g(d) = \infty$;

T1.2 if $(\lambda_S, \lambda_{PT}) \in \mathcal{C}_I$, $\exists\ 0 < \beta < \infty$ s.t.

$$\lim_{d(\mu,\nu)\to\infty} \frac{t(\mu,\nu)}{d(\mu,\nu)} = \beta \text{ a.s.}, \tag{2}$$

where the value of β depends on $(\lambda_S, \lambda_{PT})$.

Proof Sketch: To show T1.1, we use the infinite connected component[5] in $\mathcal{G}_H(\lambda_S, \lambda_{PT}, t_0)$ during some primary slot t_0 that contains μ to construct such a path from μ to ν that the multihop delay along this path is independent of the distance $d(\mu, \nu)$. To show T1.2, we first prove the existence of $\lim\limits_{d(\mu,\nu)\to\infty} \frac{t(\mu,\nu)}{d(\mu,\nu)}$ based on the Subadditive Ergodic Theorem [22] and then derive a lower bound on $\frac{t(\mu,\nu)}{d(\mu,\nu)}$ by considering the fact that the message from μ can traverse only a finite distance towards ν during each primary slot. For details, see Sec. V-A and Sec. V-B. ■

Theorem 2: Assume that $\tau > 0$. For any two secondary users $\mu,\ \nu\ \in C(\mathcal{G}_S(\lambda_S))$, where $C(\mathcal{G}_S(\lambda_S))$ is the infinite connected component of $\mathcal{G}_S(\lambda_S)$, let $t^\tau(\mu, \nu)$ denote the minimum multihop delay from μ to ν and $d(\mu, \nu)$ the distance between μ and ν, then $\exists\ \gamma = \gamma(\tau) > 0$ s.t.

$$\lim_{d(\mu,\nu)\to\infty} \frac{t^\tau(\mu,\nu)}{d(\mu,\nu)} = \gamma \geq \tau \text{ a.s..} \tag{3}$$

Furthermore, if $(\lambda_S, \lambda_{PT}) \in \mathcal{C}$,

$$\lim_{\tau\to 0} \lim_{d(\mu,\nu)\to\infty} \frac{t^\tau(\mu,\nu)}{d(\mu,\nu)} = 0 \text{ a.s.;} \tag{4}$$

if $(\lambda_S, \lambda_{PT}) \in \mathcal{C}_I$,

$$\lim_{\tau\to 0} \lim_{d(\mu,\nu)\to\infty} \frac{t^\tau(\mu,\nu)}{d(\mu,\nu)} \geq \beta > 0 \text{ a.s.,} \tag{5}$$

where $\beta = \lim\limits_{d(\mu,\nu)\to\infty} \frac{t(\mu,\nu)}{d(\mu,\nu)}$ is defined in (2).

Proof Sketch: The equality in (3) is based on the Subadditive Ergodic Theorem [22], while the inequality in (3) is established via a simple lower bound on $t^\tau(\mu, \nu)$. The basic idea behind establishing (4) is to consider the multihop delay along the path constructed in the proof of T1.1. Eqn. (5) follows immediately from the fact that $t^\tau(\mu, \nu) \geq t(\mu, \nu)$, where $t(\mu, \nu)$ is the minimum multihop delay when $\tau = 0$. For details, see Sec. V-C. ■

V. Proofs

A. Proof of Theorem 1.1

Let μ be the source and ν the destination, then we construct a specific path L_C from μ to ν which makes use of the infinite connected component in $\mathcal{G}_H(\lambda_S, \lambda_{PT}, t)$ and analyze the

[5]It is shown in [20, 21] that there exists either zero or one infinite connected component in $\mathcal{G}_H(\lambda_S, \lambda_{PT}, t)$ a.s. for any given t.

multihop delay $t^C(\mu,\nu)$ along L_C which provides an upper bound on the minimum multihop delay $t(\mu,\nu)$.

Assume that μ starts trying to send the message at time $t=0$. Since $(\lambda_S,\lambda_{PT})\in\mathcal{C}$, there exists a unique infinite connected component in $\mathcal{G}_H(\lambda_S,\lambda_{PT},t)$ for all t, which is denoted by $C(t)$. Let t' be the first primary slot such that $\mu\in C(t')$. Based on (1), we know that the probability $\theta(\lambda_S,\lambda_{PT})$ that $\mu\in C(t)$ for each t is strictly positive. It follows from the i.i.d. distribution of the primary network across slots that t' is finite a.s. Given $C(t')$, we define user w_ν as the user in $C(t')$ which is closest to ν, *i.e.*,

$$w_\nu \triangleq \arg\min_{w_i\in C(t')} d(w_i,\nu).$$

Notice that if $\nu\in C(t')$, then $w_\nu=\nu$.

As illustrated in Fig. 4, let L_C be the path from μ to ν which passes through w_ν, then the minimum multihop delay $t^C(\mu,\nu)$ along the path L_C can be expressed as:

$$\begin{aligned} t^C(\mu,\nu) &= t' + t(\mu,w_\nu) + t(w_\nu,\nu) \\ &= t' + t(w_\nu,\nu). \end{aligned}$$

In the last step, we have used $t(\mu,w_\nu)=0$, since μ, $w_\nu \in C(t')$ and $\tau=0$. It is easy to see that t' and $t(w_\nu,\nu)$ are independent of $d(\mu,\nu)$. Now it suffices to show that $t(w_\nu,\nu)$ is finite a.s.

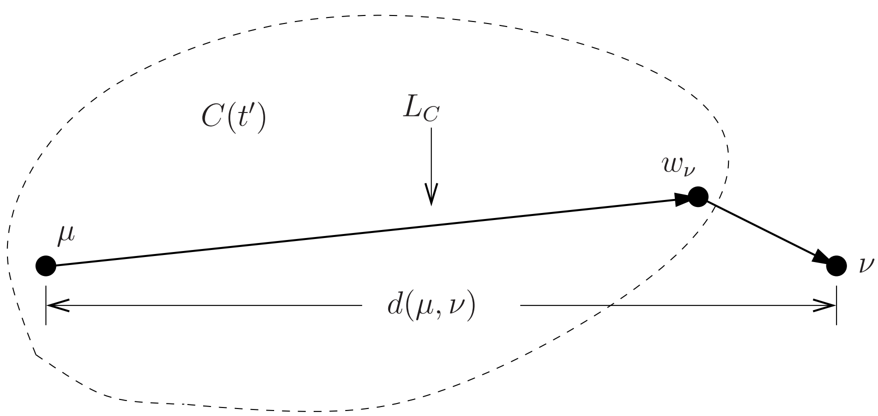

Fig. 4. An illustration of the constructed path L_C from μ to ν when $(\lambda_S,\lambda_{PT})\in\mathcal{C}$. $C(t')$ is the infinite connected component of $\mathcal{G}(\lambda_S,\lambda_{PT},t')$ which first contains μ, and w_ν is the user in $C(t')$ which is closest to ν.

Lemma 1: $t(w_\nu,\nu)$ is finite a.s.

Proof of Lemma 1: We first show that $d(w_\nu, \nu) < \infty$ a.s., and then we obtain an upper bound on the multihop delay along the shortest path $L(w_\nu, \nu)$ (in terms of the number of hops) from w_ν to ν. Notice that the multihop delay $t^L(w_\nu, \nu)$ along $L(w_\nu, \nu)$ is not necessarily the minimum multihop delay, and it only gives an upper bound on the minimum multihop delay. In other words, $t(w_\nu, \nu) \leq t^L(w_\nu, \nu)$. The proof here is inspired by the proof of Lemma 9 in [16], but with a much simpler proof of $d(w_\nu, \nu) < \infty$.

