Improving Network Connectivity in the Presence of Heavy-Tailed Interference

Pu Wang, *Member, IEEE*, and Ian F. Akyildiz, *Fellow, IEEE*

***Abstract*—The heavy tailed (HT) traffic from wireless users, caused by the emerging Internet and multimedia applications, introduces a HT interference region within which network users will experience unbounded delay with infinite mean and/or variance. Specifically, it is proven that, if the network traffic of primary networks (e.g., cellular and Wi-Fi networks) is heavy tail distributed, there always exists a critical density λ_p such that, if the density of primary users is larger than λ_p, the secondary network users (e.g., sensor devices and cognitive radio users) can experience unbounded end-to-end delay with infinite variance even though there exists feasible routing paths along the network users. To counter this problem, the mobility of network users is utilized to achieve delay-bounded connectivity, which simultaneously ensures the existence of routing paths and the finiteness of the delay variance along these paths. In particular, it is shown that there exists a critical threshold on the maximum radius that the secondary user can reach, above which delay-bounded connectivity is achievable in the secondary networks. In this case, the end-to-end latency of secondary users is shown to be asymptotically linear in the Euclidean distance between the transmitter and receiver.**

***Index Terms*—Heavy tail, connectivity, latency.**

I. Introduction

THE EMERGING Internet and multimedia applications, such as voice over Internet Protocol, telemedicine, online gaming, video conferencing, and multimedia surveillance, are expected to become dominant in current and next-generation wireless networks. Providing substantial levels of network quality of service (QoS) for these applications in the wireless domain is challenging because of the bursty nature of Internet/multimedia traffic and unreliable channel condition. Specifically, significant empirical evidence establishes that the scene length of video streams, the file size on Internet servers, and the burst duration of data center traffic can follow heavy tailed (HT) distribution [2], [8], [10], [18]. This distribution, compared with the conventional light tailed (LT) one, e.g., exponential distribution and Poisson distribution, exhibits significantly different statistical attributes that can completely change the way how wireless networks are conceived, designed, and operated.

The emerging HT traffic in wireless users necessarily invokes highly dynamic and variable wireless interference. More specifically, recent empirical results show that the channel occupancy time of both Wi-Fi and cellular networks is heavy tail distributed, which inevitably leads to extremely high channel variations [10], [18]. In particularly, it has been shown that such HT channel behavior greatly affects the channel sensing solutions [18], significantly degrades the delay and stability performance of wireless networks [15], [16], and fundamentally challenges the applicability of conventional network control schemes [17]. Despite its importance, the fundamental impact of HT traffic on the network performance limits is still an underexplored area.

As an important network performance metric, connectivity has to be maintained for reliable communications between transmitting and receiving parties in a network. Conventionally, there exist two types of connectivity: full connectivity and percolation-based connectivity [5], [20]. Specifically, full connectivity ensures that each pair of nodes in the network is connected by at least one path. However, for wireless networks, this connectivity criterion is overly restrictive or difficult to achieve because of the complicated radio environment, unplanned network topology, and severe impacts from coexisting networks. Different from full connectivity, percolation theory concerns a phase transition phenomenon where the network exhibits fundamentally different behavior for the density λ below and above some critical density λ_c. If $\lambda > \lambda_c$, the network is percolated or in the supercritical phase and it contains a giant connected component, which consists of an infinite number of nodes. Otherwise, If $\lambda < \lambda_c$, the network is in the subcritical phase, and the network is partitioned into small components containing a finite number of nodes.

Although percolation-based connectivity can characterize the existence of routing paths between network devices, it does not indicate the end-to-end QoS performance, such as delay and jitter. What is more important, as implied by our research [16], [17], under HT traffic activities, the wireless network can generate an HT interference region within which the coexisting network users will experience unbounded transmission delay with infinite variance. Such unbounded transmission delay can further lead to infinite average queuing delay [17], which can lead to significantly degraded QoS performance. Therefore, a latency-oriented connectivity definition is more meaningful in the HT network environment. Toward this, delay-bounded connectivity is of particular interest in this paper, which simultaneously ensures the existence of routing paths and the finite variance of the transmission delay along these paths.

Manuscript received September 8, 2013; revised March 27, 2014; accepted July 2, 2014. Date of publication July 22, 2014; date of current version October 8, 2014. This work was supported by the U.S. Army Research Office as a Multidisciplinary University Research Initiative under Grant W911NF-08-1-0238. The associate editor coordinating the review of this paper and approving it for publication was Y. Chen.

P. Wang is with the Department of Electrical Engineering and Computer Science, Wichita State University, Wichita, KS 67260 USA (e-mail: pu.wang@wichita.edu).

I. F. Akyildiz is with the Broadband Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: ian@ece.gatech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2014.2341635

Mobility, as the inherent nature of today's wireless networks, has been widely exploited to increase network capacity at the cost of latency. Contrary to this conventional objective, we show that mobility could actually help to improve the delay performance in the wireless networks so that delay-bounded connectivity is achievable in the presence of heavy tails. More specifically, by exploiting the spatial diversity of the spectrum availability, mobility can allow network users to evade from the HT interference region induced by the primary networks. This can further guarantee the delay boundness by preventing the rise of the HT delay. In this case, the critical question is how far network users need to move so that they can evade from such giant and irregular interference region of large-scale networks.

In this work, we study the fundamental impact of HT spectrum activities on the network connectivity as well as how and to what extent mobility can mitigate such impact. We consider a heterogeneous network setting, where there exist two networks: the primary network and the secondary network, where the primary networks either have the higher priority to access the spectrum than the secondary ones, as enforced by the emerging dynamic spectrum access scheme [1], or have much higher power to interfere the coexisting networks, such as the high-power Wi-Fi networks coexisting with the low-power wireless sensor networks. More specifically, we show that such HT traffic activities of the primary network significantly degrade the connectivity of the secondary network. Specifically, it is proven that, if the busy time of the primary users (PUs) is heavy tail distributed, there always exists a critical density λ_p such that, if the density of PUs is larger than λ_p, the secondary network cannot achieve delay-bounded connectivity surely. To encounter this problem, the mobility of secondary users is utilized to exploit the spatial diversity of the spectrum availability through the opportunistic contacts of mobile users. In particular, we prove that there exists a critical threshold on the maximum radius that the secondary user can reach, above which the secondary network can already achieve delay-bounded connectivity, independent of primary network impact such as node density and activities. Moreover, we study the latency performance of the mobility-assisted data forward schemes, which shows that their yielded end-to-end latency scales linearly in the initial distance between two mobile users.

The rest of this paper is organized as follows. In Section II, we introduce system models. In Section III, we formulate the problems and summarize the main results. Section IV gives the proof of the main results. Section V shows the simulation results. Finally, Section VI concludes this paper.

II. System Models

A. HT Distribution

Definition 1: An random variable (rv) X is HT if its tail distribution or complementary cumulative distribution function (CCDF) decays slower than exponentially, i.e., for all $\theta > 0$

$$\lim_{x\to\infty} e^{\theta x}\overline{F}(x) = \infty. \tag{1}$$

An rv X is LT if it is not HT.

Remark 1: Some typical HT distributions include Pareto and log-normal, while the typical LT distributions include exponential and Gamma.

In this paper, we focus on an important subclass of HT distributions, namely, regularly varying distribution, which is defined as follows.

Definition 2: An rv X is called regularly varying with tail index $c > 0$, denoted by $X \in \mathcal{RV}(c)$, if

$$\overline{F}(x) \sim x^{-c}\mathcal{L}(x), \tag{2}$$

where $\mathcal{L}(x)$ is a slowly varying function.

Remark 2: Regularly varying distributions are a generalization of Pareto/Ziph/power-law distributions. The tail index c indicates how heavy the tail distribution is, where smaller values of c imply a heavier tail. Moreover, for an rv $X \in \mathcal{RV}(c)$, the tail index c defines the maximum order of bounded moments that X can have. Specifically, if $0 < c < 1$, X has infinite mean and variance. If $1 < c < 2$, X has finite mean and infinite variance. In this work, we focus on regularly varying rvs because they can effectively characterize lots of network attributes such as the frame length of variable-bit-rate traffic, the session duration of licensed users in WLANs, and file sizes on Internet servers [10], [13].

B. Network Model

We consider a heterogeneous network setting, where there exist two networks: the primary network and the secondary network, where the primary networks either have the higher priority to access the spectrum than the secondary ones, as enforced by the emerging dynamic spectrum access scheme, or have much higher power to interfere the coexisting networks, such as the high-power Wi-Fi networks coexisting with the low-power wireless sensor networks.

1) Primary Network Model: We model the primary network as a random graph denoted by $G(\lambda_p, R)$ with PU density λ_p and transmission range R. Specifically, the PUs are distributed according to homogeneous Poisson point process with density λ_p. Let $\{Y_i\}_{i\leq n}$ denote the random locations of the PUs $\{1, 2, \ldots, n\}$ on the plane $\mathbb{R}^2$. With each PU i as the center, place a disk with radius equal to R if PU i is busy and with radius equal to 0 if PU i is idle. To model the activities of PUs, we associate each PU with an alternating renewal process, denoted by $S(t)$, which alternates between busy periods $\{B_i\}_{i\geq 1}$ and idle periods $\{I_i\}_{i\geq 1}$, which are mutually independent random sequences of independent and identically distributed (i.i.d.) rvs. This model has shown to be able to effectively characterize the PU behavior such as cellular and Wi-Fi users [4], [19]. To effectively characterize the high burstiness in the PU traffic, e.g., multimedia and Internet traffic, the busy periods $\{B_i\}_{i\geq 1}$ follows HT distributions, while the idle periods $\{I_i\}_{i\geq 1}$ can be either light tail or HT distributed.

2) Secondary Network Model: We model the secondary network as a random graph denoted by $G(\lambda_s, r)$ with secondary user density λ_s and transmission range r. Specifically, the secondary users are distributed according to homogeneous Poisson point process with density λ_s. Let $\{X_u\}_{u\leq n}$ denote the random locations of the secondary users $\{1, 2, \ldots, n\}$ on the plane $\mathbb{R}^2$. If the mutual distance between two secondary users, u and v, is less than the transmission range r, there exists a link e_{uv} between the two secondary users. Each link e_{uv} is associated with an rv $T(L, e_{uv})$, which denotes the time of transmitting a message of random size L over link e_{uv}. Since $T(L, e_{uv})$ is determined by the size of the message size L and the interference from the PUs, $T(L, e_{uv})$ is called interference delay over link e_{uv}.

