The Impact of Network Size and Mobility on Information Delivery in Cognitive Radio Networks

Lei Sun, *Student Member, IEEE,* Wenye Wang, *Senior Member, IEEE,* Yujin Li, *Student Member, IEEE*

***Abstract*—There have been extensive works on the design of opportunistic spectrum access and routing schemes to improve spectrum efficiency in Cognitive Radio Networks (CRNs), which becomes an integral component in the future communication regime. Nonetheless, the potentials of CRNs in boosting network performance yet remain to be explored to reach the full benefits of such a phenomenal technique. In this paper, we study the end-to-end latency in CRNs in order to find the sufficient and necessary conditions for real-time applications in finite networks and large-scale deployments. We first provide a general mobility framework which captures most characteristics of the existing mobility models and takes *spatial heterogeneity* into account. Under this general mobility framework, secondary users are mobile with an *mobility radius* α, which indicates how far a mobile node can reach in spatial domain. We find that there exists a cutoff point on α, below which the latency has a heavy tail and above which the tail of the latency is bounded by some *Gamma* distributions. As the network grows large, the latency is asymptotically scalable (linear) with respect to the dissemination *distance* (e.g., the number of hops or Euclidean distance). An interesting observation is that although the density of primary users adversely impacts the expected latency, it makes no influence on the *dichotomy* of the latency tail in finite networks and the linearity of latency in large networks. Our results encourage CRN deployment for real-time and large applications, when the mobility radius of secondary users is large enough.**

***Index Terms*—Latency, Cognitive Radio Networks, Scalability, Generic Mobility**

I. Introduction

An increasing number of users, homes, and enterprises rely on wireless technologies (such as cellular and wifi networks) for their daily activities. However, the growth of wireless networks has been hindered by the scarce and inefficient usage of radio spectrum. Currently, spectrum access is regulated by government agencies, for example, by the Federal Communications Commission (FCC) of the United States, which allocates spectrum by assigning exclusive licenses to users to operate their networks on a long term basis in large geographical regions. A recent report from FCC reveals that under this static allocation, merely $5\% \sim 15\%$ of the spectrum is utilized on average [1]. This leads new wireless applications starving for spectrum, while large chunks of it remaining idle most of the time under their current owners.

To overcome the artificial shortage of spectrum due to this misallocation, CRN has been proposed to improve spectrum efficiency for opportunistic communication and become an integral component in the future communication regime [1], [2]. In cognitive radio networks, there are two types of users: (i) *primary users* who have license from the regulator and thus have priority to utilize spectrum, and (ii) *secondary users* who opportunistically access spectrum without interfering with the coexisting primary users. Using CRNs, secondary users can gain access to the spectrum and primary users can gain financial incentives to lease their idle spectrum.

Many efforts have been made recently to understand the characteristics of CRNs such as to enable their deployment for realistic applications, including capacity limits, spectrum sensing, spectrum mobility, and spectrum sharing [3]–[8]. These works have presented a very good understanding of the potential of cognitive communications in optimizing spectrum utilization. However, the key question to the deployment of CRNs is *not* whether the spectrum efficiency is improved, but whether CRNs are able to support *applications*. For example, spectrum can be overly used, with a very high throughput, but the latency may become extremely long, falling into the traditional problem of the tradeoff between network throughput and latency [9]–[12]. To this end, we aim to study a fundamental problem, i.e., what the stochastic properties of end-to-end latency in cognitive radio networks are. The understanding of end-to-end delay offers a straightforward interpretation of the potentials of CRNs for time-sensitive applications. For example, when a CRN is used for emergency rescue in the aftermath of disasters or traffic accidents (e.g., vehicular networks), we need to ensure that *help* or *warning* messages can be disseminated to a chosen destination in time, which becomes more important than other performance metrics, such as the total network capacity in these circumstances.

Despite its importance, the latency is an under-explored problem and not well understood in multihop wireless networks. The pioneering work in [13], [14] studied the packet latency for the *fully connected* wireless ad-hoc networks and showed that there exist bounds on the latency which are tight when the number of nodes is large enough. Instead of *full connectivity*, some studies [15], [16] further showed that the latency scales asymptotically at least linearly with the transmission distance in wireless sensor networks when these networks are *percolated*. These results have greatly advanced our understanding of the nature of latency, and also laid a good foundation to approach the problem. Unfortunately, these

Lei Sun is with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh NC. E-mail: lsun3@ncsu.edu.
Wenye Wang is with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh NC. E-mail: wwang@ncsu.edu.
Yujin Li is with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh NC. E-mail: yli27@ncsu.edu.
This work is supported by the NSF Award CNS 0546289, NSF Career Award and Defense Threat Reduction Agency (DTRA) Award HDTRA1-08-1-0024. Part of the work was presented in the 31th IEEE International Conference on Computer Communications (INFOCOM' 12).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2015.2398420, IEEE Transactions on Mobile Computing

results may not be applicable to CRNs because (i) asymptotic results were obtained by assuming that wireless nodes are static; and (ii) these results [13]–[16] are only derived for *large* networks when the number of nodes approaches to infinity; (iii) these results are derived for *homogeneous networks* in which every node has the same capacity in propagation.

Particularly, mobility plays a critical role on the latency, which has been evidenced by earlier results. For instance, the seminal work [17] showed that mobility can improve the capacity in large wireless ad hoc networks at the cost of the delay. This result is obtained by assuming that nodes move according to an ergodic process that is equally likely to visit any portion of the network area. That is, the nodes are *spatially homogeneous.* With the similar assumptions, capacity-delay tradeoffs have been extensively studied under various mobility models, such as the i.i.d model [9], the Brownian motion [10], the reshuffling model [11] and variants of random walk and random way-point models [18], [19]. Later on, spatial inhomogeneity has been taken into account in [20], [21] where the nodes are either restricted to move within an randomly chosen cell or the coverage of a home point. These studies motivate an interesting question about the latency under general mobility.

It is evident that the asymptotic results, though, providing good insights into network performance, may not explain the latency properties when the number of nodes in real applications is *finite*. In other words, the stochastic properties of latency distribution in finite networks lead to understanding of real networks, rather than the large networks. As the last point, CRNs feature *heterogeneity* in wireless nodes, since there are two types of nodes, primary and secondary nodes [22], which is left open for study on the impact of latency distribution.