Since $d(w_\nu, \nu) \leq d(w_\nu, (0,0)) + d(\nu, (0,0))$ and $d(\nu, (0,0)) < \infty$ a.s., it follows that $d(w_\nu, \nu) < \infty$ a.s. if $d(w_\nu, (0,0)) < \infty$ a.s. Consider the following three events:

$$\begin{aligned} E &= \{d(w_\nu, (0,0)) < \infty\}, \\ E_r &= \{\exists\ w \in C(t') \text{ s.t. } d(w, (0,0)) \leq r\}, \\ E_{r1} &= \{\exists\ w \in \mathcal{G}_S(\lambda_S) \text{ s.t. } d(w, (0,0)) \leq r\}, \end{aligned}$$

Then we have that for a fixed $r > 0$,

$$\begin{aligned} \Pr\{E\} &\geq \Pr\{E_r\} \geq \Pr\{E_{r1}\}\theta(\lambda_S, \lambda_{PT}) \\ &= [1 - \exp(-\lambda_S \pi r^2)]\theta(\lambda_S, \lambda_{PT}) > 0, \end{aligned}$$

where $\theta(\lambda_S, \lambda_{PT})$ is defined in (1) and is strictly positive since $(\lambda_S, \lambda_{PT}) \in \mathcal{C}$. It is easy to see that the event E is invariant of the shift transformations[6]. Thus, based on the ergodicity[7] of the heterogeneous network model [20, Lemma 1], we conclude that $\Pr\{E\} = 1$, *i.e.*, $d(w_\mu, (0,0)) < \infty$ a.s.

Next we show that the number of hops $|L(w_\nu, \nu)|$ on the path $L(w_\nu, \nu)$ is finite a.s. As shown in Fig. 5, we construct a sequence of concentric squares with increasing side lengths. Specifically, all the squares are centered at the middle point of w_ν and ν, and the side length of the j-th $(j \geq 0)$ square S_j is $3^j d$. Let A_j $(j \geq 1)$ denote the square annulus inside S_j and outside

[6]For a random model in a Euclidean space $\mathbb{R}^d$ with a probability space $(\Omega, \mathcal{F}, \mu)$, the shift transformation S_x is to shift the realization $\omega \in \Omega$ by $x \in \mathbb{R}^d$.

[7]A random model under a probability space $(\Omega, \mathcal{F}, \mu)$ is said to be ergodic if there exists a transformation group $\{S_x :\ x \in \mathbb{R}^d \text{ or } \mathbb{Z}^d\}$ that acts ergodically on $(\Omega, \mathcal{F}, \mu)$. A transformation group $\{S_x :\ x \in \mathbb{R}^d \text{ or } \mathbb{Z}^d\}$ is said to act ergodically if the σ-algebra of events invariant under the whole group is trivial, *i.e.*, any invariant event has measure either zero or one. For an ergodic random model $(\Omega, \mathcal{F}, \mu)$, if an event $E \in \mathcal{F}$ invariant under the whole transformation group $\{S_x :\ x \in \mathbb{R}^d \text{ or } \mathbb{Z}^d\}$ occurs with a positive probability, *i.e.*, $\mu(E) > 0$, then it occurs a.s., *i.e.*, $\mu(E) = 1$.

S_{j-1}, and let E_j^u be the event that there exists a left-to-right crossing[8] in the upper horizontal rectangle of A_j with side length $3^j d \times 3^{j-1} d$. Similarly, define E_j^b, E_j^l, and E_j^r as the events that the bottom, left, and right rectangles of A_j are crossed from left to right or from top to bottom. By symmetry, we know that $\Pr(E_j^u) = \Pr(E_j^b) = \Pr(E_j^l) = \Pr(E_j^r)$.

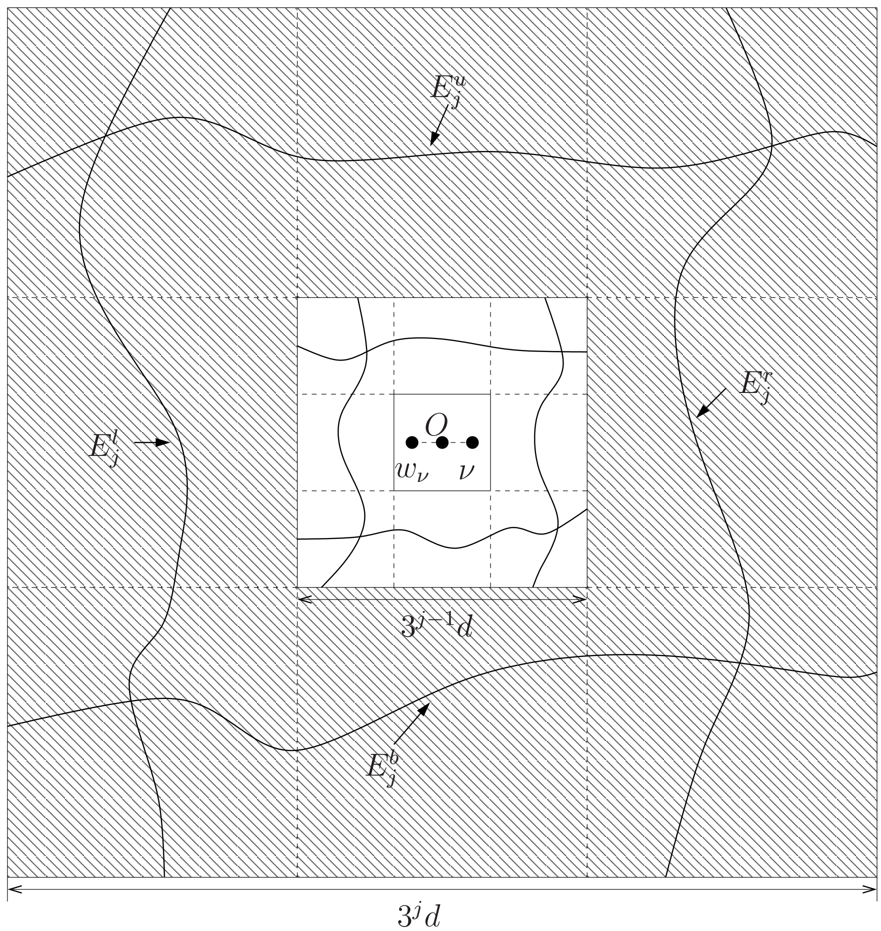

Fig. 5. A sequence $\{S_j : j \geq 0\}$ of squares cocentered at the middle point O of w_ν and ν. The shaded region is the square annulus A_j inside S_j with side length $3^j d$ and outside S_{j-1} with side length $3^{j-1} d$. In this example, the four crossings associated with the four events E_j^u, E_j^b, E_j^l, and E_j^r all exist in the corresponding four rectangles, which form a closed circuit in A_j.

Since $(\lambda_S, \lambda_{PT}) \in \mathcal{C}$, it follows that $\lambda_S > \lambda_c$ where λ_c is the critical density for a homoge-

[8] A left-to-right crossing exists in a rectangle $R = [x_1, x_2] \times [y_1, y_2]$ if and only if there exists a sequence of nodes μ_i $(1 \leq i \leq n)$ in $\mathcal{G}(\lambda_S)$ s.t. (i) $\mu_i \in R$ for all i; (ii) $d(\mu_{i+1}, \mu_i) \leq 1$ for all $1 \leq i < n$; (iii) $|x(\mu_1) - x_1| \leq \frac{1}{2}$ and $|x(\mu_n) - x_2| \leq \frac{1}{2}$, where $x(\mu_i)$ is the x-coordinate of μ_i. The top-to-bottom crossing can be defined analogously.

neous network with a unit transmission range[9]. By using Corollary 4.1 in [18], we have that $\lim_{d\to\infty} \Pr\{E_1^u\} = 1$. Then for a given $0 < \delta < 1$, we choose

$$d = d_\delta \stackrel{\Delta}{=} \max\{\inf\{d' : \ \Pr\{E_1^u\} \geq \delta \text{ if } d \geq d'\}, d(w_\nu, \nu)\}.$$

We thus have that $\Pr\{E_j^u\} \geq \delta$ for all $j \geq 1$.