C. HT Interference

To evaluate the interference delay $T(L, e_{uv})$ over link e_{uv} in the secondary network, we employ the Protocol Model in [6] as the interference model. Suppose that secondary user i located at X_i transmits to another secondary user j located at X_j. Let $\|\cdot\|$ be the Euclidean norm. The transmission is successful if the following two conditions are satisfied.

1) The distance between secondary i and j is less than the transmission range r, i.e., $\|X_i - X_j\| < r$.
2) The location Y_k of any active PU k, which is transmitting data at the same time as the secondary user i transmitting to j, should satisfy

$$\|X_j - Y_k\| > R. \tag{3}$$

Based on the above interference model, we evaluate the lower bound of the transmission time or interference delay $T(L, e_{uv})$ over a link e_{uv} in the secondary network $G(\lambda_s, r)$. We say that the link e_{uv} is interfered by the transmissions of PUs if at least one of u and v is residing within the range R of some PUs. In this case, the link e_{uv} is available for secondary users only if all the interfering PUs are idle. Therefore, $T(L, e_{uv})$ can be lower bounded by $T(L)$, where $T(L)$ is the time of transmitting message L over e_{uv}, provided that e_{uv} is interfered by exactly one PU, i.e,

$$T(L, e_{uv}) > T(L). \tag{4}$$

In particular, $T(L)$ is determined by the PU activities and can be evaluated as follows.

Suppose that L is an rv independent of the activities of PUs $\{B_i\}_{i\geq 1}$ and $\{I_i\}_{i\geq 1}$. For each message, the SU divides it into packets with constant size $L_p > 0$, which are then sent over the wireless channel. In each idle period I_i, the SU sends packets consecutively until the PU starts to transmit. In this case, if PU resumes its transmissions during the SU's transmission, SU's currently transmitted packet is lost. The SU will wait for the next idle period for the retransmission of the lost packet. An illustration of this model is given in Fig. 1. Now, the transmission time of the secondary user is formally defined as follows.

Definition 3: During an idle period I_i, the transmission time X_i of the SU is defined as

$$X_i := \sup\{nL_p : nL_p \leq I_i\}, \tag{5}$$

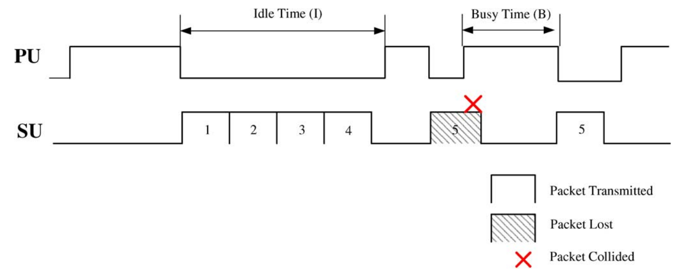

Fig. 1. Interference delay.

the total number of idle periods that the SU occupies for transmitting a message of size L is defined as

$$M := \inf\left\{m : \sum_{i=1}^{m} X_i \geq L\right\}, \tag{6}$$

and the total delay T of the SU transmitting a message of size L is defined as

$$T(L) := \sum_{i=1}^{M-1} \{I_i + B_i\} + \left(L - \sum_{i=1}^{M-1} X_i\right). \tag{7}$$

Note that $(L - \sum_{i=1}^{M-1} X_i)$ is the exact transmission time in the last idle period.

Assume that the message size L has bounded mean and variance. Under HT PU busy time, the transmission time or interference delay $T(L, e_{uv})$ over link e_{uv} is also HT with unbounded moments as discussed in the following lemma.

Lemma 1: If the PU busy time $B_i \in \mathcal{RV}(\alpha_b)$ with $\alpha_b < 2$, then $Var[T(L, e_{uv})] = \infty$.

The proof follows the similar procedures in [16] and therefore is omitted here.

Remark 3: According to Wang and Akyildiz [17], if the interference delay of a network user is HT, its queuing delay is also HT with one-order heavier tail distribution. This, combining with Lemma 1, means that, if the PU busy time $B_i \in \mathcal{RV}(\alpha_b)$ with $\alpha_b < 2$, then the secondary users will experience unbounded variance for transmission delay, which can further lead to unbounded average queuing delay.

III. Problem Formulation and Main Results

To study delay-bounded connectivity, we first define the transmission latency as follows. When two secondary users u and v are connected in the secondary network $G(\lambda_s, r)$ with density λ_s and node transmission range r, there exists at least one path between u and v consisting of links in the $G(\lambda_s, r)$. When u transmits a message of size L, this message can be delivered to v through different paths. For each path, we define the transmission latency between u and v as the total time that the message spends traveling along this path, which is formally defined as follows.

Definition 4: Given an arbitrary path $l(u, v)$ between u and v, the delay of transmitting a message of size L over path $l(u, v)$ is

$$T(u, v) = \left\{\sum_{e_{ij}\in l(u,v)} T(L, e_{ij})\right\}, \tag{8}$$

where $T(L, e_{ij})$ is the delay of transmitting a message of size L over the link e_{ij} of the $l(u, v)$.

Based on the definition of transmission latency, we define the delay-bounded connectivity for the secondary networks.

Definition 5: A secondary network $G(\lambda_s, r)$ is connected if (i) $G(\lambda_s, r)$ is percolated, i.e., there exists a giant connected component C_∞ in $G(\lambda_s, r)$, and (ii) there exists at least one path between any secondary users u and v in this giant component so that the transmission latency over this path is of finite variance.

We first consider the case where the PU busy time follows light-tail distribution, i.e., exponential distributions. It follows easily from percolation theory [5] that Lemma 2 is true.

Lemma 2: If the secondary network $G(\lambda_s, r)$ has a node density λ_s larger than a constant λ^c/r^2, it can always achieve delay-bounded connectivity independent of the node density and the transmission range of the primary network.

Remark 4: λ^c is the critical density of the random graph of $G(\lambda, 1)$, which can represent a stand-alone ad hoc network, where the users are distributed according to Poisson distribution with density of λ and a transmission range of 1. λ^c is very difficult to find analytically, and recent simulation results show that $1.43 < \lambda^c < 1.44$ [14].

Remark 5: The intuitive explanation of this phenomenon is that, if the node density of the secondary network is large enough, then the secondary network will be percolated, i.e., there is a very high probability that at least one path exists between any two secondary users. Moreover, since the PU busy time is light tail distributed, the interference delay along each link of the path is always light tail distributed no matter what node density and transmission range the primary network will have. This indicates that the end-to-end transmission latency along the path is of finite variance. Therefore, the secondary network can always achieve delay-bounded connectivity.

Consider a secondary network coexisting with a primary network, where PU density $\lambda_p = 0.1$ and secondary user density $\lambda_s = 1.6$. All PUs have a transmission range $R = 1.2$, while secondary users have a transmission range $r = 1$. According to Lemma 2, since the SU density $\lambda_s = 1.6$ is larger than the critical one λ_c, the secondary network can achieve delay-bounded connectivity if the primary network does not exist or the PUs have LT busy time. As shown in Fig. 2, the giant connected component, represented by red dots, contains the majority of secondary users in the network.

Contrary to the above LT case, the following theorem shows that, in the presence of HT PU busy time, the existence of delay-bounded connectivity of the secondary network $G(\lambda_s, r)$ depends on the node density of the primary network.

Theorem 1: Assume that the secondary network $G(\lambda_s, r)$ has a node density $\lambda_s > \lambda^c/r^2$ and the busy periods of PUs are HT with tail index α, i.e., $B_1 \in \mathcal{RV}(\alpha)$. If $\alpha < 2$ and $\lambda_p > \lambda_p^* = \lambda^c/(4(R^2 - r^2/4))$, then the delay-bounded connectivity is not achievable in the secondary network $G(\lambda_s, r)$.

The above theorem indicates that the HT nature of PUs can induce infinite delay variance between a pair of secondary users even if their mutual distance is finite. In other words, a message sent by a secondary user can be only delivered to a small portion of nodes in the secondary network within bounded delay. The intuitive explanation of this phenomenon is that, since the busy

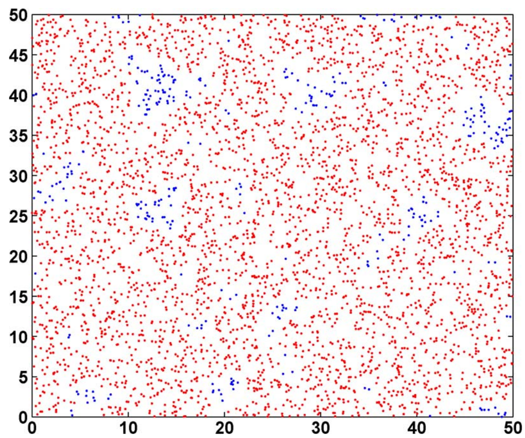

Fig. 2. Largest connected component (in red dots) of the secondary network.

periods of PUs are heavy tail distributed, the transmission time over the interfered links in $G(\lambda_s, r)$ could be HT and of infinite variance. Thus, as the density of PUs increases, more links in $G(\lambda_s, r)$ will exhibit HT behavior so that all the paths between two secondary users u and v will have at least one link with the transmission time of infinite variance. This implies the infiniteness of $Var[T(u, v)]$.

However, when the secondary users are mobile, there is a probability that the transmission latency $T(u, v)$ can have finite mean and variance even if the busy periods of PUs are heavy tail distributed. For instance, when a secondary user u broadcasts a message, all the secondary users that reside inside the connected component of u and outside the interference range of any PUs can receive the message within finite average delay. As time goes on, nodes move, and the message is passed from message-carrying nodes to new nodes whenever they are within communication range and outside the interference range of any PUs. As this process goes on, the message can be delivered within the whole network without being affected by the HT activities of the PUs.