Putting all together, in this paper we study the latency distribution in *finite* networks, and the *scaling law* for large networks with *infinite* number of secondary nodes under general mobility where spatial inhomogeneity is considered in addition to common features of a variety of mobility models. We find that in finite CRNs, the latency of information dissemination depends on the *mobility radius* α, which indicates how far a mobile node can reach in spatial domain. Moreover, there exists a *cutoff point* on α, below which the latency has a *heavy tail*; and above which the tail of its distribution is bounded by some Gamma distribution. In addition, as the network grows large, the latency asymptotically scales linearly with respect to the *distance* in terms of the number of hops or the Euclidean distance between the source and destination nodes if the network remains fully connected or percolated. It is interesting, though not surprising, that the density of primary nodes presents an adverse influence on the expected latency, but showing no obvious effect on the dichotomy of the latency tail in finite networks and linear scaling law of the asymptotic latency with respect to dissemination distance.

The rest of this paper is organized as follows. We describe the mobility and network models, and formulate the latency problem in Section II. Our main results are summarized in Section III. In Sections IV and V, we present the proofs and simulations results of our findings on dissemination latency in detail. Section VI concludes the paper with two promising applications of CRNs.

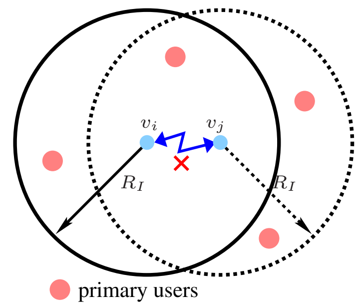


Fig. 1. *Primary-secondary* interference.

II. System Models and Problem Formulation

In this section, we first describe the network models and then collect basic assumptions, notations and definitions of the metric of interest that will be used throughout the paper.

A. Network Models

We consider a CRN consisting of n mobile secondary users $\{v_1, \ldots, v_n\}$ in a torus surface $\Omega_n = [0, \sqrt{\frac{n}{\lambda}}]^2$ (λ is the spatial density of secondary users). Denote $V(t) = (v_1(t), \ldots, v_n(t))$ as the positions of secondary users at time t. A set of m channels $\{ch_1, \ldots, ch_m\}$ are assumed to be accessible by secondary users. For any $1 \leq k \leq m$, an overlay network of primary users with spatial density λ_{pk} are transmitting with channel ch_k. We assume that $\lambda_{pk} = \lambda_p$ for any k for simplicity. To model the dynamics of the primary traffic, we adopt a synchronized slotted structure, which has been used in [23] to study the connectivity of a large single-channel CRN. Particularly, time is slotted into units and at any time slot, primary users transmitting on any channel ch_k are assumed to be uniformly and independently distributed in Ω_n, and such distribution is i.i.d across slots.

1) Interference Models: In CRNs, there are two types of interference for information dissemination among secondary users: *secondary-secondary* and *primary-secondary* interference. The former interference can be characterized by the well-known *protocol model* [24], which has been widely adopted in homogeneous networks [9], [14], [17], [24]. Particularly, without interference with primary users, a successful transmission from a mobile secondary user v_i to v_j is achievable at time t if $\|v_i(t) - v_j(t)\| \leq r$ and for any other simultaneously transmitting node on the same channel v_l, $\|v_l(t) - v_j(t)\| \geq (1+\Delta)r$, where r is the transmission radius of secondary users, and Δ models the guard zone around v_j in which any simultaneously transmission on the same channel causes collision at v_j.

In terms of the latter interference, let us denote R_I as the interference range of primary users. And as Fig. 1 shows, two secondary users v_i and v_j are permitted to use the channel ch_k only when there are no primary users on ch_k in the neighborhood, i.e., $\|v_i(t) - u(t)\| > R_I$ for any primary user u transmitting with ch_k, where $u(t)$ is the position of u at time t.

We assume that $r < R_I$ *W.O.L.G.*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2015.2398420, IEEE Transactions on Mobile Computing

2) Mobility Models: We consider a general mobility model, $\mathcal{M}(\Phi, \Psi, \alpha)$, which is characterized by three parameters Φ, Ψ, and α over the region Ω_n. First, *spatial heterogeneity*, which accounts for the scenario that mobile nodes are more likely to be found in some area (e.g., the neighborhood of their home in the case of people, or the neighborhood of the garage in the case of vehicles), is taken into account. Particularly, we consider that a node spends most of its time in a small region, and rarely visits the areas far away from it. We model this behavior by assuming that each node v_i has a *home point* [21], located at v_i^h. Nodes move "around" their home points according to independent stationary and ergodic processes. We assume that each home point v_i^h is associated with a fixed point v_i^c, which is called the *center point* of v_i. The center points are regularly placed in Ω_n. For example, $\{v_1^c, \ldots, v_n^c\}$ are placed regularly at positions $(\frac{1}{2\sqrt{\lambda}} + \frac{i}{\sqrt{\lambda}}, \frac{1}{2\sqrt{\lambda}} + \frac{j}{\sqrt{\lambda}})$ with $0 \le i \le \sqrt{n} - 1$ and $0 \le j \le \sqrt{n} - 1$ (since we do not have knowledge about the specific positions of center points, this layout represents an average scenario and simplifies our analysis.). We describe the distribution of the home point v_i^h around v_i^c by a non-increasing probability density function $\Phi_i(x) = \Phi(x - v_i^c)$, which is assumed to be invariant in all directions and used as the first parameter in the mobility model. The second parameter, $\Psi_i(x) = \Psi(x - v_i^h)$ is used to describe the probability density of a node v_i around v_i^h, which is again a non-increasing and direction-invariant function. We assume that Ψ_i is non-zero in and only in a region characterized by a constant α; that is, $\Psi_i(x) = \Psi(x - v_i^h) > 0$ when $\|x - v_i^h\| < \alpha$ and $\Psi_i(x) = \Psi(x - v_i^h) = 0$, otherwise. We refer α as ***mobility radius***.