Let E_j $(j \geq 1)$ be the event that there exists a closed circuit of connected users in $\mathcal{G}(\lambda_S)$ within A_j. If E_j^u, E_j^b, E_j^l, and E_j^r all occur, then E_j occurs (see Fig. 5). Since E_j^u, E_j^b, E_j^l, and E_j^r are all increasing events[10], it follows from the FKG inequality [18, Theorem 2.2] that

$$\begin{aligned}\Pr\{E_j\} &\geq \Pr\{E_j^u \cap E_j^b \cap E_j^l \cap E_j^r\} \\ &\geq \Pr\{E_j^u\}\Pr\{E_j^b\}\Pr\{E_j^l\}\Pr\{E_j^r\} \geq \delta^4.\end{aligned}$$

When E_j occurs, we claim that there exists a path $L'(w_\nu, \nu)$ from w_ν to ν within S_j. If all the paths from w_ν to ν go outside S_j, they will intersect the closed circuit in A_j and then we can construct a path $L'(w_\nu, \nu)$ within S_j by using part of the closed circuit.

As illustrated in Fig. 6, we place a circle with radius $\frac{1}{2}$ at each user along $L'(w_\nu, \nu)$, and it is easy to see that any two circles centered at the two users w_1 and w_2 on $L'(w_\nu, \nu)$ which are not connected via a topological link do not overlap, otherwise we can shorten the path by skipping the users between w_1 and w_2. Thus, given the number of hops $|L'(w_\nu, \nu)|$, at least $\left\lceil \frac{|L'(w_\nu,\nu)|}{2} \right\rceil$ nonoverlapping circles centered at alternating nodes on $L'(w_\nu, \nu)$ can be found, and they are all contained within the square with side length $3^j d_\delta + 1$. It follows that

$$|L'(w_\nu, \nu)| \leq 2 \left\lceil 4(3^j d_\delta + 1)^2/\pi \right\rceil < \infty,$$

where $2 \lceil 4(3^j d_\delta + 1)^2/\pi \rceil$ is the maximum number of nonoverlapping circles with radius $\frac{1}{2}$ within the square with side length $3^j d_\delta + 1$.

On the other hand, Since E_j are independent, and

$$\sum_{j=1}^{\infty} \Pr\{E_j\} \geq \sum_{j=1}^{\infty} \delta^4 = \infty,$$

[9]As illustrated in Fig. 1, the critical density λ_S^* of the secondary users is the the infimum density of the secondary users to ensure connectivity, which is shown in [20, 21] to be equal to λ_c.

[10]Consider two realizations ω and ω' of $\mathcal{G}(\lambda_S)$. A partial ordering '$\preceq$' is defined as $\omega \preceq \omega'$ if and only if every node in ω is also present in ω'. In other words, ω can be obtained from ω' by removing some secondary users. An event E is said to be increasing if for every $\omega \preceq \omega'$, $\mathbb{I}_E(\omega) \leq \mathbb{I}_E(\omega')$, where $\mathbb{I}_E$ is the indicator function of the event E.

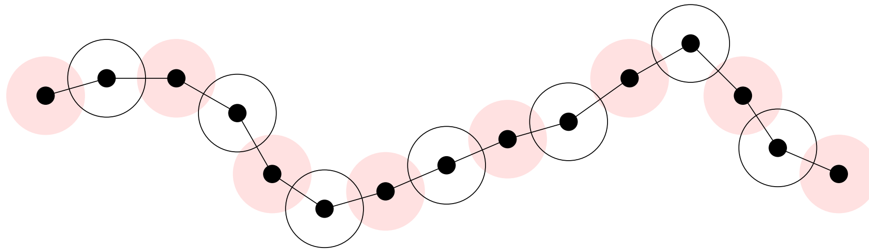

Fig. 6. An illustration of nonoverlapping colored circles with radii $\frac{1}{2}$ centered at alternating nodes on the path $L'(w_\nu, \nu)$.

it follows from the Borel-Cantelli Lemma that E_j occurs for some j a.s. We thus have that $|L'(w_\nu, \nu)| < \infty$ a.s. It implies that $|L(w_\nu, \nu)| \leq |L'(w_\nu, \nu)| < \infty$ a.s., which, together with the a.s. finiteness of the single-hop transmission delay (see Sec. III-A), yields the a.s. finiteness of the multihop delay $t^L(w_\nu, \nu)$ along the shortest path $L(w_\nu, \nu)$. ∎

B. Proof of Theorem 1.2

Let μ be the source and ν the destination. Take μ as the origin, and the line connecting μ and ν as the x-axis. Define an auxiliary node $\tilde{w}_i$ in $C(\mathcal{G}_S(\lambda_S))$ for every integer i:

$$\tilde{w}_i \stackrel{\Delta}{=} \arg\min_{w \in C(\mathcal{G}_S(\lambda_S))} d(w, (i, 0)).$$

Obviously, $\tilde{w}_0 = \mu$. Let n be the closest integer to ν, then

$$\frac{t(\tilde{w}_0, \tilde{w}_n) - t(\tilde{w}_n, \nu)}{n+1} \leq \frac{t(\mu, \nu)}{d(\mu, \nu)} \leq \frac{t(\tilde{w}_0, \tilde{w}_n) + t(\tilde{w}_n, \nu)}{n-1}.$$

If $\tilde{w}_n = \nu$, then $t(\tilde{w}_n, \nu) = 0$; if $\tilde{w}_n \neq \nu$, then $t(\tilde{w}_n, \nu)$ is at most the single-hop transmission delay because $d(\tilde{w}_n, \nu) \leq d(\tilde{w}_n, (n, 0)) + d(\nu, (n, 0)) \leq 2d(\nu, (n, 0)) \leq 1$.

Let

$$t_{m,n} = t(\tilde{w}_m, \tilde{w}_n) \text{ for any two integers } m, n. \tag{6}$$

Then to show T1.2, it suffices to show that

$$\lim_{n \to \infty} \frac{t_{0,n}}{n} = \beta > 0 \text{ a.s.} \tag{7}$$

The proof of (7) is divided into two steps, which are expressed in the form of the following two lemmas.

Lemma 2: $\beta \stackrel{\Delta}{=} \lim_{n \to \infty} \frac{t_{0,n}}{n}$ exists a.s.

Lemma 3: $0 < \beta = \lim_{n\to\infty} \frac{t_{0,n}}{n} < \infty$.

Proof of Lemma 2: The proof is based on Liggett's Subadditive Ergodic Theorem as follows:

Fact 1: [22, Theorem 1.10] Let $\{t_{m,n}\}$ be a collection of random variables indexed by integers satisfying $0 \le m < n$. Suppose $\{t_{m,n}\}$ has the following properties:

(i) $t_{0,n} \le t_{0,m} + t_{m,n}$.

(ii) For each n, $\mathbb{E}(|t_{0,n}|) < \infty$ and $\mathbb{E}(t_{0,n}) \ge cn$ for some constant $c > -\infty$.