It is worth to mention that utilizing opportunistic contacts of mobile users for mobile data content exchange has been widely investigated under the context of delay-tolerant networking or mobile opportunistic networking. So far, a lot of opportunistic routing protocols have been proposed, which aim to find the optimal forwarding strategy that selects the best message-carrying node to establish a suitable data forwarding among all possible paths so that the desired network performance can be achieved. However, in this work, our intention is not to design a new opportunistic routing protocol. Instead, we focus on the general design guideline, namely, the maximum mobility radius, which all opportunistic routing protocols can follow. Specifically, we show that, above such maximum mobility radius, even under challenged HT interference, there will be at least one possible path with bounded forwarding delay between any two nodes in the network independent of the routing algorithms.

In this paper, we consider that the secondary users can move around in a circular region centered at their initial locations, where the maximum radius that the secondary user can reach is denoted by a. For theoretical analysis, we assume that the secondary users will follow an i.i.d. mobility model [7], [12] such that, at the beginning of each time slot, the secondary user moves to a randomly and uniformly selected location within the circular region and the positions of secondary users are independent of each other and independent over time slots. More specifically, the i.i.d. mobility model can be formally defined as follows.

Definition 6 (i.i.d. Mobility Model): Given the initial locations $\{X_1^0, X_2^0, \ldots, X_n^0\}$ of secondary users, $u = 1, 2, \ldots, n$ at time 0. At each time $t = 1, 2, \ldots$, the location X_u^t of the secondary user u is uniformly distributed in $A(X_u^0, a)$, which is the circular region centered at the initial location X_u^0 of u with radius $a > 0$. The positions X_u^t are mutually independent among all secondary users and independent of all previous locations $X_u^{t'}, t = 0, 1, \ldots, t-1$.

The following theorem states the sufficient condition on the maximum radius a that the secondary user should reach, under which the secondary network can achieve delay-bounded connectivity. To distinguish from the static secondary network $G(\lambda_s, r)$, we denote the mobile secondary network by $G_m(\lambda_s, r)$.

Theorem 2: If the maximum radius a that the secondary user can reach is larger than the threshold value $a^* < \infty$, where

$$a^* = \arg\min_{a>0}\left(\left(1-e^{-a^2/5\lambda_s}\right)^2\left(1-\left(1-e^{-|R^I|\lambda_p}\right)^{N^I}\right) > \frac{5}{6}\right)$$

where $|R^I| = 2(r/\sqrt{5}+R)(2R + r/\sqrt{5})$ and $N_I = \lfloor (2/5)a^2/R^I \rfloor$, then the mobile secondary network $G_m(\lambda_s, r)$ achieves delay-bound connectivity such that there exists a giant component C_∞ containing an infinite number of secondary users, in which the transmission delay between any secondary users u and v with $\|u - v\| < \infty$ is of finite mean and variance, i.e., $E[T(u, v)] < \infty, \forall\, u, v \in C_\infty$ and $Var[T(u, v)] < \infty, \forall\, u, v \in C_\infty$

Remark 6: By analyzing the expression a^*, it is easy to see that a^* always exists under any user densities, i.e., λ_s and λ_p, the transmission range r of secondary users, and the interference range R of PUs. This means that, by properly adjusting the mobility radius a, finite mean transmission latency can be always achieved. In other words, the maximum transmission distance d^c (shown in Theorem 1) induced by HT spectrum activities vanishes by exploiting the mobility of secondary users.

Now, assume that the mobile secondary network $G_m(\lambda_s, r)$ achieves delay-bounded connectivity. Consider two secondary users u and v that are connected in $G_m(\lambda_s, r)$. When u broadcasts a message of size L, this message can be delivered to v through different paths. We define the first-passage latency between u and v as the first time that v receives the message, which is formally defined as follows.

Definition 7: Consider any two secondary users u and v. The first-passage latency from u to v is

$$T_p(u, v) = \inf_{l(u,v)\in\mathcal{P}(u,v)} \left\{ \sum_{e_{ij}\in l(u,v)} T(L, e_{ij}) \right\} \quad (9)$$

where $l(u, v)$ is an arbitrary path between u and v and $\mathcal{P}(u, v)$ is the set of paths from u and v.

The following theorem characterizes the first-passage latency between the SU u and v.

Theorem 3: For any secondary users $u, v \in C_\infty$, $T(u, v)$ scales linearly with the Euclidean distance between u and v as the distance approaches infinity, i.e.,

$$\Pr\left\{ \lim_{\|u-v\|\to\infty} \frac{T_p(u, v)}{\|u - v\|} = \rho' \right\} = 1 \quad (10)$$

where ρ' is a strictly positive value.

IV. Delay-Bounded Connectivity Under HT PU Activities

In this section, we prove Theorem 1 which shows that, under the HT activities of PUs, if PU density $\lambda_p > \lambda_p^* = \lambda^c/4(R^2 - r^2/4)$ and the busy periods of PUs are HT with tail index $\alpha < 2$, the secondary network cannot achieve delay-bounded connectivity with bounded end-to-end delay variance. Equivalently, we show that, given $\alpha < 2$, there always exists a positive distance d^c such that, for any secondary users u and v with their mutual distance larger than d^c, i.e., $\|u - v\| > d^c$, then the end-to-end latency between u and v is of infinite variance, i.e., $Var[T(u, v)] = \infty$.

Consider a secondary user u in the giant component C_u of $G(\lambda_s, r)$. Assume $\alpha < 2$. It follows by Lemma 1 that the transmission delay of the secondary users within the interference region of the PUs is of infinite variance. Therefore, to prove that $Var[T(u, v)] = \infty\, \forall\, v \in C_u$ with $\|u - v\| > d^c$, it is sufficient to prove that there exists a continuous interference region surrounding u such at all paths starting from u and ending at any v with $\|u - v\| < d^c$ are disconnected by this interference region. By (3), to disconnect a path between u and v, the possible narrowest width of this interference region should be larger than the transmission radius r of the secondary user. Thus, the basic idea of the proof is to show that, if condition $\lambda_p > \lambda_p^* = \lambda^c/4(R^2 - r^2/4)$ is satisfied, such interference region exists almost surely. Since the secondary network $G(\lambda_s, r)$ is induced by a homogeneous Poisson point process, all the nodes are probabilistically indistinguishable. We choose an arbitrary node in $G(\lambda_s, r)$ as the source node u.

We start by placing a square lattice S on $\mathbb{R}^2$, with edge length d_p. Consider a sequence $\{G_i\}_{i\geq 1}$ of annuli around the origin. Each annulus G_i is made up of four rectangles

$$\begin{aligned} A_i^+ &= [-d_p 2^i, d_p 2^i] \times [d_p 2^{i-1}, d_p 2^i] \\ A_i^- &= [-d_p 2^i, d_p 2^i] \times [-d_p 2^i, -d_p 2^{i-1}] \\ B_i^+ &= [d_p 2^{i-1}, d_p 2^i] \times [-d_p 2^i, d_p 2^i] \\ B_i^- &= [-d_p 2^i, -d_p 2^{i-1}] \times [-d_p 2^i, d_p 2^i]. \end{aligned} \quad (11)$$

Fig. 4(a) shows examples of annuli G_1 and G_2, respectively, and Fig. 4(b) illustrates the arrangement of the four rectangles of annulus G_1. According to (11), G_1 is made of the rectangles $A_1^+ = [-2d_p, 2d_p] \times [d_p, 2d_p]$, $A_1^- = [-2d_p, 2d_p] \times [-2d_p, -d_p]$, $B_1^+ = [d_p, 2d_p] \times [-2d_p, 2d_p]$, and $B_1^- = [-2d_p, -d_p] \times [-2d_p, 2d_p]$. Similarly, G_2 is made of the rectangles $A_2^+ = [-4d_p,$

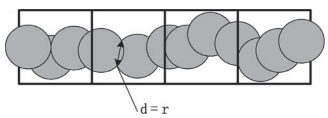

Fig. 3. Closed rectangle with the narrowest width equal to the transmission radius r.

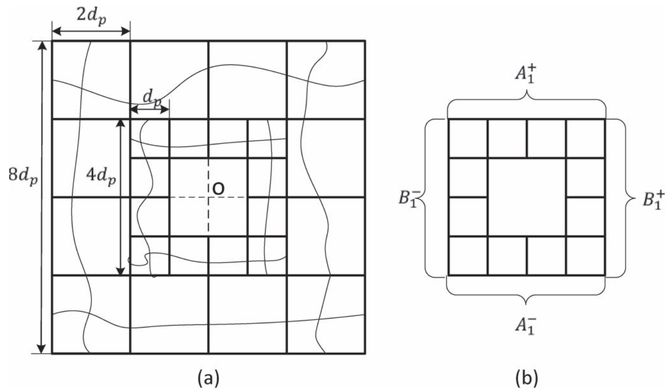

Fig. 4. (a) Annuli (inside) G_1 and (outside) G_2. Each annulus has four closed (crossed) rectangles. (b) Arrangement of the four rectangles of annulus G_1.

$4d_p] \times [2d_p, 4d_p]$, $A_2^- = [-4d_p, 4d_p] \times [-4d_p, -2d_p]$, $B_2^+ = [2d_p, 4d_p] \times [-4d_p, 4d_p]$, and $B_2^- = [-4d_p, -2d_p] \times [-4d_p, 4d_p]$.

For each rectangle, we define the crossing events as follows.

Definition 8 [5]: A rectangle $R = [x_1, x_2] \times [y_1, y_2]$ being crossed from left to right by a connected component in $G(\lambda_s, r)$ means that there exists a sequence of nodes $v_1, v_2, \ldots, v_m \in G(\lambda_s, r)$ contained in R, with $\|X_{v_i} - X_{v_j}\| < 2\sqrt{R^2 - r^2/4}$, $i = 1, 2, \ldots, m-1$, and $0 < X_{v_1} - x_1 < \sqrt{R^2 - r^2/4}$, $0 < x_2 - X_{v_m} < \sqrt{R^2 - r^2/4}$, where X_{v_1} and X_{v_m} are the x-coordinates of the most left node v_1 and the most right v_m, respectively. A rectangle being crossed from top to bottom can be defined analogously.

We define that $A_i^+(A_i^-)$ is closed if $A_i^+(A_i^-)$ is crossed from left to right by a connected component in the primary network $G(\lambda_p, R)$, as illustrated in Fig. 3. Similarly, we declare that $B_i^+(B_i^-)$ is closed if $B_i^+(B_i^-)$ is crossed from bottom to top by a connected component in $G(\lambda_p, R)$. The structure of the annulus is shown in Fig. 4.