Remark 1: The idea of "home points" is not new [21] and it has been used to describe the *spatial inhomogeneity* incurred by the mobility of a particular wireless node. We introduce an additional concept, "center points" to model the heterogeneously spatial distribution of the home points, which characterizes the *spatial inhomogeneity* incurred by *heterogeneous mobility* of different wireless nodes. This two-level mobility model accounts for a wide range of mobility patterns. For example, if the probability density function $\Phi(x)$ is a constant function independent of x (i.e., home points are uniformly distributed over Ω_n), $\mathcal{M}(\Phi, \Psi, \alpha)$ reduces to the *Uniform Anisotropic* model in [21]. Furthermore, if the probability density function $\Psi_i(x) = \Psi(x - v_i^h) = \delta(x - v_i^h)$, where $\delta(x)$ is the Dirac impulse function, $\mathcal{M}(\Phi, \Psi, \alpha)$ reduces to the static model in [24], where nodes are assumed to be static and uniformly distributed; if $\Psi(x)$ is also a constant function independent of x and α, $\mathcal{M}(\Phi, \Psi, \alpha)$ reduces to the homogeneous mobility model in [17]; and if $\Psi(x)$ is a threshold function whose value is zero when $x \ge \alpha$ and a nonzero constant when $x < \alpha$, $\mathcal{M}(\Phi, \Psi, \alpha)$ reduces to the *constrained i.i.d* model used in [16].

Mobility of Secondary Users: In this paper, we assume that secondary users are mobile under the *general mobility* $\mathcal{M}(\Phi, \Psi, \alpha)$.

To facilitate the study of the dissemination latency of secondary users, we consider three classes according to the *spatial inhomogeneity* of home points:

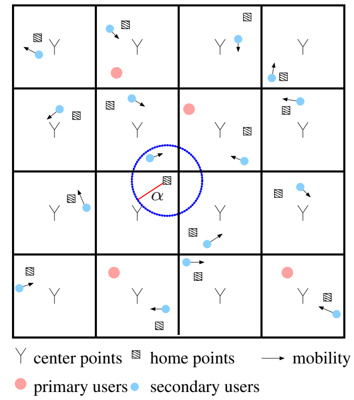


Fig. 2. An illustration of the general mobility $\mathcal{M}(\Phi, \Psi, \alpha)$.

- Extremely Inhomogeneous Home Points (**EIHP**) mobility $\mathcal{M}(\Phi_E, \Psi, \alpha)$: Home points are fixed and regularly placed over Ω_n. Here $\Phi_E(x) = \delta(x)$.
- Partial Inhomogeneous Home Points (**PIHP**) mobility $\mathcal{M}(\Phi_P, \Psi, \alpha)$: As shown in Fig. 2, center points $\{v_i^c\}_{i=1}^n$ partition the Ω_n into n *subregions* $\{\mathcal{O}_i\}_{i=1}^n$ as Voronoi diagrams. In this class, the home point v_i^h is randomly distributed in $\mathcal{O}_i$.
- Homogeneous Home Points (**HHP**) mobility $\mathcal{M}(\Phi_H, \Psi, \alpha)$: Home points $\{v_i^h\}_{i=1}^n$ are independently and uniformly distributed over Ω_n. Here $\Phi_H(x)$ is a constant density function independent of x.

Mobility of Primary Users: Mobility of primary users (spatial dynamics) influences the probability of the existence of links between secondary users (spectrum dynamics, see Section II-A1). As shown in [16], a network of mobile nodes is equivalent to a network of stationary nodes with dynamic links, in terms of spectrum dynamics, and the latter is much easier to be analyzed. Thus instead of mobile primary users, we assume stationary primary users with dynamic links in this paper.

B. Problem Formulation

We denote $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi, \Psi, \alpha), (\lambda, \lambda_p)]$ as a CRN $\mathcal{F}_{m,n}$, where n secondary users opportunistically access m channels and are mobile under $\mathcal{M}(\Phi, \Psi, \alpha)$, and the spatial densities of secondary users and primary users on any channel are λ and λ_p respectively. We further denote $\mathbb{L}(t)$ as the set of communication links among secondary users at time t and $\mathbb{L}(t)$ is obviously dynamic due to the mobility of the secondary and primary users.

In this paper, we focus on the *dissemination latency*, i.e., *how fast* information can be disseminated from the source to the destination secondary user. Therefore, rebroadcasting and "store-carry-and-forward" communication paradigm (also named mobility-assisted routing) are considered. Specifically, by omitting the propagation delay, when the source v_s broadcasts a message at time 0, all the secondary users connected to v_s in $\mathbb{L}(0)$ receive the message instantly. Denote $l_{i,j}$ as a communication link between secondary users v_i and v_j and $\mathcal{V}(t)$ as the set of secondary users that have received the message at time t.

Definition 1: The *first hitting time* between v_i and v_j is defined as $\mathcal{T}_h(v_i, v_j) \triangleq \inf\{t \ge 0 : l_{i,j} \in \mathbb{L}(t)\}$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2015.2398420, IEEE Transactions on Mobile Computing

Definition 2: The *dissemination latency* $\mathcal{T}_d$ from the source v_s and the destination v_d is defined as:

$$\mathcal{T}_d \triangleq \inf\{t \geq 0 : v_d \in \mathcal{V}(t)\}.$$

In $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi, \Psi, \alpha), (\lambda, \lambda_p)]$, three metrics can be used to characterize *how far* two nodes v_i and v_j are apart:

- $d^{(t)}(v_i, v_j)$: the distance between v_i and v_j at time t.
- $d_h(v_i, v_j)$ and $d_c(v_i, v_j)$: the distance between home points and center points of v_i and v_j respectively.

Here the *distance* can be any *p-norm* metric function and we consider two of the most popular metrics *transmission hops* and *Euclidean distance*.

Definition 3: Denote $\mathcal{D}$ as the distance, i.e., either the number of hops or the Euclidean distance between v_s and v_d, which can be one of the three distances explained above. We define $\mathcal{S}_d \triangleq \frac{\mathcal{T}_d}{\mathcal{D}}$, which characterizes how fast information disseminates and is called *dissemination speed* in this paper.

Based on the definitions and system models, we can formulate the problem as

1) In a finite $\mathcal{F}_{m,n}$, what the distribution of the dissemination latency $\mathcal{T}_d$ is;
2) As the network grows large, say to infinity, whether the dissemination latency $\mathcal{T}_d$ is scalable with respect to dissemination distance $\mathcal{D}$, i.e., whether $\mathcal{S}_d$ is going to converge.

III. Main Results

The key question in this study is how fast information is disseminated in both finite and large CRNs under general mobility $\mathcal{M}(\Phi, \Psi, \alpha)$. We first study the dissemination latency $\mathcal{T}_d$ in CRNs where secondary users are mobile under the three subclasses of models EIHP, PIHP and HHP, respectively. Then based on the generalization of these results, we obtain the fundamental properties of the dissemination latency $\mathcal{T}_d$ when secondary users are mobile under the general mobility $\mathcal{M}(\Phi, \Psi, \alpha)$. We summarize our main results as follows.