(iii) The distribution of $\{t_{m,m+k} : \ k \ge 1\}$ does not depend on m.

(iv) For each $k \ge 1$, $\{t_{nk,(n+1)k} : \ n \ge 0\}$ is a stationary sequence.

Then:

(a) $\eta \stackrel{\Delta}{=} \lim_{n\to\infty} \frac{\mathbb{E}[t_{0,n}]}{n} = \inf_{n\ge 1} \frac{E[t_{0,n}]}{n}$.

(b) $T \stackrel{\Delta}{=} \lim_{n\to\infty} \frac{t_{0,n}}{n}$ exists a.s.

(c) $\mathbb{E}[T] = \eta$.

Furthermore, if

(v) the stationary sequence in (iv) is ergodic,

then

(d) $T = \eta$ a.s.

By the definition of the minimum multihop delay and the stationarity of the heterogeneous network model, it is obvious that conditions (i), (iii), and (iv) hold for the sequences $\{t_{m,n}\}$ defined in (6). We only need to show that conditions (ii) and (v) also hold for $\{t_{m,n}\}$. We first show that $\mathbb{E}(|t_{0,n}|) < \infty$ for each n.

By using the techniques similar to showing $d(w_\mu, (0,0)) < \infty$ a.s. in the proof of Lemma 1, we can easily see that $d(\tilde{w}_0, (0,0)) < \infty$ a.s. as well as $d(\tilde{w}_n, (n,0)) < \infty$ a.s. It follows that a.s.

$$d(\tilde{w}_0, \tilde{w}_n) \le d(\tilde{w}_0, (0,0)) + n + d(\tilde{w}_n, (n,0)) < \infty.$$

Let $L(\tilde{w}_0, \tilde{w}_n)$ be the shortest path from $\tilde{w}_0$ to $\tilde{w}_n$. Let $|L|$ denote the number of hops of $L(\tilde{w}_0, \tilde{w}_n)$ and $t^L_{0,n}$ the multihop delay along $L(\tilde{w}_0, \tilde{w}_n)$. Consider the sequence $\{S_j : \ j \ge 0\}$ of squares constructed in the proof of Lemma 1 (see Fig. 5). For any given $\sqrt[4]{\frac{8}{9}} < \delta < 1$, we choose

$$d = d_\delta \stackrel{\Delta}{=} \max\{\inf\{d' : \ \Pr\{E^u_1\} \ge \delta \text{ if } d = d'\}, d(\tilde{w}_0, \tilde{w}_n)\}.$$

Similarly, when the event E_j $(j \geq 1)$ occurs, we have $|L| \leq 2\left\lceil 4(3^j d_\delta+1)^2/\pi \right\rceil$.

If $|L(\tilde{w}_0, \tilde{w}_n)| > 2\left\lceil 4(3^j d_\delta+1)^2/\pi \right\rceil$, then none of the events E_1, E_2,...,E_j occur. Thus

$$\Pr\{|L| > 2\left\lceil 4(3^j d_\delta+1)^2/\pi \right\rceil\} \leq \prod_{i=1}^{j} \Pr\{E_i^c\} \leq (1-\delta^4)^j.$$

Let $M = 2\left\lceil 4(3^j d_\delta+1)^2/\pi \right\rceil$, then we have

$$\begin{aligned}
\mathbb{E}[|L|] &= \sum_{k=0}^{\infty} \Pr\{|L|>k\} = \sum_{k=0}^{M} \Pr\{|L|>k\} + \sum_{k=M+1}^{\infty} \Pr\{|L|>k\} \\
&\leq M + \sum_{j=1}^{\infty} 2\left\lceil 4(3^{j+1} d_\delta+1)^2/\pi \right\rceil \Pr\{|L| > 2\left\lceil 4(3^j d_\delta+1)^2/\pi \right\rceil\} \\
&\leq M + \sum_{j=1}^{\infty} 2\left\lceil 4(3^{j+1} d_\delta+1)^2/\pi \right\rceil (1-\delta^4)^j \\
&= M + \frac{72 d_\delta^2}{\pi} \sum_{j=1}^{\infty} 9^j (1-\delta^4)^j + \frac{48 d_\delta}{\pi} \sum_{j=1}^{\infty} 3^j (1-\delta^4)^j + 2\left(\frac{4}{\pi}+1\right) \sum_{j=1}^{\infty} (1-\delta^4)^j.
\end{aligned}$$

If $\delta > \sqrt[4]{\frac{8}{9}}$, $(1-\delta^4)^j < 9^{-j}$. It implies that $\mathbb{E}[|L|] < \infty$. Let $t_M = \max_{0 \leq d \leq 1} \{\mathbb{E}[t_s(d)]\}$ be the maximum expected single-hop transmission delay of one hop for all hop lengths $0 \leq d \leq 1$, then for all $n \geq 1$,

$$\mathbb{E}[t_{0,n}] \leq \mathbb{E}[t_{0,n}^L] \leq t_M \mathbb{E}[|L|] < \infty,$$

$\{t_{0,n}\}$ satisfies the condition (ii).

Next we show that $\{t_{nk,(n+1)k} : \ n \geq 0\}$ is mixing[11], which implies its ergodicity. As illustrated in Fig. 7, we construct two squares S_n and S_{n+j} centered at $\left(\frac{(2n+1)k}{2}, 0\right)$ and $\left(\frac{[2(n+j)+1]k}{2}, 0\right)$ with side length d_n and d_{n+j}. Let L_n^* be the minimum path (in terms of the multihop delay) from $\tilde{w}_{nk}$ to $\tilde{w}_{(n+1)k}$. We claim that the two paths L_n^* and L_{n+j}^* are contained a.s. in S_n and S_{n+j} respectively for some $d_n > 0$ and $d_{n+j} > 0$. If, for example,

$$\Pr\{E_n\} = \Pr\{L_n^* \text{ is not contained in any finite } S_n\} > 0,$$

[11]A measure preserving transformation T is said to be mixing on a probability space $(\Omega, \ \mathcal{F}, \ \mu)$ if for all $E, F \in \mathcal{F}$, $\mu(T^n E \cap F) - \mu(E)\mu(F) \to 0$ as $n \to \infty$. A sequence $\{x_k\}$ is said to be mixing if the unit right-shift transformation is mixing on its probability space. The mixing property of a sequence implies its ergodicity [23].

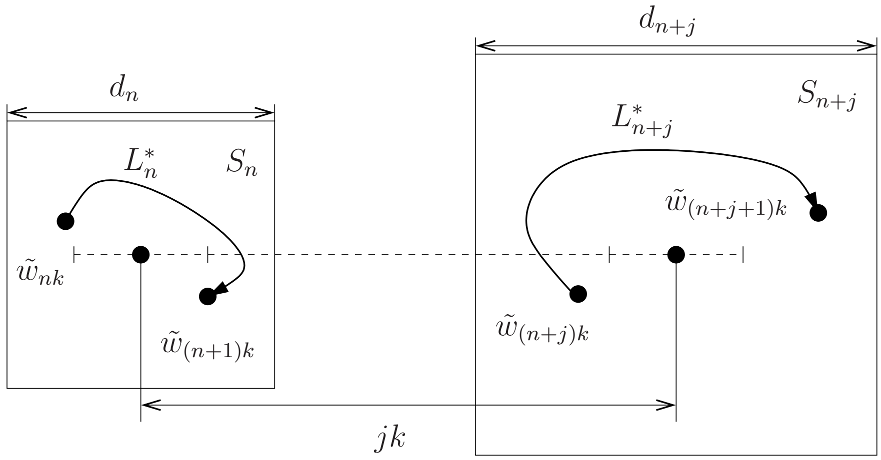

Fig. 7. The two minimum paths L^*_n (from $\tilde{w}_{nk}$ to $\tilde{w}_{(n+1)k}$) and L^*_{n+j} (from $\tilde{w}_{(n+j)k}$ to $\tilde{w}_{(n+j+1)k}$) are contained in the two squares S_n and S_{n+j} centered at $((2n+1)k/2, 0)$ and $((2n+2j+1)k/2, 0)$ with finite side length d_n and d_{n+j} respectively. As $j \to \infty$, S_n and S_{n+j} becomes nonoverlapping, and thus the multihop delay along L^*_n is asymptotically independent of the one along L^*_{n+j}.