By Definition 8, it is clear that, if A_i^+ is closed, the primary network $G(\lambda_p, R)$ generates a continuous interference region in A_i^+ with the possible narrowest width equal to the transmission radius r of the secondary users. Let $\tilde{A}_i^+$, $\tilde{A}_i^-$, $\tilde{B}_i^+$, and $\tilde{B}_i^-$ be the events that A_i^+, A_i^-, B_i^+, and B_i^- are closed, respectively. According to Russo–Seymour–Welsh (RSW) theorem and the scaling property of random geometric graphs [11], when $\lambda_p > \lambda^c/4(R^2 - r^2/4)$, i.e., $G(\lambda_p, R)$ is in the supercritical phase, we can choose d_p to be large enough so that events $\tilde{A}_i^+$, $\tilde{A}_i^-$, $\tilde{B}_i^+$, and $\tilde{B}_i^-$ occur with the probability arbitrarily close to 1. This means that, for any $0 < \delta < 1$, there always exists d'_p so that, if $d_p \geq d'_p$, $\Pr(\tilde{A}_i^+) = \Pr(\tilde{A}_i^-) = \Pr(\tilde{B}_i^+) = \Pr(\tilde{B}_i^-) \geq \delta$.

If events $\tilde{A}_i^+$, $\tilde{A}_i^-$, $\tilde{B}_i^+$, and $\tilde{B}_i^-$ occur simultaneously, the annulus G_i must contain a continuous interference region generated by the PUs, and hence, all the paths starting from u are necessarily confined within the outer boundary of G_i. Denote the latter event by $\tilde{G}_i$. Since $\tilde{A}_i^+$, $\tilde{A}_i^-$, $\tilde{B}_i^+$, and $\tilde{B}_i^-$ are dependent [11], which means that they are positively correlated, utilizing the Fortuin–Kasteleyn–Ginibre (FKG) inequality [11] yields

$$\begin{aligned}\Pr(\tilde{G}_i) &= \Pr\left(\tilde{A}_i^+ \cap \tilde{A}_i^- \cap \tilde{B}_i^+ \cap \tilde{B}_i^-\right) \\ &\geq \Pr\left(\tilde{A}_i^+\right)\Pr\left(\tilde{A}_i^-\right)\Pr\left(\tilde{B}_i^+\right)\Pr\left(\tilde{B}_i^-\right) \geq \delta^4.\end{aligned}$$

Thus, we have $\sum_{i=1}^{\infty} \Pr(\tilde{G}_i) \geq \sum_{i=1}^{\infty} \delta^4 = \infty$. Since the construction of the annuli $\{G_i\}_{i\geq 1}$ guarantees that events $\{G_i\}_{i\geq 1}$ are independent, by the Borel–Cantelli lemma [11], there exists $j < \infty$ so that $\tilde{G}_j$ occurs with the probability of 1. Since the outer boundary of annulus G_j is $[-d_p 2^j, d_p 2^j] \times [-d_p 2^j, d_p 2^j]$, this means that there exists a finite value $d^c = d_p 2^j$ such that $E[T(u, v)] = \infty \; \forall \, v \in C_u$ with $\|u - v\| > d^c$.

V. Mobility Improves Delay-Bounded Connectivity

In this section, we prove Theorem 2 to show that mobility can help to achieve delay-bounded latency. To prove Theorem 2, we transform a random network with mobile nodes $G_m(\lambda_s, r)$ into a random static network with stationary nodes $G(\lambda_s)$. The node positions in the latter static network are given by the initial node positions of the former mobile network. A link between a node u and v in the latter static network exists if, in the former mobile network, node u and v can exchange the message within finite delay variance, without being assisted by other mobile nodes. Let $T_{uv}(L)$ define this message exchanging time between u and v without being assisted by other mobile nodes. We formally define $G(\lambda_s)$ as follows.

Definition 9: Given the initial locations $\mathcal{X}^0 = \{X_1^0, X_2^0, \ldots, X_n^0\}$ of secondary users $u = 1, 2, \ldots, n$ at time 0 and the maximum mobility radius a of the secondary users, let $G(\lambda_s)$ be a static random graph, in which secondary users are located at $\mathcal{X}'$ and a links exists between u and v if the message exchanging time between u and v is of finite mean, i.e., $E[T_{uv}(L)] < \infty$.

Remark 7: Note that, in the definition above, we require that a link exists between u and v if the message exchanging time between them is of finite mean. Similarly, we can also enforce that a link exists between u and v if the message exchanging time between them is of finite variance. In both cases, the proofs follow the similar procedures.

Now, proving Theorem 2 is equivalent to deriving the conditions on the maximum mobility radius a under which the static random graph $G(\lambda_s)$ is percolated. Towards this, we first perform the following mapping processes as illustrated in Fig. 5 to place a specially constructed square lattice $\mathcal{L}$ on the plane $\mathbb{R}^2$. Then, utilizing such lattice, we will prove Theorem 2 by translating the presence of continuum percolation of $G(\lambda_s)$ on the plane $\mathbb{R}^2$ into the presence of bond percolation on the lattice.

The square lattice $\mathcal{L}$ has the edge length D. All the vertices of L are located at $(D \times i, D \times j)$, where $(i, j) \in \mathbb{Z}$. We choose the edge length $D = a/\sqrt{5}$, where a is the maximum radius that a mobile secondary user can reach. Any two squares

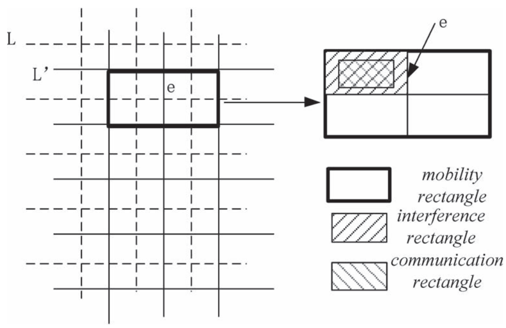

Fig. 5. Lattice $\mathcal{L}'$ and its dual $\mathcal{L}$ with four interference rectangles in one mobility rectangle.

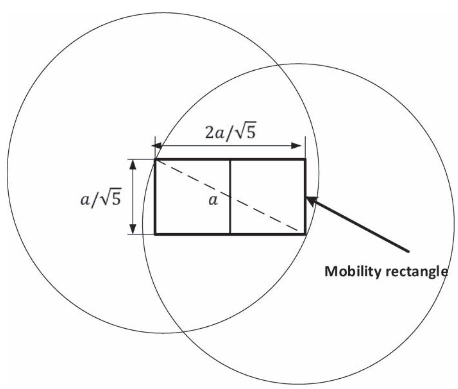

Fig. 6. Mobility rectangle.

adjacent to an edge form a *mobility rectangle*, and each mobility rectangle is evenly divided into multiple *interference rectangles* with dimension $2(r/\sqrt{5}+R)\times(2R+r/\sqrt{5})$, where r is the transmission radius of the secondary user and R is the interference range of the PU. At the center of each *interference rectangle*, we place a *communication rectangle* with dimension $2r/\sqrt{5}\times r/\sqrt{5}$.

Based on Lemma 3, we can determine whether there exists a link between any two nodes u and v in $G(\lambda_s)$ so that the topology of $G(\lambda_s)$ can be characterized. Such topology information will be utilized in the following sections to prove the percolation condition of $G(\lambda_s)$.

Lemma 3: Given two secondary users u and v of $G_m(\lambda_s, r)$, if the following two conditions are met, i.e.,

1) The initial location u and v at time 0 are within the same mobility rectangle.
2) At least one of the interference rectangles in the corresponding mobility rectangle has no PUs.

then there is a link between u and v in $G(\lambda_S)$.

Proof: To prove the above lemma, we show that the time $T_{uv}(L)$ of exchanging a message of size L between secondary users u and v without being assisted by other mobile secondary users is of finite mean. As shown in Fig. 5, the edge length of the lattice is $a/\sqrt{5}$. According to Fig. 6, this ensures that, if two secondary users u and v are in the mobility region at time 0, where the first condition of the above lemma is satisfied, satisfying this condition implies that there exists a positive probability that, as time proceeds, u and v will be within the transmission range of each other. Moreover, by our lattice construction, each mobility rectangle is divided into multiple interference rectangles, and each interference rectangle contains a communication rectangle. Therefore, if u and v are in the mobility rectangle at time 0, at each time slot, the probability that they can be simultaneously within a particular communication rectangle i is $p=2r^2/5\pi R^2$ because of the constrained i.i.d. mobility model. This means that the first time $T_1(v,u)$ that u and v are within this communication rectangle follows geometric distribution with mean $E[T_1(v,u)]=p$. Assume that the interference rectangle containing communication rectangle i has no PUs, which means that the second condition of the lemma is satisfied. This implies by the construction of the interference rectangle that there exists no PU interference in the communication rectangle i. Consequently, the mean delay $E[T_{tr}(u,,v)]$ for transmitting one packet between u and v is less than $E[T_1(v,u)]=p$, i.e., $E[T_{tr}(u,,v)]\le E[T_1(v,u)]$. This is due to the fact that the size of the communication rectangle is $2r/\sqrt{5}\times r/\sqrt{5}$, and thus, any two secondary users residing in the communication rectangle are necessarily within the transmission range of each other. Assume that the mean size of the message L is finite, i.e., $E[L]<\infty$. This implies that the mean transmission delay for message L is $E[T_{uv}(L)]\le E[L]E[T_{tr}(u,v)]\le E[L]p$. This completes the proof. □

A. Proof of Theorem 2

In this section, we prove Theorem 2 based on the random static network $G(\lambda_s)$ characterized by Definition 9 and Lemma 3. Theorem 2 states that, if the maximum mobility radius a is larger than threshold a^*, then the mobile secondary network $G_m(\lambda_s, r)$ achieves delay-bounded connectivity. To prove Theorem 2-1 is sufficient to prove the following two lemmas.

1) (Lemma 4) If the moving radius a of secondary users is larger than a threshold a^*, then $G(\lambda_s)$ is percolated, i.e., there is a giant connected component in $G(\lambda_s)$.
2) (Lemma 5) If $G(\lambda_s)$ is percolated, then the minimum transmission latency $T(u,v)$ between any two secondary users u and v inside the giant connected component of $G(\lambda_s)$ is of finite mean.