Theorem 1: In a CRN $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi, \Psi, \alpha), (\lambda, \lambda_p)]$ with finite users, there exists a **cutoff point** on the *mobility range* α, above which the tail distribution of dissemination latency $\mathcal{T}_d$ is bounded by some *Gamma* distribution; below which $\mathcal{T}_d$ has a *heavy-tailed* distribution and $\mathbb{P}(\mathcal{T}_d = \infty) > 0$.

Remark 2: $\mathbb{P}(\mathcal{T}_d = \infty) > 0$ indicates a positive probability that the destination will not receive the message from the source. Thus the requirement $\mathbb{P}(\mathcal{T}_d < \infty) = 1$ in mobile wireless networks is equivalent to the *connectivity* in the wired networks, which is used as a prerequisite to evaluate network functions. Moreover, a *heavy tail* of the dissemination latency $\mathcal{T}_d$ implies a significant probability that it takes long time to disseminate a message from the source to the destination. Thus in addition to a bounded dissemination latency (i.e., $\mathbb{P}(\mathcal{T}_d < \infty) = 1$), a *light-tailed* distributed dissemination latency $\mathcal{T}_d$ (i.e., $E(\mathcal{T}_d) < \infty$) is also crucial for time-critical applications in CRNs. Therefore, a *light-tailed* distribution of $\mathcal{T}_d$ is assumed or required in many deployments and performance studies of wireless networks in the literature. For example, the authors in [17] implicitly assume that the dissemination latency is exponentially bounded (light-tailed) so as to make their delay-capacity tradeoff analysis tractable.

Theorem 1 tells that to achieve a *light-tailed* dissemination latency (note that *Gamma distribution* is a type of light-tailed distribution), the *mobility radius* of secondary users α need to be larger than some cutoff points, which are specifically identified in Proposition 1 for EIHP, Proposition 2 for PIHP and Proposition 3 for HHP, respectively. This result encourages the existing endeavor of deploying CRN for practical applications, including time-critical applications, such as emergency networks and military networks.

We must emphasize that the goal of this paper is to investigate the fundamental properties of the dissemination latency $\mathcal{T}_d$ in CRNs where secondary users are mobile according to the **general** mobility patterns. However, if given more knowledge of the CRN, e.g.,the specific mobility patterns, the same proof can also be used to derive more specific distributions of $\mathcal{T}_d$

As the network size increases, we have the following theorem on the scalability of the dissemination latency $\mathcal{T}_d$.

Theorem 2: Given a large connected[1] cognitive radio network $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi, \Psi, \alpha), (\lambda, \lambda_p)]$, there exists a finite constant κ such that $\mathbb{P}(\lim_{\mathcal{D}\to\infty} \mathcal{S}_d = \lim_{\mathcal{D}\to\infty} \frac{\mathcal{T}_d}{\mathcal{D}} = \kappa) = 1$.

Remark 3: Scalability has been one of the most fundamental problems that has discouraged the deployment of large wireless networks [9], [24]. Theorem 2 reveals that in large connected CRNs, the dissemination latency $\mathcal{T}_d$ asymptotically scales linearly with the initial *distance* between the source and destination, i.e., the message sent by a source reaches its destination at a fixed asymptotic speed. This result enables the feasible deployment of CRNs for large applications.

It is worthy of noting that we aim to understand the fundamental properties of the dissemination latency $\mathcal{T}_d$ in CRNs under general mobility. However, besides the theoretical importance of our findings, our results can be used practically not only in the initial deployment of a CRN, but also in evaluating the performance of network applications. For example, in a large deployment of CRNs as sensor-actuator networks, the result in Theorem 2 can be used to estimate the delay elapsed between the time at which an incoming event is sensed and the time that this event report is retrieved by the data collecting sink. In the next two sections, we present the proofs for Theorem 1 and Theorem 2, which studies the distribution and scalability of the dissemination latency $\mathcal{T}_d$ in finite and large CRNs under EIHP, PIHP, and HHP mobility, respectively.

IV. The Distribution of $\mathcal{T}_d$ in Finite Networks

Here, we present the proof for Theorem 1 described in Section III. This result tells that there exists a cutoff point on the *mobility capability* α in a finite cognitive radio network under general mobility $\mathcal{M}(\Phi, \Psi, \alpha)$, above which the tail of the dissemination latency $\mathcal{T}_d$ is bounded by some *Gamma* distribution and below which $\mathcal{T}_d$ has a heavy tail. For the convenience of analysis, we first study the distribution of $\mathcal{T}_d$ under these three subclasses of mobility models, i.e., EIHP $\mathcal{M}(\Phi_E, \Psi, \alpha)$, PIHP $\mathcal{M}(\Phi_P, \Psi, \alpha)$, and HHP mobility

[1] We consider two types of connectivity in large CRNs: *full connectivity* and *percolation-based connectivity*. The former is that there exists a communication path between any two nodes; and the latter is that there exists a large component well scattered over the entire network.

$\mathcal{M}(\Phi_H, \Psi, \alpha)$ based on the *spatial inhomogeneity* of home points defined in Section II, respectively. Then we move on to identify the fundamental properties of $\mathcal{T}_d$ under the general mobility. The scalability of the dissemination latency $\mathcal{T}_d$ in large CRNs (Theorem 2) will be considered in the next section. To proceed, we find the following definitions useful toward the derivation of tail distribution of the dissemination latency $\mathcal{T}_d$.

Definition 4: If Z and Z' are random variables such that $\mathbb{P}(Z > z) \leq \mathbb{P}(Z' > z)$ for all z, we say that Z is *stochastically dominated* by Z' and write $Z \overset{\mathcal{D}}{<} Z'$; and if $Z \overset{\mathcal{D}}{<} Z'$, there exists a random variable $\hat{Z}'$, which has the same distribution of Z' such that $Z \leq \hat{Z}'$ ($\hat{Z}'$ is called a *coupling* of Z' [25].)

Definition 5: If Z and Z' are random variables such that $\mathbb{P}(Z > z) \leq \mathbb{P}(Z' > z)$ for large z, we say that the Z's **tail** is *stochastically dominated* by Z''s **tail**.