then with a positive probability $|L^*_n| = \infty$, which implies that

$$\begin{aligned} \mathbb{E}[t_{nk,(n+1)k}] &\geq \mathbb{E}[t_{nk,(n+1)k}|\ E_n]\text{Pr}\{E_n\} \\ &\geq t_m \mathbb{E}[|L^*_n| \mid E_n]\text{Pr}\{E_n\} = \infty, \end{aligned}$$

with $t_m = \min\limits_{0\leq d\leq 1} \{\mathbb{E}[t_s(d)]\} > 0$ being the minimum expected single-hop transmission delay of one hop for all hop lengths[12]. This makes a contradiction to $\mathbb{E}[t_{nk,(n+1)k}] < \infty$. Then we have that as $j \to \infty$, not only the two minimum paths L^*_n and L^{n+j} do not share any common secondary users a.s., but also the subsets of the primary transmitter-receiver pairs that affect their multihop delay become disjoint a.s. Thus, $t_{nk,(n+1)k}$ and $t_{(n+j)k,(n+j+1)k}$ are asymptotically independent of each other as $j \to \infty$, *i.e.*,

$$\begin{aligned} &\lim_{j\to\infty} \text{Pr}\{\left(t_{nk,(n+1)k} < t\right) \cap \left(t_{(n+j)k,(n+j+1)k} < t'\right)\} \\ = \;& \text{Pr}\{t_{nk,(n+1)k} < t\}\text{Pr}\{t_{nk,(n+1)k} < t'\} \end{aligned}$$

The mixing property of $\{t_{nk,(n+1)k} :\ n \geq 0\}$ follows immediately. Since all the five conditions in Fact 1 are satisfied by $\{t_{m,n}\}$, we conclude that $\exists \beta \geq 0$ s.t. $\lim\limits_{n\to\infty} \frac{t_{0,n}}{n} = \beta$ a.s. ■

[12]The inequality $t_m > 0$ is shown in Proposition 2 in Appendix A.

Proof of Lemma 3: From Fact 1, we know that

$$\beta = \inf_{n\geq 1} \frac{E[t_{0,n}]}{n} \leq \mathbb{E}[t_{0,1}] < \infty.$$

To show that $\beta > 0$, we need the following fact which gives an upper bound on the CDF of the diameter[13] of the connected component in $\mathcal{G}_H(\lambda_S, \lambda_{PT}, t)$ when $(\lambda_S, \lambda_{PT}) \notin \mathcal{C}$. This fact can be easily proven by using techniques similar to the ones used in proving Theorem 2.4 in [18].

Fact 2: Given $\mathcal{G}_H(\lambda_S, \lambda_{PT}, t)$ for any t with $(\lambda_S, \lambda_{PT}) \notin \mathcal{C}$, let $B_h = [-h, h]^2$ $(h > 0)$ and take an arbitrary secondary user as the origin. Then $\exists\ C_1,\ C_2 > 0$ s.t. Pr$\{O \leftrightsquigarrow (B_h)^c\} \leq C_1 \exp(-C_2 h)$, where $\{O \leftrightsquigarrow (B_h)^c\}$ denotes the event that the origin is connected with some secondary user outside B_h, *i.e.*, the origin and some node in $(B_h)^c$ belong to the same connected component.

Now choose $H > 0$ s.t. $C_1 \exp(-C_2 H) < \frac{1}{2}$, where C_1 and C_2 are the constants specified in Fact 2. As show in Fig. 8, for any path L from $\tilde{w}_0$ to $\tilde{w}_n$, we partition it into several segments in the following way. Define a sequence $\{R_i :\ i \geq 1\}$ of uniformly distributed ribbons on $\mathbb{R}^2$ as

$$\begin{aligned} R_i \;\;=\;\; & \{(x, y) \in \mathbb{R}^2 : \\ & H + (i-1)(H+1) \leq x - x(\tilde{w}_0) < i(H+1)\}, \end{aligned}$$

where $x(\tilde{w}_0)$ is the x-coordinate of user $\tilde{w}_0$. Since the width of each ribbon is 1, there exits at least one user z_i within each R_i that lies between $\tilde{w}_0$ and $\tilde{w}_n$. Assume that these z_i partition the path L into m segments, then the multihop delay t^L along the path L can be written as

$$t^L = \sum_{i=1}^{m}\ t^L(z_{i-1}, z_i), \tag{8}$$

where $z_0 = \tilde{w}_0$ and $z_m = \tilde{w}_n$.

Based on fact 2, with a probability greater than $\frac{1}{2}$ at least one hop on the segment of L from z_{i-1} to z_i does not see the opportunity. We thus have that for all $1 \leq i \leq m-1$,

$$\mathbb{E}[t^L(z_{i-1}, z_i)] > \frac{1}{2} t_m, \tag{9}$$

where $t_m = \min\limits_{0\leq d\leq 1} \{\mathbb{E}[t(d)]\} > 0$.

[13]The diameter of a connected component C is defined as $\max\limits_{\mu,\nu\in C}\ d(\mu,\nu)$.

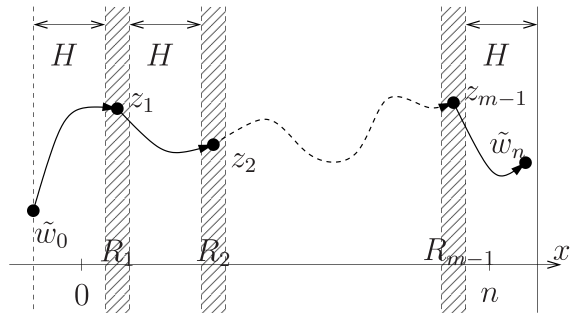

Fig. 8. A path L from $\tilde{w}_0$ to $\tilde{w}_n$ which is partitioned into m segments by users z_i $(1 \le i \le m-1)$ within these shaded ribbons R_i between $\tilde{w}_0$ and $\tilde{w}_n$.