In the rest of this section, we state and prove these two lemmas, respectively.

Lemma 4: Given a secondary network $G(\lambda_s,)$, there exists a strictly positive value $a^*<\infty$, i.e.,

$$a^*=\arg\min_{a>0}\left(\left(1-e^{-a^2/5\lambda_s}\right)^2\left(1-\left(1-e^{-|R^I|\lambda_p}\right)^{N^I}\right)>\frac{5}{6}\right)$$

where $|R^I|=2(r/\sqrt{5}+R)(2R+r/\sqrt{5})$ and $N_I=\lfloor(2/5)a^2/R^I\rfloor$ such that, if $a>a^*$, then $G(\lambda_s)$ is percolated.

Before giving the proof of Lemma 4, we introduce several useful definitions.

Definition 10: A vertical edge e of $\mathcal{L}$ in Fig. 5 is said to be open if the following conditions are satisfied.

1) At initial time $t=0$, both squares adjacent to e contain at least one secondary user.
2) In the mobility region formed by the adjacent two squares, there exists at least one of the interference rectangles, which does not contain any PUs.

Define similarly the open horizontal edge of $\mathcal{L}$ by rotating the rectangle R_e by 90°. Next, we construct the dual lattice of $\mathcal{L}$, which is denoted by $\mathcal{L}'$. $\mathcal{L}'$ is obtained from $\mathcal{L}$ in the following way. A vertex is placed in the center of each square of $\mathcal{L}$, and two such neighboring vertices are joined by a straight line segment. This line segment becomes an edge of $\mathcal{L}'$. As $\mathcal{L}$ is a square lattice, the dual lattice $\mathcal{L}'$ is the same lattice shifted by $(D/2, D/2)$.

Definition 11: An edge of $\mathcal{L}'$. is said to be open if and only if its corresponding edge of $\mathcal{L}$ is open.

Definition 12: A path (in $\mathcal{L}$ or $\mathcal{L}'$) is said to be open if and only if all its edges are open; a path (in $\mathcal{L}$ or $\mathcal{L}'$) is said to be close if and only if all its edges are close.

Proof of Lemma 4: The basic idea of the proof for Lemma 4 is to translate the presence of continuum percolation on $G(\lambda_s)$ into the presence of bond percolation on the lattice $= \mathcal{L}'$. More specifically, we first show that the secondary network $G(\lambda_s)$ will have an infinite connected component on the continuous plane $\mathbb{R}^2$ if bond percolation occurs on $\mathcal{L}'$, i.e., if there exists an infinite open path on $\mathcal{L}'$.

Let E_1 and E_2 be the events when the conditions 1) and 2) in Definition 10 are satisfied, respectively. Let C_e denote the event that an edge e is closed. The probability that C_e occurs is upper bounded by

$$\begin{aligned}\Pr(C_e) &= 1 - \Pr(E_1 \cap E_2) \overset{a}{=} 1 - \Pr(E_1)\Pr(E_2)\\ &= 1 - \left(1 - e^{-D^2\lambda_s}\right)^2 \left(1 - \left(1 - e^{-|R^I|\lambda_p}\right)^{N^I}\right)\end{aligned}$$

where $D = a/\sqrt{5}$, $|R^I| = 2(r/\sqrt{5} + R)(2R + r/\sqrt{5})$ is the area of the interference region R_I and $N_I = \lfloor 2D^2/R^I \rfloor$ is the number of interference rectangles in the mobility region. Equality a in (12) comes from the independence of the locations of the primary and secondary users.

Now, let us consider a path $P_n = \{e_i\}_{i=1}^n$ of length n in L. Because the states (i.e., open or closed) of any set of non-adjacent edges are independent, there exist at least $m \geq n/2$ edges in P_n, e.g., $\{e_j\}_{j=1}^m \subseteq \{e_i\}_{i=1}^n$, such that their states are independent from each other. Let X_{e_i} denote the event that e_i is closed. Then, the probability that the path P_n is closed is upper bounded by

$$\Pr(\mathit{closed}\,\mathrm{P}_n) = \Pr\left(\bigcap_{i=1}^n X_{e_i}\right) \leq \Pr\left(\bigcap_{i=1}^m X_{e_i}\right) \leq q^{\frac{n}{2}}, \tag{12}$$

where $q = \Pr(C_e)$ as given in (12).

By the duality between $\mathcal{L}$ and $\mathcal{L}'$, if an open path starting from a vertex (e.g., the origin) in $\mathcal{L}'$ is finite, the origin is necessarily surrounded by a closed circuit (a closed path with the same starting and ending vertex) in the dual lattice L. Hence, by letting the latter event be O_L, the probability that there exists an infinite open path starting from the origin is $1 - \Pr(O_L)$. Furthermore, from (12), we have

$$\Pr(O_\mathcal{L}) = \sum_{n=2}^{\infty} \sigma(n)\Pr(\mathit{closed}\,\mathrm{P}_{2n}) \leq \sum_{n=2}^{\infty} \sigma(n) q^n, \tag{13}$$

where $\sigma(n)$ is the number of closed circuits of the length $2n$ surrounding the origin. It is easy to show that $\sigma(n)$ is upper bounded by

$$\sigma(n) \leq (n-1)3^{2(n-1)}. \tag{14}$$

Hence, we have

$$\Pr(O_\mathcal{L}) = \sum_{n=2}^{\infty}(n-1)3^{2(n-1)}q^n = \frac{9q^2}{(1-9q)^2}. \tag{15}$$

Therefore, from (15) and (12), if $q = \Pr(C_e) < 1/6$ or $a > a*$, where

$$a^* = \arg\min_{a>0}\left(\left(1 - e^{-a^2/5\lambda_s}\right)^2\left(1 - \left(1 - e^{-|R^I|}\lambda_p\right)^{N^I}\right) > \frac{5}{6}\right)$$

then $\Pr(O_\mathcal{L})$ converges to a number less than one. As a consequence, the probability that there exists an infinite open path starting from the origin in $\mathcal{L}'$ is positive. According to Kolmogorov's zero–one law, this implies that an infinite path exists in $\mathcal{L}'$ with a probability of one.

We now consider an infinite open path $P_\infty = \{e_i'\}_{i=1}^\infty$ in $\mathcal{L}'$. Along each edge e_i', there exists two adjacent squares in the dual lattice $\mathcal{L}$. Therefore, along P_∞, there exists a sequence of squares $\{S_i\}_{i=1}^\infty$ in $\mathcal{L}$ such that any two consecutive squares, denoted by S_i and S_{i+1}, are adjacent. By Definitions 10 and 11, the region comprising S_i and S_{i+1} contains at least two mutually connected SUs that belong to $G(\lambda_s)$. Thus, the sequence of squares $\{S_i\}_{i=1}^\infty$ forms an infinite connected component in $G(\lambda_s)$, which indicates that $G(\lambda_s)$ is percolated. □

Now, we are ready to prove the first part of Theorem 2. We first introduce some useful notations. We denote by C_∞ the infinite connected component in $G(\lambda_s)$. For each coordinate $(i, 0)$ on the square lattice $\mathcal{L}$ with $i \in Z$, denote the location of the nearest secondary users in C_∞ by $\widetilde{X}_i$, i.e., $\widetilde{X}_i = \arg\min_{X_j \in C_\infty} \|X_j| - (i,0)\|$. Let $T_{m,n} = T(\widetilde{X}_m, \widetilde{X}_n)$.

Now, we assume $a > a^*$ which means that $G(\lambda_s)$ is percolated. Then, to prove Theorem 2-1 is sufficient to prove following lemma.

Lemma 5: Given two SUs $\widetilde{X}_0$ and $\widetilde{X}_n$ within the giant connected component in $G(\lambda_s, r)$, if the mutual distance $d_{0,n} = \|\widetilde{X}_0 - \widetilde{X}_n\| < \infty$, then $E(|T_{0,n}|) = E(|T(\widetilde{X}_0, \widetilde{X}_n)|) < \infty$.

Proof: To compute the upper bound of $E(|T_{0,n}|)$, we consider the shortest path (in links) $l_{0,n}$ from $\widetilde{X}_0$ to $\widetilde{X}_n$. Denote the number of hops or links on such a path by $|l_{0,n}|$ and the delay on each hop i by T_i. Since the smallest delay $T_{0,n}$ cannot be greater than the delay on any particular path, $E(|T_{0,n}|)$ is upper bounded by $E(|T_{0,n}|) \leq E(\sum_{i=1}^{|l_{0,n}|} T_i) = E(T_i)E(|l_{0,n}|)$.

By the construction of $G(\lambda_s)$ and Lemma 3, $E(T_i) < \infty$. Therefore, to show $E(|T_{0,n}|) < \infty$ is sufficient to prove $E(|l_{0,n}|) < \infty$. Let $v_{\widetilde{X}_0}$ denote the vertex of $\mathcal{L}'$, which is closest to $\widetilde{X}_0$, and let $v_{\widetilde{X}_n}$ denote the vertex of $\mathcal{L}'$, which is closest to $\widetilde{X}_n$. By our construction of $G(\lambda_s)$, the shortest path $l_{0,n}$ between $\widetilde{X}_0$ and $\widetilde{X}_n$ is the open path on $\mathcal{L}'$ consisting of the smallest number of edges. To evaluate $|l_{0,n}|$, we construct a new square lattice on the top of square lattice $\mathcal{L}'$ (shown in Fig. 5). This new square lattice has the edge length $d_{0,n} = nL_D$, where L_D is the edge length of $\mathcal{L}'$. Then, a sequence annuli $\{G_i(d_{0,n})\}_{i\geq 1}$ around the origin is constructed in the same way as illustrated in Fig. 4. More specifically, each annulus $G_i(d_{0,n})$ is made up of four rectangles $A_i^+(d_{0,n})$, $A_i^-(d_{0,n})$, $B_i^+(d_{0,n})$, and $B_i^-(d_{0,n})$, according to (11), in which we substitute $d_{0,n}/2$ for d_p. We define the crossing events for the rectangles as follows.

Definition 13 [5]: A rectangle $[x_1, x_2] \times [y_1, y_2]$ is crossed from left to right if $[x_1, x_2] \times [y_1, y_2]$ contains an open path on the square lattice $\mathcal{L}'$ that joins the left and right borders of $[x_1, x_2] \times [y_1, y_2]$. A rectangle $[x_1, x_2] \times [y_1, y_2]$ being crossed from top to bottom can be defined analogously.