Remark 4: Coupling is a very important tool in probability theory which is used throughout the paper. To use the coupling method, *stochastic domination* is required (as shown in Definition 4). However, in finite CRNs, we are interested in the tail distribution of the dissemination latency $\mathcal{T}_d$, which implies that only *stochastic tail domination* needs to be considered. Therefore, in order to use *coupling*, we need the following lemma, which bridges the gap between *stochastic domination* and *stochastic tail domination*.

Lemma 1: Given non-negative i.i.d random variables $\{X_i\}_{i=1}^{\infty}$ and $\{Y_i\}_{i=1}^{\infty}$ where $\mathbb{P}(X_i > t) \leq \mathbb{P}(Y_i > t)$ for large t, i.e., the *tails* of the former are *stochastically dominated* by the latter, there exist i.i.d random variables $\{\bar{X}_i\}_{i=1}^{\infty}$, which have the same tail distribution with $\{X_i\}_{i=1}^{\infty}$ and are *stochastically dominated* by $\{Y_i\}_{i=1}^{\infty}$. Furthermore, for any finite k, $\sum_{i=1}^{k} X_i$ has the same tail with $\sum_{i=1}^{k} \bar{X}_i$.

Proof: (Sketch.) Assume $\mathbb{P}(X_i > t) \leq \mathbb{P}(Y_i > t)$ when $t > t_c$ for some finite constant t_c. We construct $\{\bar{X}_i\}_{i=1}^{\infty}$ as $\bar{X}_i = 0$ when $X_i \leq t_c$ and $\bar{X}_i = X_i$ otherwise. This proves the first part. For the second part, we only need to show $\mathbb{P}(\bar{X}_1 + \bar{X}_2 > t) = \mathbb{P}(X_1 + X_2 > t)$ for large t:

$$\begin{aligned}\mathbb{P}(\bar{X}_1 + \bar{X}_2 > t) = \mathbb{P}(\bar{X}_1 < t_c)\mathbb{P}(\bar{X}_1 + \bar{X}_2 > t | \bar{X}_1 < t_c) + \\ \mathbb{P}(\bar{X}_1 > t - t_c)\mathbb{P}(\bar{X}_1 + \bar{X}_2 > t | \bar{X}_1 > t - t_c) \\ + \mathbb{P}(t_c < \bar{X}_1 < t - t_c)\mathbb{P}(\bar{X}_1 + \bar{X}_2 > t | t_c < \bar{X}_1 < t - t_c).\end{aligned}$$

Note that the third item on the right hand side is equal to its counterpart of $\mathbb{P}(X_1 + X_2 > t)$ and the first two items are on the higher order of the third item as $t \to \infty$. Thus $\mathbb{P}(\bar{X}_1 + \bar{X}_2 > t) \to \mathbb{P}(X_1 + X_2 > t)$ for large t and this completes the proof. ∎

A. Distribution of $\mathcal{T}_d$ under EIHP Mobility $\mathcal{M}(\Phi_E, \Psi, \alpha)$

Prior studies [15], [16] have shown that *propagation delay* in networks whose topologies change frequently (e.g., due to mobility) is negligible, in comparison with the latency incurred by the topology dynamics. Therefore, $\mathcal{T}_d$ can be coupled as the sum of a sequence of the first hitting time $\mathcal{T}_h$ between secondary users along a communication path from the source to the destination node. Hence we study $\mathcal{T}_h$ first. In EIHP mobility, secondary users move around home points, which are overlaid with center points, with the Euclidean distance between neighboring home points being $\sqrt{\frac{1}{\lambda}}$ (see Fig. 2). The following lemma studies the property of the first hitting time $\mathcal{T}_h(v_i, v_j)$ between v_i and v_j with neighboring home points.

Lemma 2: Given secondary users v_i and v_j in a finite CRN $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_E, \Psi, \alpha), (\lambda, \lambda_p)]$ with $d_c(v_i, v_j) = \sqrt{\frac{1}{\lambda}}$, we have i) $\mathbb{P}(T_h(v_i, v_j) = \infty) = 1$ if $\alpha < \frac{\sqrt{\frac{1}{\lambda}} - r}{2}$; ii) otherwise, $E(T_h(v_i, v_j)) < \infty$ and $\mathbb{P}(T_h(v_i, v_j) > t) \leq e^{-c_1 t}$ for sufficiently large t and some positive constant c_1.

Proof: At time t if and only if $d^{(t)}(v_i, v_j) < r$, v_i and v_j may communicate directly. Also, $d^{(t)}(v_i, v_j) > d_h(v_i, v_j) - 2\alpha$ for all t. Thus if $\alpha < \frac{\sqrt{\frac{1}{\lambda}} - r}{2}$, $d^{(t)}(v_i, v_j) > r$ for all t, which implies that v_i and v_j cannot communicate with each other. This completes the proof of part i).

Now, suppose $\alpha > \frac{\sqrt{\frac{1}{\lambda}} - r}{2}$. Denote $\mathcal{E}_t$ as the event that there exists no communication link between v_i and v_j at time t and $\bar{\mathcal{E}}_t$ as its complement. Let $\varepsilon = \mathbb{P}(\mathcal{E}_t)$ be the probability of $\mathcal{E}_t$. As shown in Fig. 3, a necessary condition for $\bar{\mathcal{E}}_t$ is that there exist no primary users within the bigger circle centered at o, and a sufficient condition for $\bar{\mathcal{E}}_t$ is that v_i lies in the shaded region S_1, v_j in S_2 and no primary users in the bigger circle. Therefore,

$$\begin{aligned}0 < 1 - (1 - \pi R_I^2/(n/\lambda))^m < \varepsilon = \mathbb{P}(\mathcal{E}_t) = 1 - \mathbb{P}(\bar{\mathcal{E}}_t) \\ < 1 - (1 - \pi R_I^2/(n/\lambda))^m \bar{\Psi}(S_1)\bar{\Psi}(S_2) < 1, \quad (1)\end{aligned}$$

where $(1 - \pi R_I^2/(n/\lambda))^m$ characterizes the probability of the necessary condition and $(1 - \pi R_I^2/(n/\lambda))^m \bar{\Psi}(S_1)\bar{\Psi}(S_2)$ characterizes the probability of the sufficient condition for $\bar{\mathcal{E}}_t$, respectively. $\bar{\Psi}(S) = \int_S \Psi dS$ and n/λ is the area of Ω_n. To proceed, we next find an index set $\mathcal{I}$ such that $\{\mathcal{E}_t\}_{t\in\mathcal{I}}$ are independent and let $\varepsilon = \mathbb{P}(\mathcal{E}_t)$ for convenience.