Since $d(\tilde{w}_0, \tilde{w}_n) \ge n - d(\tilde{w}_0, (0,0)) - d(\tilde{w}_n, (n,0))$, and both $d(\tilde{w}_0, (0,0))$ and $d(\tilde{w}_n, (n,0))$ are finite a.s., it follows that $\lim_{n\to\infty} \Pr\{d(\tilde{w}_0, \tilde{w}_n) > \frac{n}{2}\} = 1$. When $d(\tilde{w}_0, \tilde{w}_n) > \frac{n}{2}$ holds, any path from $\tilde{w}_0$ to $\tilde{w}_n$ has at least $\left\lfloor \frac{n}{2(H+1)} \right\rfloor$ segments. By recalling (8, 9), we conclude that

$$\begin{aligned}\beta &= \lim_{n\to\infty} \frac{\mathbb{E}[t_{0,n}]}{n} \\ &> \lim_{n\to\infty} \frac{t_m}{2n} \left\lfloor \frac{n}{2(H+1)} \right\rfloor \Pr\left\{ d(\tilde{w}_0, \tilde{w}_n) < \frac{n}{2} \right\} \\ &> \lim_{n\to\infty} \frac{t_m}{2} \left(\frac{1}{2(H+1)} - \frac{1}{n} \right) > 0.\end{aligned}$$

This implies that $\beta = \lim_{n\to\infty} \frac{t_{0,n}}{n} > 0$. ■

C. *Proof of Theorem 2*

When the propagation delay $\tau > 0$, the single-hop transmission delay $t_s^\tau(d)$ is still a proper random variable with finite mean. Similarly, in order to show the a.s. existence of $\lim_{d(\mu,\nu)\to\infty} \frac{t^\tau(\mu,\nu)}{d(\mu,\nu)}$, it suffices to prove that $\lim_{n\to\infty} \frac{t^\tau_{0,n}}{n}$ exists a.s. By the same argument as used in the proof of Lemma 2, we can easily verify the five conditions in Fact 1. Then the a.s. existence of $\lim_{n\to\infty} \frac{t^\tau_{0,n}}{n}$ follows.

Let

$$\gamma = \gamma(\tau) = \lim_{d(\mu,\nu)\to\infty} \frac{t^\tau(\mu,\nu)}{d(\mu,\nu)}.$$

Since the minimum number of hops between μ and ν is $\lfloor d(\mu,\nu) \rfloor$ and the minimum transmission delay of one hop is τ, we have

$$t^\tau(\mu,\nu) \ge \tau \lfloor d(\mu,\nu) \rfloor,$$

which implies that $\gamma \geq \tau$.

Obviously, $\mathbb{E}[t^{\tau}_{0,n}]$ decreases as τ decreases. Then from Fact 1 and basic series theory, we have that

$$\begin{aligned} \lim_{\tau\to 0}\lim_{d(\mu,\nu)\to\infty}\frac{t^{\tau}(\mu,\nu)}{d(\mu,\nu)} &= \lim_{\tau\to 0}\gamma(\tau) \\ &= \lim_{\tau\to 0}\lim_{d(\mu,\nu)\to\infty}\frac{\mathbb{E}[t^{\tau}(\mu,\nu)]}{d(\mu,\nu)} \end{aligned}$$

exists a.s.

If $(\lambda_S, \lambda_{PT}) \in \mathcal{C}_I$, then

$$\lim_{\tau\to 0}\lim_{d(\mu,\nu)\to\infty}\frac{t^{\tau}(\mu,\nu)}{d(\mu,\nu)} \geq \lim_{\tau\to 0}\lim_{d(\mu,\nu)\to\infty}\frac{t(\mu,\nu)}{d(\mu,\nu)} = \beta,$$

where $t(\mu,\nu)$ is the minimum multihop delay from μ to ν for $\tau = 0$, and β is defined in (2).

If $(\lambda_S, \lambda_{PT}) \in \mathcal{C}$, then we consider the path L_C from μ to ν constructed in the proof of T1.1 which contains some nodes of the infinite connected component $C(t')$ in $\mathcal{G}_H(\lambda_S, \lambda_{PT}, t')$. Notice that for fixed $d(\mu,\nu)$, only a finite number of hops on L_C belong to $C(t')$. Thus if τ is sufficiently small, it takes at most one primary slot for the message to transmit from the source μ in $C(t')$ to the end node w_ν, and we have that for some small $\tau_0 = \tau_0(d(\mu,\nu)) > 0$,

$$t^{C}_{\tau_0}(\mu,\nu) \leq t' + t^{\tau_0}(w_\nu,\nu) + 1,$$

where $t^{C}_{\tau_0}(\mu,\nu)$ denotes the multihop delay along the path L_C when the propagation delay is τ_0. It implies that

$$\begin{aligned} \lim_{\tau\to 0}\gamma(\tau) &= \lim_{\tau\to 0}\lim_{d(\mu,\nu)\to\infty}\frac{\mathbb{E}[t^{\tau}(\mu,\nu)]}{d(\mu,\nu)} \\ &= \lim_{d(\mu,\nu)\to\infty}\lim_{\tau\to 0}\frac{\mathbb{E}[t^{\tau}(\mu,\nu)]}{d(\mu,\nu)} \\ &\leq \lim_{d(\mu,\nu)\to\infty}\frac{\mathbb{E}[t^{C}_{\tau_0}(\mu,\nu)]}{d(\mu,\nu)} \\ &\leq \lim_{d(\mu,\nu)\to\infty}\frac{\mathbb{E}[t'] + \mathbb{E}[t^{\tau_0}(w_\nu,\nu)] + 1}{d(\mu,\nu)} = 0, \end{aligned}$$

since both $\mathbb{E}[t']$ and $\mathbb{E}[t^{\tau_0}(w_\nu,\nu)]$ are finite and independent of $d(\mu,\nu)$. In the second equality, we can interchange the order of the two limits because $\mathbb{E}[t^{\tau}(\mu,\nu)] < \infty$. Consequently, we conclude that a.s.

$$\lim_{\tau\to 0}\lim_{d(\mu,\nu)\to\infty}\frac{t^{\tau}(\mu,\nu)}{d(\mu,\nu)} = \lim_{\tau\to 0}\gamma(\tau) = 0.$$

VI. Conclusion

We have studied the connectivity and multihop delay of ad hoc cognitive radio networks. The criterion for connectivity is the occurrence of percolation, *i.e.*, the almost sure existence of an infinite connected component. The impact of connectivity on the multihop delay has been examined by establishing the asymptotic behavior of the minimum multihop delay as the source-destination distance tends to infinity. Specifically, depending on whether the cognitive radio network is connected or intermittently connected, the scaling of the minimum multihop delay behaves distinctly, in terms of either the scaling law when the propagation delay is negligible or the scaling rate when the propagation delay is nonnegligible. This result on scaling is independent of the random positions of the source and the destination, and it only depends on the network parameters (e.g., the density of the secondary users and the traffic load of the primary network). In establishing these results, we have used theories and techniques from continuum percolation and ergodicity including the concept of critical density and the Subadditive Ergodic Theorem.

Appendix A: Probability of Bidirectional Spectrum Opportunity

The expression for the probability of a bidirectional spectrum opportunity is presented in the following proposition.

Proposition 1: Let λ_S be the density of the secondary users. Let r_I and R_I be the interference range of the secondary and primary users, respectively, and r_p and R_p the transmission range of the secondary and primary users, respectively. Then for any two secondary users μ and ν with distance $d \leq r_p$ apart, the probability p_0 of having a bidirectional spectrum opportunity at the topological link $e_{\mu\nu}$ connecting μ and ν is given by

$$p_0 = \exp\Bigg\{ -\lambda_{PT}\Big[2\pi(r_I^2+R_I^2) - S_I(d,r_I,r_I) - S_I(d,R_I,R_I) - \iint_{\mathcal{S}_{U2}(d,R_I,R_I)} \frac{S_{I2}(r,\theta,R_p,d,r_I)}{\pi R_p^2} r\mathrm{d}r\mathrm{d}\theta\Big]\Bigg\}, \tag{A1}$$

where $S_I(d,r_1,r_2)$ the common area of two circles with radii r_1 and r_2 and centered d apart (see Fig. 9(a)), and $\mathcal{S}_{U2}(d,r_1,r_2)$ is the union of two circles with radii r_1 and r_2 and centered d apart (see Fig. 9(b)). $S_{I2}(r,\theta,R_p,d,r_I)$ is the intersection area between one circle with radius R_p and the union of the two circles with both radii r_I (see Fig. 9(c)). For $S_{I2}(r,\theta,R_p,d,r_I)$, the two identical circles are centered d apart, and the other circle is centered at (r, θ), where the middle point of the centers of the two identical circles is chosen to be the origin O.