Let $\widetilde{A}_i^+(d_{0,n})$, $\widetilde{A}_i^-(d_{0,n})$, $\widetilde{B}_i^+(d_{0,n})$, and $\widetilde{B}_i^-(d_{0,n})$ be the event that $A_i^+(d_{0,n})$, $A_i^-(d_{0,n})$, $B_i^+(d_{0,n})$, and $B_i^-(d_{0,n})$ are open, respectively. By RSW theorem for independent bound percolation [5], if the states of edges of $\mathcal{L}'$ are independent and the open probability of the edge is larger than 1/2, for any $0 < \delta < 1$, there always exists $i < \infty$ such that $P(\widetilde{A}_i^+(d_{0,n})) > \delta$. Although the adjacent edges of $\mathcal{L}'$ are dependent, by Definitions 10 and 11, the open states of adjacent edges in $\mathcal{L}'$ are increasing events. Moreover, by the proof of Lemma 4, the open probability of the edge is larger than 5/6, provided that $a > a^*$. This implies that, if $a > a^*$, $P(\widetilde{A}_i^+(d_{0,n})) > \delta$ still holds. Let $\widetilde{G}_i(d_{0,n})$ denote the event that all the four rectangles of an annulus $G_i(d_{0,n})$ are open. Since $\widetilde{A}_i^+(d_{0,n})$, $\widetilde{A}_i^-(d_{0,n})$, $\widetilde{B}_i^+(d_{0,n})$, and $\widetilde{B}_i^-(d_{0,n})$ are increasing events, by the FKG inequality [11], we have $\Pr(\widetilde{G}_i(d_{0,n})) = \Pr(\widetilde{A}_i^+(d_{0,n}) \cap \widetilde{A}_i^-(d_{0,n}) \cap \widetilde{B}_i^+(d_{0,n}) \cap \widetilde{B}_i^-(d_{0,n})) \geq \delta^4$.

If all the four rectangles of an annulus $G_i(d_{0,n})$ are open, i.e., $\widetilde{G}_i(d_{0,n})$ occurs, then the annulus $G_i(d_{0,n})$ contains an open circuit. Consequentially, the shortest path $L_{0,n}$ is necessarily included in the square $[-d_{0,n}2^i, d_{0,n}2^i] \times [-d_{0,n}2^i, d_{0,n}2^i]$. Thus, the number of hops on $l_{0,n}$, e.g., $|l_{0,n}|$, can be upper bounded by a certain value. Since $l_{0,n}$ is included in a square of the size of $4^i d_{0,n}^2$, $|l_{0,n}|$ is upper bounded by $4^i n^2 + 2^i n$. Therefore, we have

$$\Pr\left(|l_{0,n}| > 4^i n^2 + 2^i n\right) < \Pr\left(\bigcap_{j=1}^{i} \widetilde{G}_j(d_{0,n})\right) < (1-\delta^4)^i.$$

$E(|L_{0,n}|)$ can be upper bounded by

$$\begin{aligned} E\left(|l_{0,n}|\right) &= \sum_{k=0}^{N-1} \Pr\left(|l_{0,n}|\right) > k + \sum_{k=N}^{\infty} \Pr\left(|l_{0,n}|\right) > k \\ &\leq 4^{N-1}n^2 + 2^{N-1}n + \sum_{k=N}^{\infty} \frac{4^k n^2 + 2^k n}{(1-\delta^4)^{-k}} < \infty \end{aligned}$$

where N is the minimum value of the annulus index i such that $\Pr(\widetilde{A}_i^+(d_{0,n}) > \delta > (3/4)^{1/4}$. This completes the proof. □

B. Proof of Theorem 3

Next, we prove Theorem 3 regarding the linear scaling property of the first-passage latency $T_p(u, v)$ with the Euclidean distance between u and v as the distance approaches ∞.

The proof of Theorem 2-2 relies on the following lemma.

Lemma 6:

$$\lim_{n\to\infty}\left(\frac{T_{0,n}}{n}\right) = \rho \tag{16}$$

with a probability of one, where $\rho = \lim_{n\to\infty}(E(T_{0,n})/n) = \inf_{n\geq 1}(E(T_{0,n})/n)$.

The main tools to prove Lemma 6 is Liggett's subadditive ergodic theorem [9], which states as follows.

Theorem 4 (Subadditive Ergodic Theorem): Let $\{T_{m,n}\}$ be a collection of rvs indexed by integers satisfying $0 \leq m < n$. Suppose that $\{T_{m,n}\}$ has the following properties: (i) $T_{0,n} \leq T_{0,m} + T_{m,n}$. (ii) The distribution of $\{T_{m,m+k} : k \geq 1\}$ does not depend on m. (iii) $\{T_{nk,(n+1)k} : n \geq 0\}$ is a stationary sequence for each $k \geq 1$. (iv) $E(|T_{0,n}|) < \infty$ for each n. Then, (a) $\eta = \lim_{n\to\infty}(E(T_{0,n})/n) = \inf_{n\geq 1}(E(T_{0,n})/n)$. (b) $T = \lim_{n\to\infty}(T_{0,n}/n)$ with a probability of one, and $E(T) = \eta$. Furthermore, (v) if, for $k \geq 1$, $\{T_{nk,(n+1)k} : n \geq 0\}$ are ergodic, then (c) $T = \eta$.

Evidently, if all above conditions are satisfied, Lemma 6 is proved. It is easy to check that condition (i) is satisfied because $T_{0,n}$ is defined as the smallest transmission latency between the SUs located at $\widetilde{X}_m$ and $\widetilde{X}_n$, and it is easy to see that $T_{0,n}$ cannot exceed $T_{0,m} + T_{m,n}$; otherwise, $T_{0,n}$ is not the smallest transmission latency. Moreover, the conditions (ii) and (iii) are clearly fulfilled because of the stationarity of the homogeneous Poisson point process. Moreover, condition (iv) is met because of Lemma 5. To show that condition (v) of Theorem 4 is satisfied, as in [3], we prove that $\{T_{nk,(n+1)k} : n \geq 0\}$ is asymptotically independent, which is a stronger statement than the statement that $\{T_{nk,(n+1)k} : n \geq 0\}$ is ergodic.

Lemma 7: The sequence $T_{nk,(n+1)k} : n \geq 0$ is strong mixing.

Proof: Similar to the proof of Lemma 5, we place two annuli centered at $(nk, 0)$ and $((n+m)k, 0)$, respectively. Since $G(\lambda_s)$ is percolated, by the Borel–Cantelli lemma, there always exists an annulus containing an open circuit such that the path with the shortest latency from $\widetilde{X}_{nk}$ to $\widetilde{X}_{(n+1)k}$ is circumscribed within a square S_1 with finite edge length $2^{i+1}d$ and the corresponding path from $\widetilde{X}_{(n+m)k}$ to $\widetilde{X}_{(n+m+1)k}$ is circumscribed within a square S_2 with finite edge length $2^{j+1}d$. As m goes to infinity, the two squares are not overlapping, which means that the path from $\widetilde{X}_{nk}$ to $\widetilde{X}_{(n+1)k}$ does not share any links with the path from $\widetilde{X}_{(n+m)k}$ to $\widetilde{X}_{(n+m+1)k}$. Thus, $T_{nk,(n+1)k}$ and $T_{(n+m)k,(n+m+1)k}$ are asymptotically independent, i.e.,

$$\begin{aligned} &\lim_{m\to\infty} \Pr\left(T_{nk,(n+1)k} < t \cap T_{(n+m)k,(n+m+1)k} < t\right) \\ &\quad = \Pr\left(T_{nk,(n+1)k} < t\right) \cap \Pr\left(T_{(n+m)k,(n+m+1)k} < t\right). \end{aligned}$$

This completes the proof. □

Proof of Theorem 2-2: Without loss of generality, take a straight line passing through X_u and X_v as the x-axis. Consider X_u as the origin. This means that $X_u = \widetilde{X}_0$. We denote by n the integer closest to the x-axis coordinate of X_v, which implies that $\|X_v - (n, 0)\| < 1/2$ and $\|X_u - X_v\| < n + (1/2)$.

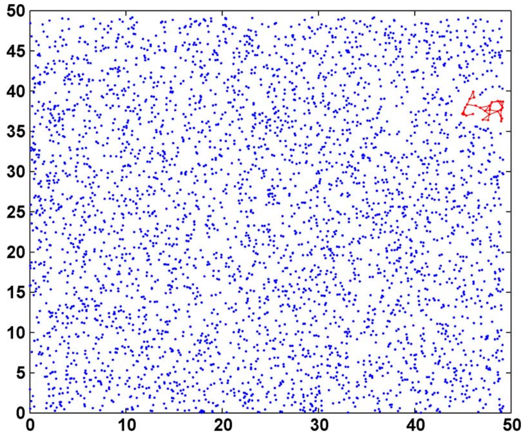

Fig. 7. Largest connected component (in red dots) of a secondary network ($\lambda_s = 1.6$) coexisting with a primary network ($\lambda_p = 0.1$).