Denote ρ as a *renewal* interval for secondary users, i.e., for any $t > 0$, $\{v_i(t') : t' \leq t\}$ and $\{v_i(t'' + \rho) : t'' \geq t\}$ are independent. Denote $\{\rho_i\}_{i=1}^{\infty}$ as a sequence of i.i.d random variables with the same distribution as ρ. Now we consider the index set $\mathcal{I}_t = \{0, t_1, \ldots, t_{N(t)}\} \subset (0, t]$, where $t_k = \sum_{i=1}^{k} \rho_i$ and $N(t) = |\mathcal{I}_t| = \max\{k : t_k \leq t\}$. Observe that

$$\mathbb{P}(T_h(v_i, v_j) > t) \leq \mathbb{P}(\cap_{s\in\mathcal{I}_t} \mathcal{E}_s) = \prod_{s\in\mathcal{I}_t} \mathbb{P}(\mathcal{E}_s) \qquad (2)$$

where the last equality is by the independency of $\{\mathcal{E}_s\}_{s\in\mathcal{I}_t}$ and by conditioning on $N(t)$, we have

$$\mathbb{P}(T_h(v_i, v_j) > t) \leq E(\varepsilon^{N(t)}) = E(e^{-\beta N(t)}), \qquad (3)$$

where $\beta = -\log\varepsilon > 0$. In addition, for any $\tau > 0$,

$$\begin{aligned}E(e^{-\beta N(t)}) = E(e^{-\beta N(t)} I_{\{N(t)\leq\tau t\}}) + \\ E(e^{-\beta N(t)} I_{\{N(t)\leq\tau t\}}) \leq \mathbb{P}(N(t) \leq \tau t) + e^{-\beta\tau N(t)}.\end{aligned}$$

Note that the finite sum of exponentially bounded random variables is still exponentially bounded [25], [26]. Thus, if we can show that $\mathbb{P}(N(t) \leq \tau t)$ is exponentially bounded, we will finish the proof. In order to proceed, we assume that the tails of renewals $\{\rho_i\}_{i=1}^{\infty}$ are exponentially bounded. This assumption is reasonable considering the network is finite, which has been well-explained in many mobility models [16]–[18], [21], [26]. We next show that $\mathbb{P}(N(t) \leq \tau t)$ is exponentially bounded.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2015.2398420, IEEE Transactions on Mobile Computing

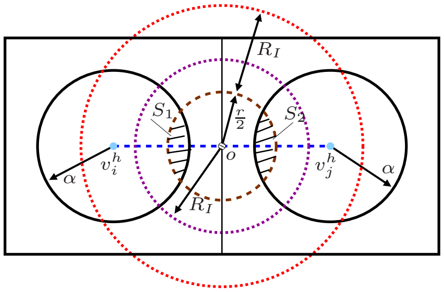


Fig. 3. Calculation of the first hitting time under EIHP Mobility.

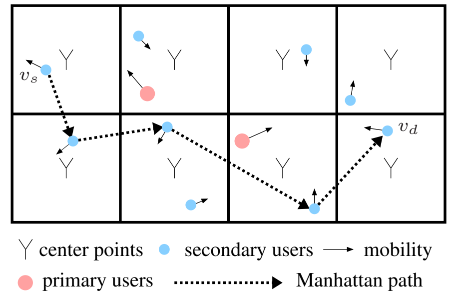


Fig. 4. A Manhattan path between v_s and v_d under EIHP mobility.

By letting $k = \tau t$, $\mathbb{P}(N(t) \leq \tau t) = \mathbb{P}(\sum_{i=1}^{\tau t} \rho_i > t) = \mathbb{P}(\sum_{i=1}^{k} r_i > \frac{k}{\tau})$. The last item is obviously bounded by some exponential variable considering $\{\rho_i\}_{i=1}^{\infty}$ are exponentially bounded. This completes the proof. ■

We next present our main result on the tail distribution of the dissemination latency $\mathcal{T}_d$ under EIHP mobility.

Proposition 1: Given $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_E, \Psi, \alpha), (\lambda, \lambda_p)]$ with finite users, if $\alpha > \frac{\sqrt{\frac{1}{\lambda}} - r}{2}$, the tail of the dissemination latency $\mathcal{T}_d$ is stochastically dominated by a Gamma distribution, $\Gamma(2\sqrt{n}, c_2)$, and the tail of the *dissemination speed* $\mathcal{S}_d$ is *stochastically dominated* by $\Gamma(\sqrt{\lambda}\mathcal{D}, \frac{c_2}{\mathcal{D}})$ for some positive constant c_2; otherwise, $\mathcal{T}_d$ has a heavy tail and $\mathbb{P}(\mathcal{T}_d = \infty) > 0$.

Proof: As the end to end latency, $\mathcal{T}_d$ is clearly bounded by the transmission delay along any path from the source v_s to destination v_d. Theorem 2 shows that, if $\alpha > \frac{\sqrt{\frac{1}{\lambda}} - r}{2}$, a link exists between two neighboring secondary users with positive probability. Therefore, we can identify a *Manhattan path* through which v_s first transmits the message vertically until the message reaches the secondary user whose center point has the same horizontal coordinate with v_d^c, and then transmits the message horizontally to v_d as shown in Fig. 4. Denote $\{X_k\}_{k=1}^{\infty}$ as a sequence of random variables with identical distributions as the first hitting time between neighboring secondary users. Note that a *Manhattan path* consists of at most $2\sqrt{n}$ communication links and thus $\mathcal{T}_d \leq \sum_{k=1}^{2\sqrt{n}} X_k$.

The next challenge is that the first hitting time of neighboring links, i.e., X_i and X_{i+1} are not independent. To tackle this challenge, we assume that after receiving the message, each secondary user will hold this message for a *renewal* time ρ before it tries to relay the message. Let $\{\rho_i\}_{i=1}^{\infty}$ be a sequence of *renewals* and $Y_k = X_k + \rho_k$. It is clear that $\mathcal{T}_d \leq \sum_{k=1}^{2\sqrt{n}} Y_k$. Note that Y_k is bounded by $exponential(c_2)$ (since both X_k and ρ_k are both exponentially bounded) and $\{Y_k\}_{k=1}^{\infty}$ are clearly independent. Let $\{\hat{Y}_k\}_{k=1}^{\infty}$ be a sequence of independent random variables distributed as $exponential(c_2)$, we have

$$\mathbb{P}(\mathcal{T}_d > t) \leq \mathbb{P}(\sum_{k=1}^{2\sqrt{n}} Y_k > t) \leq \mathbb{P}(\sum_{k=1}^{2\sqrt{n}} \hat{Y}_k > t), \quad (4)$$

where the last inequality is from Lemma 1 and *coupling* (Definition 4). By the *moment generating function* technique [27], we know that Y follows a Gamma distribution, $\Gamma(2\sqrt{n}, c_2)$. This completes the proof for $\mathcal{T}_d$.