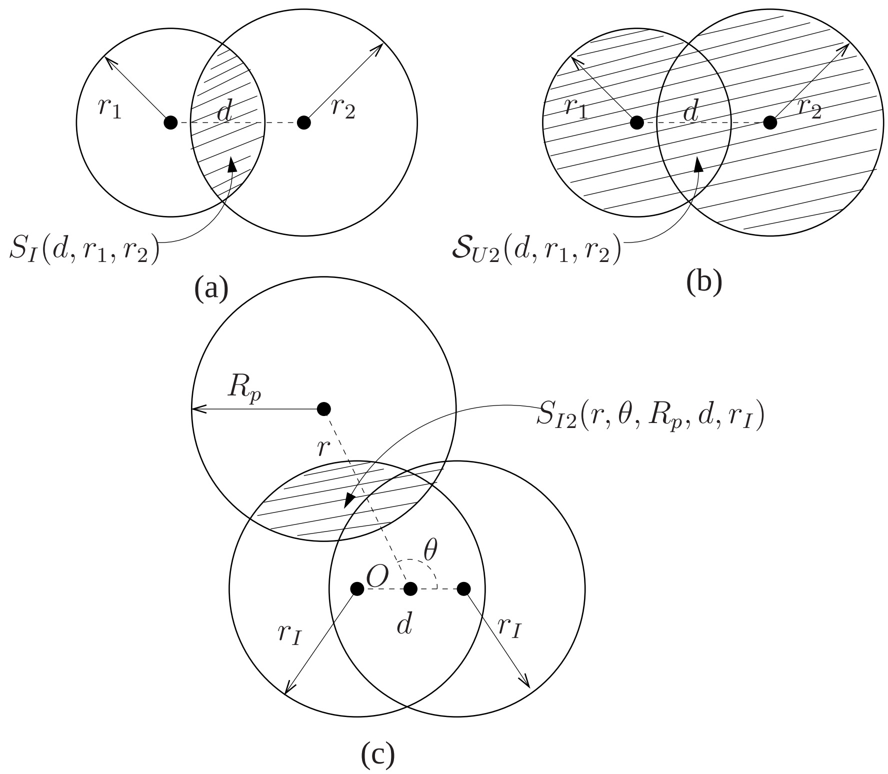

Fig. 9. An illustration of $S_I(d, r_1, r_2)$ (the common area of two circles with radii r_1 and r_2 and centered t apart), $\mathcal{S}_{U2}(d, r_1, r_2)$ (the union area of two circles with radii r_1 and r_2 and centered d apart), and $S_{I2}(r, \theta, R_p, d, r_I)$ (the intersection area between one circle with radius R_p and the union of the two identical circles with radii r_I).

The expression for $S_I(d, r_1, r_2)$ can be obtained in explicit form, which can be found in [24, Appendix A]. The expression for $S_{I2}(r, \theta, R_p, d, r_I)$ depends on the expression for the common area of three circles which is tedious and is given in [25].

From (A1), we can easily see that the probability p_0 of a bidirectional opportunity is a function of the distance d between μ and ν. Moreover, although p_0 is an exponentially decreasing function of λ_{PT}, it is strictly positive no matter how large λ_{PT} is.

Proof: From the definition of spectrum opportunity given in Sec. III-A, we know that a bidirectional spectrum opportunity occurs at $e_{\mu\nu}$ if and only if there are no primary transmitters within distance R_I of either μ or ν and no primary receivers within distance r_I of either μ or ν.

Let $\mathbb{I}(\mu, d, \text{rx/tx})$ denote the event that there exists primary receivers/transmitters within distance d of a secondary user μ. Let $\overline{\mathbb{I}(\mu, d, \text{rx/tx})}$ denote the complement of $\mathbb{I}(\mu, d, \text{rx/tx})$. Then the probability p_0 of a bidirectional opportunity at $e_{\mu\nu}$ is given by

$$\begin{aligned} p_0 &= \Pr\{\overline{\mathbb{I}(\nu, r_I, \text{rx})} \cap \overline{\mathbb{I}(\nu, R_I, \text{tx})} \cap \overline{\mathbb{I}(\mu, r_I, \text{rx})} \cap \overline{\mathbb{I}(\mu, R_I, \text{tx})}\} \\ &= \Pr\{\overline{\mathbb{I}(\nu, R_I, \text{tx})} \cap \overline{\mathbb{I}(\mu, R_I, \text{tx})} |\ \overline{\mathbb{I}(\nu, r_I, \text{rx})} \cap \overline{\mathbb{I}(\mu, r_I, \text{rx})}\} \Pr\{\overline{\mathbb{I}(\nu, r_I, \text{rx})} \cap \overline{\mathbb{I}(\mu, r_I, \text{rx})}\}. \quad \text{(A2)} \end{aligned}$$

Next, we compute the two probabilities in (A2) one by one. Since the primary receivers admit a Poisson point process X_{PR} with density λ_{PT}, we have

$$\Pr\{\overline{\mathbb{I}(\nu, r_I, \text{rx})} \cap \overline{\mathbb{I}(\mu, r_I, \text{rx})}\} = \exp[-\lambda_{PT}(2\pi r_I^2 - S_I(d, r_I, r_I))], \quad \text{(A3)}$$

where $S_I(d, r_I, r_I)$ is the common area of two circles with both radii r_I and centered t apart (see Fig. 9(a)).

Let X_{PT} denote the Poisson point process formed by primary transmitters. If we remove from X_{PT} primary transmitters whose receivers are within distance r_I of either ν or μ, then it follows from Coloring Theorem [17, Chapter 5] that all the remaining primary transmitters form another Poisson point process with density $\lambda_{PT}\left[1 - \frac{S_{I2}(r,\theta,R_p,d,r_I)}{\pi R_p^2}\right]$, where $S_{I2}(r, \theta, R_p, d, r_I)$ is the area of the circle with radius R_p and centered at (r, θ) intersecting the two circles with both radii r_I and centered d apart (see Fig. 9(c)). We thus have

$$\begin{aligned} &\Pr\{\overline{\mathbb{I}(\nu, R_I, \text{tx})} \cap \overline{\mathbb{I}(\mu, R_I, \text{tx})} |\ \overline{\mathbb{I}(\nu, r_I, \text{rx})} \cap \overline{\mathbb{I}(\mu, r_I, \text{rx})}\} \\ = &\ \exp\left\{ -\lambda_{PT} \iint\limits_{\mathcal{S}_{U2}(d,R_I,R_I)} \left[1 - \frac{S_{I2}(r, \theta, R_p, d, r_I)}{\pi R_p^2} r\mathrm{d}r\mathrm{d}\theta\right]\right\} \\ = &\ \exp\left\{ -\lambda_{PT} \left[2\pi R_I^2 - S_I(d, R_I, R_I) - \iint\limits_{\mathcal{S}_{U2}(d,R_I,R_I)} \frac{S_{I2}(r, \theta, R_p, d, r_I)}{\pi R_p^2} r\mathrm{d}r\mathrm{d}\theta \right]\right\}, \quad \text{(A4)} \end{aligned}$$

where $\mathcal{S}_{U2}(d, R_I, R_I)$ is the union of two circles with both radii R_I and centered d apart (see Fig. 9(b)). Then plugging (A3, A4) into (A2) yields (A1). ■

Based on Proposition 1, we establish an inequality on the minimum expected single-hop transmission delay for all link lengths in the following proposition.