Let d_n be the Euclidean distance between $\widetilde{X}_n$ and $(n, 0)$, i.e., $d_n = \|\widetilde{X}_n - (n, 0)\|$. By triangle inequality, we have $\|X_v - \widetilde{X}_n\| < \|\widetilde{X}_n - (n, 0)\| + \|X_v - (n, 0)\| < d_n + (1/2) < \infty$. As a result, $T(X_u, X_v) > T_{0,n} - T(X_v, \widetilde{X}_n)$, and thus

$$T(X_u, X_v) = T_{0,n} - \Delta_t, \tag{17}$$

where $\Delta_t < T(X_v, \widetilde{X}_n) < \infty$, which implies that

$$\lim_{\|X_u - X_v\| \to \infty} \frac{T(X_u, X_v)}{\|X_u - X_v\|} = \lim_{n \to \infty} \frac{T_{0,n}}{n}.$$

Thus, from Lemma 5, we obtain

$$\Pr\left(\lim_{\|X_u - X_v\| \to \infty} \frac{T(X_u, X_v)}{\|X_u - X_v\|} = \rho\right) = 1. \tag{18}$$

Now, we derive the upper and lower bounds of ρ. From Proposition 2, ρ is upper bounded by

$$\rho = \inf_{n \geq 1} \frac{E(T_{0,n})}{n} \leq E(T_{0,1}) < \infty. \tag{19}$$

To prove the upper bound $\rho > 0$, let $d_0 = \|\widetilde{X}_0 - (n, 0)\|$ and $d_n = \|\widetilde{X}_n - (n, 0)\|$. It is clear that $d_0 < \infty$ and $d_n < \infty$. By triangle inequality, we have $\|\widetilde{X}_n - \widetilde{X}_0\| > n - d_0 - d_n$. Thus, the number of hops from $\widetilde{X}_n$ to $\widetilde{X}_0$ is at least $(n - d_0 - d_n)/r$. Therefore, ρ is upper bounded by

$$\rho = \inf_{n \geq 1} \frac{E(T_{0,n})}{n} \geq \lim_{n \to \infty} \frac{E(T_i)(n - d_0 - d_n)}{rn} = \frac{E(T_i)}{r} > 0.$$

□

VI. Simulation Results

In this section, we evaluate the impact of node mobility on delay-bounded connectivity through simulations. Consider a secondary network coexisting with a primary network, where PU density $\lambda_p = 0.1$ and secondary user density $\lambda_s = 1.6$. All PUs have a transmission range $R = 1.2$, while secondary users have a transmission range $r = 1$.

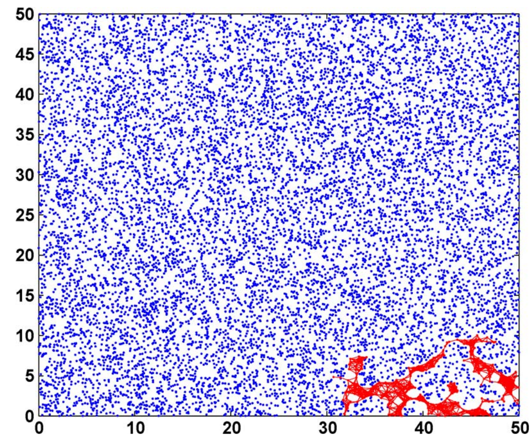

Fig. 8. Largest connected component (in red dots) of a secondary network ($\lambda_s = 5$) coexisting with a primary network ($\lambda_p = 0.1$).

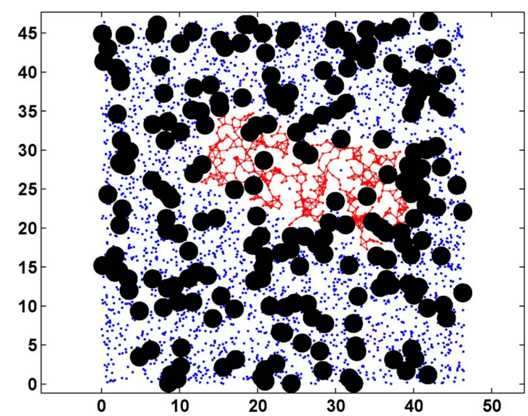

Fig. 9. Largest connected component of static secondary network under $\lambda_p = 0.1$.

As indicated by Theorem 1, because of the HT interference from the primary network, there exists a critical density λ_p^* of the primary network, above which delay-bounded connectivity is not achievable. According to the network settings, we have $\lambda_p^* = 0.3025$. In this case, as shown in Fig. 7, the largest connected component, represented by red dots, only contains a very small portion of the secondary users, which implies that the whole network is partitioned by the HT interference into small components which are disconnected from each other. Moreover, as implied by Theorem 1, the critical density λ_p^* of the primary network does not depend on the density of the secondary network, which implies that increasing the density of secondary users cannot improve the connectivity of the secondary network. This can be observed in Fig. 8, where the secondary network is still disconnected even if the secondary network in Fig. 7 has much higher density $\lambda_s = 5$ than the node density $\lambda_s = 1.6$ of the secondary network in Fig. 7.

Next, we investigate the impact of node mobility on the secondary network connectivity. To combat the impact of HT interference of PUs, the spatial diversity of the spectrum can be exploited by allowing mobile secondary users to exchange messages when they opportunistically move into the same white space, the region without PU interference, and are close enough for data communications. As indicated in Theorem 2, there exists a critical mobility radius a^* above which the secondary

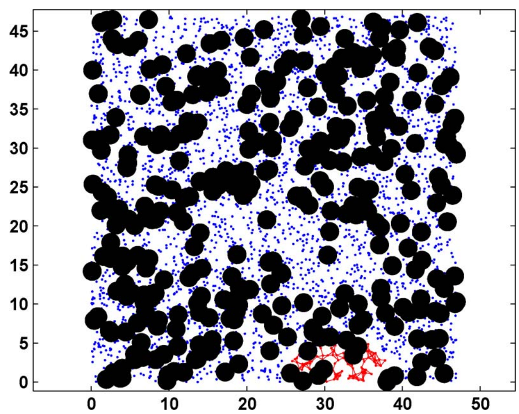

Fig. 10. Largest connected component of static secondary network under $\lambda_p = 0.15$.

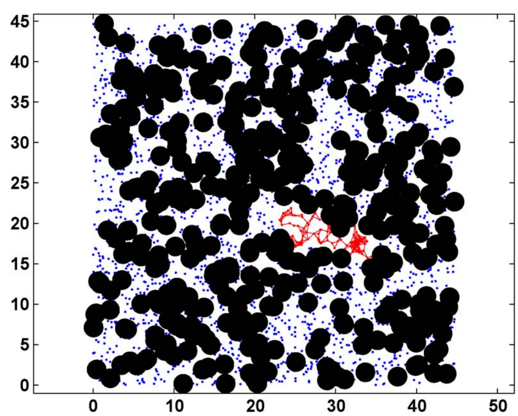

Fig. 11. Largest connected component of static secondary network under $\lambda_p = 0.2$.

network can achieve delay-bounded connectivity surely. Based on the network settings at the beginning of this section, we have $a^* = 12.8$. To validate the existence of a^*, we let all secondary users have the maximum mobility radius equal to $a^* = 12.8$ and evaluate the network connectivity in Fig. 12 accordingly. It is shown that, with the help of node mobility, the secondary network can achieve delay-bounded connectivity so that almost every secondary user resides in a giant connected component. In the contrary, as shown in Fig. 9, without node mobility, the same secondary network is disconnected in the sense that the largest connected component only contains a small portion of secondary users because of the HT interference region of PUs, which is denoted by black disks.

Moreover, we validate the existence of a^* by varying the density λ_p of the primary network, while keeping other network settings unchanged. Accordingly, we have $a^* = 14.8$ for $\lambda_p = 0.15$ and $a^* = 19.1$ for $\lambda_p = 0.2$. It is shown in Figs. 13 and 14 that, by letting all secondary users possess the critical mobility radius, the secondary network can achieve delay-bounded connectivity. On the contrary, as shown in Figs. 10 and 11, the static secondary network without exploiting the node mobility has much worse connectivity performance.

Aside from exploiting mobile opportunistic communications, extending the transmission range of the secondary users is another natural way to increase network connectivity. Next, we study the connectivity performance of static long-range

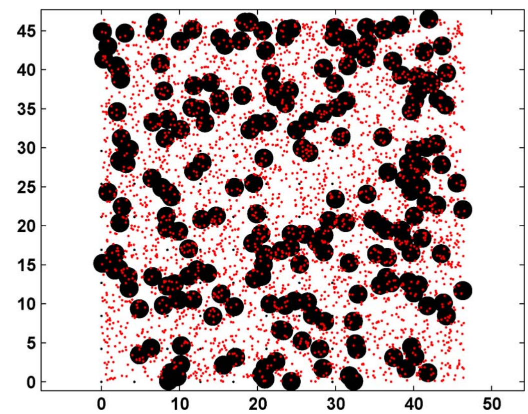

Fig. 12. Largest connected component of mobile secondary network under $\lambda_p = 0.1$.

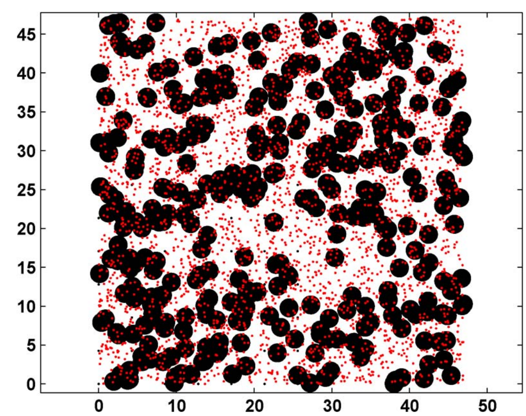

Fig. 13. Largest connected component of mobile secondary network under $\lambda_p = 0.15$.

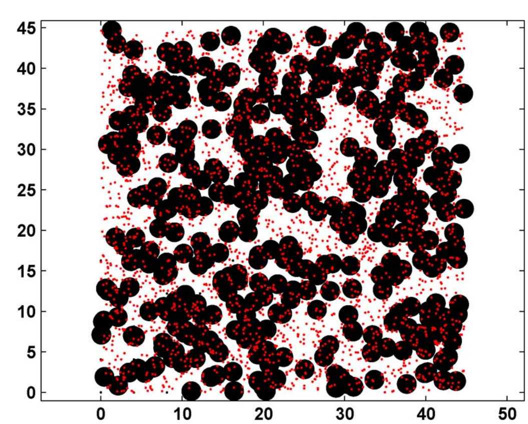

Fig. 14. Largest connected component of mobile secondary network under $\lambda_p = 0.2$.

communications by letting secondary users stay stationary while extending their transmission range in such a way that the communication coverage of the static long-range communications is equal to that of mobility-assisted communications. More specifically, for mobile opportunistic communications, we adopt the same network settings as that in Fig. 14, and thus, we have a^*=19.1. For static long-range communications, we consider the same network settings as that in Fig. 11, except that we increase the transmission range r of the secondary users

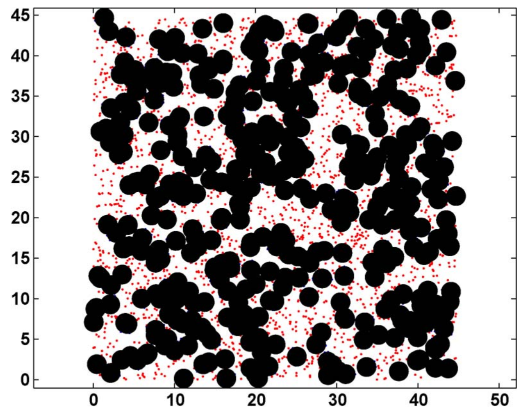

Fig. 15. Largest connected component of the static secondary network with static long-range communication.