Proposition 2: Given the parameters of the primary and secondary network (*i.e.*, λ_{PT}, r_p, r_I, R_p, and R_I are fixed), let $t_s(d)$ is the single-hop transmission delay for the link length d, and

t_m denote the minimum expected single-hop transmission delay for all link lengths when the propagation delay $\tau = 0$, *i.e.*,

$$t_m = \min_{0 \le d \le r_p} \{\mathbb{E}[t_s(d) |\ \tau = 0]\}.$$

Then we have $t_m > 0$.

Proof: It follows from Proposition 1 that for fixed λ_{PT}, r_p, r_I, R_p, and R_I, the probability p_0 of the bidirectional opportunity is a function of the link length d, *i.e.*, $p_0 = p_0(d)$. Let $t_{sw}(d)$ be the waiting time for the bidirectional opportunity at the link with length d, then given that $\tau < T_S$, $t_{sw}(d)$ is a geometric random variable with parameter $p_0(d)$ due to the i.i.d. distribution of the primary network across slots. The mean of $t_{sw}(d)$ is thus given by

$$\mathbb{E}[t_{sw}(d)] = \frac{1 - p_0(d)}{p_0(d)},$$

which is a monotonically decreasing function of $p_0(d)$.

From (A2), we have that for all $0 \le d \le r_p$,

$$\begin{aligned} p_0(d) &= \Pr\{\overline{\mathbb{I}(\nu, r_I, \text{rx})} \cap \overline{\mathbb{I}(\nu, R_I, \text{tx})} \cap \overline{\mathbb{I}(\mu, r_I, \text{rx})} \cap \overline{\mathbb{I}(\mu, R_I, \text{tx})}\} \\ &\le \Pr\{\overline{\mathbb{I}(\mu, r_I, \text{rx})} \cap \overline{\mathbb{I}(\mu, R_I, \text{tx})}\} \\ &= \exp[-\lambda_{PT}\pi(r_I^2 + R_I^2 - I(R_I, R_p, r_I))] < 1, \end{aligned}$$

where the last equality has been obtained by setting the distance $d = 0$ in the expression for the probability of a unidirectional opportunity between two secondary users with distance d apart given in Proposition 1 in [24], and

$$I(R_I, R_p, r_I) = 2\int_0^{R_I} t \frac{S_I(t, R_p, r_I)}{\pi R_p^2} \text{d}t < r_I^2.$$

The expression for $I(R_I, R_p, r_I)$ in explicit form can be found in [24, Appendix A].

Let $\delta = \exp[-\lambda_{PT}\pi(r_I^2 + R_I^2 - I(R_I, R_p, r_I))]$. Obviously, $0 < \delta < 1$. When $\tau = 0$, we have that $\forall\ 0 \le d \le r_p$,

$$\mathbb{E}[t_s(d)] = \mathbb{E}[t_{sw}(d)] = \frac{1 - p_0(d)}{p_0(d)} \ge \frac{1 - \delta}{\delta}.$$

Thus,

$$t_m = \min_{0 \le d \le r_p} \{\mathbb{E}[t_s(d) |\ \tau = 0]\} \ge \frac{1 - \delta}{\delta} > 0.$$

■

References

[1] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, May 2007.

[2] I. J. Mitola and J. G. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Personal Communications*, vol. 6, pp. 13–18, August 1999.

[3] O. Lévêque and I. E. Telatar, "Information-theoretic upper bounds on the capacity of large extended ad hoc wireless networks," *IEEE Transactions on Information Theory*, vol. 51, no. 3, pp. 858–865, March 2005.

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

[5] C. Yin, L. Gao, and S. Cui, "Scaling laws of overlaid wireless networks: A cognitive radio network vs. a primary network," *IEEE/ACM Transactions on Networking (submitted). http://arxiv.org/abs/0805.1209*, May 2008.

[6] S.-W. Jeon, N. Devroye, M. Vu, S.-Y. Chung, and V. Tarokh, "Cognitive networks achieve throughput scaling of a homegeneous network," *IEEE Transactions on Information Theory (submitted). http://arxiv.org/pdf/0801.0938*, January 2008.

[7] M. Vu, N. Devroye, M. Sharif, and V. Tarokh, "Scaling laws of cognitive networks," in *Proc. of IEEE CrownCom*, August 2007, pp. 2–8.

[8] N. Bansal and Z. Liu, "Capacity, delay and mobility in wireless ad-hoc networks," in *Proc. of IEEE INFOCOM*, vol. 2, April 2003, pp. 1553–1563.

[9] G. Sharma and R. Mazumdar, "Scaling laws for capacity and delay in wireless ad hoc networks with random mobility," in *Proc. of IEEE ICC*, vol. 7, June 2004, pp. 3869–3873.

[10] L. B. Le, K. Jagannathan, and E. Modiano, "Delay analysis of maximum weight scheduling in wireless ad hoc networks," in *Proc. of IEEE CISS*, March 2009, pp. 389–394.

[11] M. J. Neely and E. Modiano, "Capacity and delay tradeoffs for ad hoc mobile networks," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 1917–1937, June 2005.

[12] G. Sharma, R. Mazumdar, and N. B. Shroff, "Delay and capacity trade-offs in mobile ad hoc networks: A global perspective," *IEEE/ACM Transactions on Networking*, vol. 15, no. 5, pp. 981–992, October 2007.

[13] L. Ying, S. Chao, and R. Srikant, "Optimal delay-throughput tradeoffs in mobile ad hoc

networks," *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 4119–4143, September 2008.

[14] O. Dousse, P. Mannersalo, and P. Thiran, "Latency of wireless sensor networks with uncoordinated power saving mechanisms," in *Proc. of ACM MobiHoc*, May 2004, pp. 109–120.

[15] Z. N. Kong and E. M. Yeh, "Distributed energy management algorithm for large-scale wireless sensor networks," in *Proc. of ACM MobiHoc*, September 2007, pp. 209–218.

[16] ——, "Connectivity, percolation, and information dissemination in large-scale wireless networks with dynamic links," *IEEE Transactions on Information Theory (submitted)*, Feburary 2009.

[17] J. F. C. Kingman, *Poisson Processes*. New York: Oxford University Press, 1993.

[18] R. Meester and R. Roy, *Continuum Percolation*. New York: Cambridge Univerisity Press, 1996.

[19] Q. Zhao, "Spectrum opportunity and interference constraint in opportunistic spectrum access," in *Proc. of IEEE ICASSP*, vol. 3, April 2007, pp. 605–608.

[20] W. Ren, Q. Zhao, and A. Swami, "Connectivity of heterogeneous wireless networks," *IEEE Transactions on Information Theory (submitted). http://arxiv.org/abs/0903.1684.*, August 2009.

[21] ——, "Connectivity of cognitive radio networks: Proximity vs. opportunity," in *Proc. of ACM MobiCom Workshop on Cognitive Radio Networks*, September 2009.

[22] T. M. Liggett, "An improved subadditive ergodic theorem," *The Annals of Probability*, vol. 13, no. 4, pp. 1279–1285, 1985.

[23] K. Petersen, *Ergodic Theory*. New York: Cambridge Univerisity Press, 1989.

[24] W. Ren, Q. Zhao, and A. Swami, "Power control in cognitive radio networks: How to cross a multi-lane highway," *IEEE Journal on Selected Areas in Communications (JSAC): Special Issue on Stochastic Geometry and Random Graphs for Wireless Networks (to appear)*, September 2009.

[25] M. P. Fewell, "Area of common overlap of three circles," *Technical Note (http://handle.dtic.mil/100.2/ADA463920)*, October 2006.