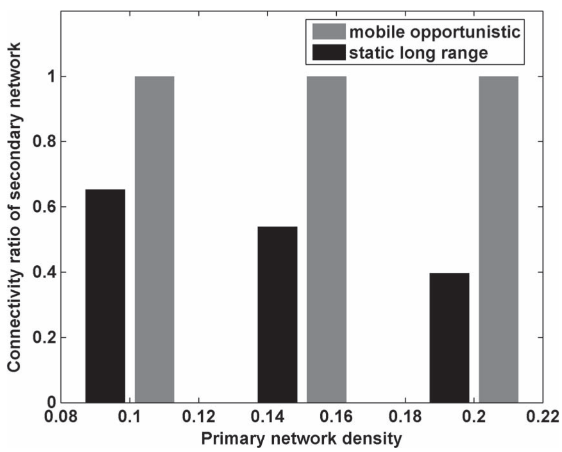

Fig. 16. Largest connected component of the static secondary network with static long-range communication.

from r $= 1$ to $r_{ex} = a^* + r = 20.1$. It is easy to check that such increased transmission range leads to the same communication coverage as the mobile opportunistic communication with the critical mobility radius $a^* = 19.1$. As shown in Fig. 15, even though we significantly increase the transmission range, only a small portion of secondary users are connected, i.e., the largest connected component only consists of a small number of secondary users. If we define the connectivity ratio as the percentage of users residing within the largest connected components, it is shown in Fig. 16 that mobile opportunistic communication can considerably increase the connectivity ratio compared with static long-range communication.

We now evaluate the end-to-end average latency in the secondary network with mobile users. More specifically, we assume that all secondary users have their maximum mobility radius larger than the critical one a^* so that delay-bounded connectivity is guaranteed with the rise of a giant connected component in the network. Then, we randomly select a secondary user from this giant component and evaluate the paths with the smallest average latency between this secondary user and every other secondary user that it can connect to. As indicated by Theorem 3, the latency along those paths, formally defined as the first-passage latency, scales linearly as the initial distance between the sending and receiving parties become large. Such

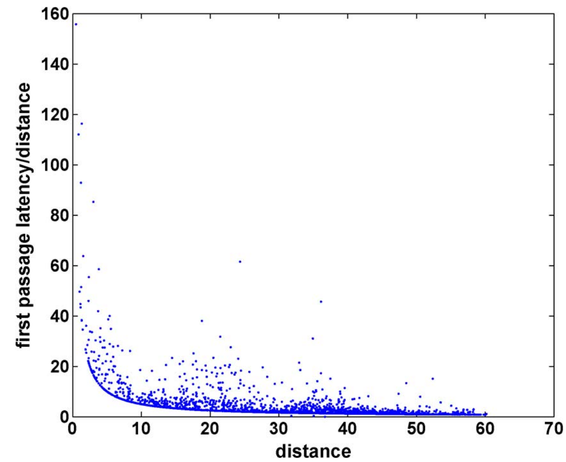

Fig. 17. First-passage latency of a secondary network with mobile users.

asymptotic linear relationship can be seen in Fig. 17, where the ratio between the first-passage latency and the initial node distance approaches a constant as the distance increases.

VII. Conclusion

In this paper, we investigate the fundamental impact of HT interference on the delay-bounded connectivity in wireless networks. Specifically, it is shown that, because of the HT interference induced by the multimedia and Internet services in the primary network, there always exists a critical density λ_p such that, if the density of PUs is larger than λ_p, the delay-bounded connectivity is not achievable since the transmission latency that the secondary users experience is of unbounded mean and variance. To solve this problem, the mobility of secondary users along with the spatial diversity of the wireless spectrum is exploited to avoid the impact of the HT interference. In particular, it is shown that there exists a critical threshold on the maximum radius that the secondary user can reach, above which delay-bounded connectivity can be achieved surely. In this case, the first-passage latency is shown to scale linearly in the Euclidean distance between the transmitter and receiver.

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," *Comput. Netw. J.*, vol. 50, no. 13, pp. 2127–2159, Sep. 2006.

[2] T. Benson, A. Akella, and D. A. Maltz, "Network traffic characteristics of data centers in the wild," in *Proc. 10th ACM SIGCOMM Conf. Internet Meas.*, Melbourne, Australia, Nov. 2008, pp. 267–280.

[3] O. Dousse, F. Baccelli, and P. Thiran, "Latency of wireless sensor networks with uncoordinated power saving mechanisms," in *Proc. ACM MOBIHOC*, 2004, pp. 109–120.

[4] S. Geirhofer and L. Tong, "Dynamic spectrum access in the time domain: Modeling and exploiting white space," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 66–72, May 2007.

[5] G. Grimmett, *Percolation,* 2nd ed. New York, NY, USA: Springer-Verlag, 1999.

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

[7] Z. Kong and E. M. Yeh, "On the latency for information dissemination in mobile wireless networks," in *Proc. MobiHoc*, May 2008, pp. 139–148.

[8] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self similar nature of Ethernet traffic (extended version)," *IEEE/ACM Trans. Netw.*, vol. 2, no. 1, pp. 1–15, Feb. 1994.

[9] T. Liggett, "An improved subadditive ergodic theorem," *Annals of Prob.*, vol. 13, no. 4, pp. 1279–1285, Nov. 1985.

[10] S. Luo, J. Li, K. Park, and R. Levy, "Exploiting heavy-tailed statistics for predictable QoS routing in *ad hoc* wireless networks," in *Proc. IEEE INFOCOM*, Mar. 2008, pp. 2387–2395.
[11] R. Meester and R. Roy", *Continuum Percolation,* 1st ed. New York, NY, USA: Cambridge Univ. Press, 1996.
[12] M. Neely and E. Modiano, "Capacity and delay tradeoffs for *ad hoc* mobile networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1917–1937, Jun. 2005.
[13] K. Park and W. Willinger, *Self-Similar Network Traffic and Performance Evaluation*. Hoboken, NJ, USA: Wiley, 2000.
[14] J. Quintanilla, S. Torquato, and R. M. Ziff, "Efficient measurement of the percolation threshold for fully penetrable discs," *Phys. A*, vol. 33, no. 42, pp. 399–407, Oct. 2000.
[15] P. Wang and I. F. Akyildiz, "Network stability of dynamic spectrum access networks in the presence of heavy tailed traffic," in *Proc. IEEE SECON*, Seoul, Korea, 2012, pp. 165–173.
[16] P. Wang and I. F. Akyildiz, "On the origins of heavy tailed delay in dynamic spectrum access networks," *IEEE Trans. Mobile Comput.*, vol. 11, no. 2, pp. 204–217, Feb. 2012.
[17] P. Wang and I. F. Akyildiz, "Asymptotic queueing analysis for dynamic spectrum access networks in the presence of heavy tails," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 514–522, Mar. 2013.
[18] M. Wellens, J. Riihijarvi, and P. Mahonen, "Empirical time and frequency domain models of spectrum use," *Phys. Commun.*, vol. 2, no. 1/2, pp. 10–32, Mar. 2009.
[19] D. Willkomm, S. Machiraju, J. Bolot, and A. Wolisz, "Primary user behavior in cellular networks and implications for dynamic spectrum access," *IEEE Commun. Mag.*, vol. 47, no. 3, pp. 88–95, Mar. 2009.
[20] F. Xing and W. Wang, "On the critical phase transition time of wireless multi-hop networks with random failures," in *Proc. MOBICOM*, San Francisco, CA, USA, 2008, pp. 175–186.

Pu Wang (M'10) received the B.S. degree in electrical engineering from the Beijing Institute of Technology, Beijing, China, in 2003, the M.Eng. degree in computer engineering from the Memorial University of Newfoundland, St. Johns, NL, Canada, in 2008, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2013.

He is currently an Assistant Professor with the Department of Electrical Engineering and Computer Science, Wichita State University, Wichita, KS, USA. He was named Broadband Wireless Networking (BWN) Lab Researcher of the Year 2012 by Georgia Institute of Technology. He was also named Fellow of the School of Graduate Studies by the Memorial University of Newfoundland in 2008. His research interests include wireless sensor networks, cognitive radio networks, software defined networks, nanonetworks, multimedia communications, wireless communications in challenged environment, and cyber-physical systems

Dr. Wang received the technical program committee (TPC) top ranked paper award of IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN) 2011.

Ian F. Akyildiz (M'86–SM'89–F'96) received the B.S., M.S., and Ph.D. degrees in computer engineering from the University of Erlangen-Nürnberg, Erlangen, Germany, in 1978, 1981, and 1984, respectively.

Currently, he is the Ken Byers Chair Professor in Telecommunications with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA, the Director of the Broadband Wireless Networking Laboratory, and the Chair of the Telecommunication Group at Georgia Tech. He is an honorary professor with the School of Electrical Engineering at Universitat Politècnica de Catalunya (UPC) in Barcelona, Catalunya, Spain, and the founder of the Nanonetworking Center in Catalunya. Since 2011, he has been a Consulting Chair Professor at the Department of Information Technology, King Abdulaziz University in Jeddah, Saudi Arabia. Since January 2013, he has also been a Finland Distinguished Professor Program (FiDiPro) Professor (FiDiPro supported by the Academy of Finland) at the Tampere University of Technology, Department of Communications Engineering, Finland. His current research interests are in nanonetworks, long-term-evolution advanced networks, cognitive radio networks, and wireless sensor networks.

Dr. Akyildiz is the Editor-in-Chief of Computer Networks (Elsevier) Journal and the founding Editor-in-Chief of the Ad Hoc Networks (Elsevier) Journal, the Physical Communication (Elsevier) Journal, and the Nano Communication Networks (Elsevier) Journal. He is an Association for Computing Machinery (ACM) fellow (1997). He received numerous awards from IEEE and ACM.