To further describe *how fast* information can be disseminated, we study the dissemination speed $\mathcal{S}_d = \frac{\mathcal{T}_d}{\mathcal{D}}$. We need first to specify the *distance* $\mathcal{D}$ between v_s and v_d. As analyzed above, since for any v_i under EIHP mobility, the number of secondary users that can communicate with v_i directly is finite, and thus when v_d is beyond the transmission range of v_i, hop by hop communication is necessary, in which case *transmission hops* can describe "how far" more accurately than the Euclidean distance. Therefore, $\mathcal{D}$ here denotes the *Manhattan distance* between v_s^c and v_d^c by which the maximum number of transmission hops between v_s and v_d can be expressed as $\sqrt{\lambda}\mathcal{D}$. Then it follows $\mathcal{S}_d = \frac{\mathcal{T}_d}{\mathcal{D}} \leq \frac{\sum_{K=1}^{\sqrt{\lambda}\mathcal{D}} Y_k}{\mathcal{D}}$. Similarly, let $Y' = \frac{\sum_{K=1}^{\sqrt{\lambda}\mathcal{D}} \hat{Y}_k}{\mathcal{D}}$ and $\mathbb{P}(\mathcal{S}_d > t) \leq \mathbb{P}(Y' > t)$ by *coupling*. Then Y' is distributed as $\Gamma(\sqrt{\lambda}\mathcal{D}, \frac{c_2}{\mathcal{D}})$ from [27], which obtains the result for $\mathcal{S}_d$.

When $\alpha < \frac{\sqrt{\frac{1}{\lambda}} - r}{2}$, Theorem 2 says that the first hitting time between any two secondary users $\mathcal{T}_h(v_i, v_j) = \infty$. Therefore, $\mathcal{T}_d = \infty$, which completes the proof. ■

Remark 5: Proposition 1 shows that the tail of the dissemination speed $\mathcal{S}_d$ is bounded by $Y' \sim Gamma(\sqrt{\lambda}\mathcal{D}, \frac{c_2}{\mathcal{D}})$. Note that the mean $E(Y') = \sqrt{\lambda}c_2$ and the variance, $var(Y') = \lambda c_2^2/\mathcal{D}$. As $\mathcal{D}$ increases, $var(Y') \to 0$, which leads to $Y' = \sqrt{\lambda}c_2$ for large $\mathcal{D}$. Intuitively, this implies that in large CRNs where $\mathcal{D}$ is usually large, $\mathcal{S}_d$ may be bounded by some constant and its tail disappears as $\mathcal{D}$ increases. Actually, in Section V, we will rigorously prove that in large CRNs, $\mathcal{S}_d$ approaches to some constant, which agrees with our intuition here.

In addition to the distribution, we further study the average delay $E(\mathcal{T}_d)$. We begin with studying the average first hitting time $E(T_h(v_i, v_j))$. Based on $E(T_h(v_i, v_j)) = \int t \cdot d\mathbb{P}(T_h(v_i, v_j) > t)$, we obtain an upper bound on the average first hitting time from Eq. (3). From the proof of Proposition 1, we have $E(\mathcal{T}_d) \leq \sum_{k=1}^{2\sqrt{n}} E(T_h(v_i, v_j))$, which presents an upper bound on the average delay $E(\mathcal{T}_d)$. Note that this result is for generic mobility and we can derive more accurate delay for specific models. For example, We next derive the average delay for discrete models. Particularly, if we assume that at each time $t = 0, 1, 2, \ldots$, the nodes are independently distributed, and independent of all locations at previous time slot (note that such discrete assumption is well adopted and the i.i.d models in [16] is one of such examples), then we have:

Corollary 1: Given $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_E, \Psi, \alpha), (\lambda, \lambda_p)]$ with fi-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2015.2398420, IEEE Transactions on Mobile Computing

nite users, if $\alpha > \frac{\sqrt{\frac{1}{\lambda}}-r}{2}$, the expected delay $E(\mathcal{T}_d) < \frac{2\sqrt{n}}{\varepsilon}$ for discrete models, where ε is given by Eq.(1) and n is the number of secondary users.

Proof: Note that the probability that there exists no communication link between v_i and v_j at any time slot t $\varepsilon = \mathbb{P}(\mathcal{E}_t)$ is given by Eq.(1). Thus, $\mathcal{P}(T_h(v_i, v_j) = k) = \varepsilon^{k-1} \cdot (1 - \varepsilon)$ and

$$E(T_h(v_i, v_j)) = \sum_{k=1}^{\infty} \mathcal{P}(T_h(v_i, v_j) = k) \cdot k = \frac{1}{\varepsilon}. \quad (5)$$

Therefore, based on the proof of Proposition 1, we have

$$E(\mathcal{T}_d) \le \sum_{k=1}^{2\sqrt{n}} E(T_h(v_i, v_j)) = \frac{2\sqrt{n}}{\varepsilon}. \quad (6)$$

This completes the proof. ■

B. Distribution of $\mathcal{T}_d$ under PIHP Mobility $\mathcal{M}(\Phi_P, \Psi, \alpha)$

Note that the main difference between PIHP and EIHP mobility is that home points in the former are randomly located, and thus for neighboring secondary users v_i and v_j with $d_c(v_i, v_j) = \sqrt{\frac{1}{\lambda}}$, $d_h(v_i, v_j) \neq \sqrt{\frac{1}{\lambda}}$ in PIHP mobility. But under PIHP mobility, $d_h(v_i, v_j)$ is still bounded and we have $\mathbb{P}(d_h(v_i, v_j) \le \sqrt{\frac{5}{\lambda}}) = 1$. Thus, by similar proof to Theorem 2, we are able to see that for any v_i and v_j with $d_c(v_i, v_j) = \sqrt{\frac{1}{\lambda}}$, if $\alpha > \frac{\sqrt{\frac{5}{\lambda}}-r}{2}$, the first hitting time $T_h(v_i, v_j)$ is *exponentially* bounded; and if $\alpha < \frac{\sqrt{\frac{5}{\lambda}}-r}{2}$, $\mathbb{P}(T_h(v_i, v_j) = \infty) > 0$. Therefore, through the similar proof as that of Proposition 1, we have the following results about the dissemination latency $\mathcal{T}_d$ and speed $\mathcal{S}_d$ in finite CRNs under EIHP mobility:

Proposition 2: Given $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_P, \Psi, \alpha), (\lambda, \lambda_p)]$ with finite users, if $\alpha > \frac{\sqrt{\frac{5}{\lambda}}-r}{2}$, the tail of $\mathcal{T}_d$ is stochastically dominated by a Gamma distribution $\Gamma(2\sqrt{n}, c_3)$ and the tail of $\mathcal{S}_d$ is *stochastically dominated* by $\Gamma(\sqrt{\lambda}\mathcal{D}, \frac{c_3}{\mathcal{D}})$; otherwise, $\mathcal{T}_d$ has a heavy tail and $\mathbb{P}(\mathcal{T}_d = \infty) > 0$. c_3 is some positive constant.

Remark 6: In Propositions 1 and 2, the distance $\mathcal{D}$ between the source and destination nodes v_s and v_d has been considered as *Manhattan distance* between their *center points* v_s^c and v_d^c. Note that under both EIHP and PIHP mobility, the mobile region of any secondary user v_i is constrained to the coverage of home point v_i^h and thus the number of nodes that can communicate with v_i directly is *finite*. Hence information can only be delivered from the source to the destination *hop by hop* under these two mobility. Thus *Manhattan distance*, which characterizes the number of transmission hops between the source and destination, is an appropriate metric to measure the *distance* between nodes v_s and v_d. Also note that we can derive average delay for PIHP mobility by similar proof to that of EIHP model. We skip this due to the page limit.

C. Distribution of $\mathcal{T}_d$ under HHP Mobility $\mathcal{M}(\Phi_H, \Psi, \alpha)$

When HHP mobility is considered, home points are homogeneously distributed in the whole network Ω_n. Therefore, unlike EIHP and PIHP mobility, the distance between home points of secondary users v_i and v_j is homogeneous and may be any value in the interval $(0, \sqrt{\frac{n}{2\lambda}})$ (see Fig. 5, where Ω_n is a torus surface without border effect). We next show that, to overcome the *randomness* of $d_h(v_i, v_j)$, secondary users need to move around the whole network (large mobility capability α) to eliminate the *heavy tail* of the first hitting time.

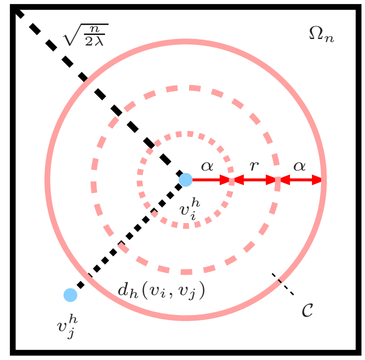


Fig. 5. An illustration of the first hitting time in HHP mobility.

Lemma 3: Given a CRN $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_H, \Psi, \alpha), (\lambda, \lambda_p)]$ with finite users, if $\alpha > \frac{\sqrt{\frac{n}{2\lambda}}-r}{2}$, the first hitting time $T_h(v_i, v_j)$ is exponentially bounded; and if $\alpha < \frac{\sqrt{\frac{n}{2\lambda}}-r}{2}$, $T_h(v_i, v_j)$ has a heavy tail and $\mathbb{P}(\mathcal{T}_d = \infty) > 0$.

Proof: As shown in Fig. 5, there may exist a communication link between v_i and v_j (i.e., $\mathbb{P}(T_h(v_i, v_j) < \infty) = 1$), if and only if v_j^h is located in the solid circle $\mathcal{C}$ centered at v_i^h. Thus the probability that v_j^h is distributed outside $\mathcal{C}$ (i.e., $\mathbb{P}(T_h(v_i, v_j) = \infty) > 0$), will incur a heavy tail of $T_h(v_i, v_j)$ (that is, $E(T_h(v_i, v_j)) = \infty$). Therefore, to eliminate the *heavy tail*, $\mathcal{C}$ must cover the whole network Ω_n, which requires $2\alpha + r > \sqrt{\frac{n}{2\lambda}} \Rightarrow \alpha > \frac{\sqrt{\frac{n}{2\lambda}}-r}{2}$. When $\alpha > \frac{\sqrt{\frac{n}{2\lambda}}-r}{2}$, v_i and v_j may communicate with each other with some positive probability at any time. Hence, with the similar proof of Theorem 2, we can show that $T_h(v_i, v_j)$ is exponentially bounded. This completes the proof. ■

Different from the above two scenarios (i.e., EIHP and PIHP), under HHP mobility, any secondary user v_i may receive the message directly from the source v_s, and any v_i that carries the message may in turn copy this message to all secondary users it encounters along its trajectory. Hence, we cannot apply the coupling method in calculating $\mathcal{T}_d$ hop by hop along the end to end path for HHP mobility. Instead, we use a stochastic model to analyze $\mathcal{T}_d$ and obtain the following result:

Proposition 3: Given $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_H, \Psi, \alpha), (\lambda, \lambda_p)]$ with finite users, if $\alpha < \frac{\sqrt{\frac{n}{2\lambda}}-r}{2}$, $\mathcal{T}_d$ has a heavy tail and $\mathbb{P}(\mathcal{T}_d = \infty) > 0$; and if $\alpha > \frac{\sqrt{\frac{n}{2\lambda}}-r}{2}$, the tail of $\mathcal{T}_d$ is stochastically dominated by a *Gamma distribution.*

Proof: When $\alpha < \frac{\sqrt{\frac{n}{2\lambda}}-r}{2}$, there exists some positive probability that all the home points $\{v_j^h, j \neq s\}$ are located outside the circle centered at v_s^h with radius $2\alpha + r$, which implies $\mathbb{P}(\mathcal{T}_d = \infty) > 0$ and thus a heavy tail.

When $\alpha > \frac{\sqrt{\frac{n}{2\lambda}}-r}{2}$, Lemma 3 shows that the tail of the first hitting time $\mathcal{T}_h(v_i, v_j)$ between any v_i and v_j is stochastically

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2015.2398420, IEEE Transactions on Mobile Computing

dominated by $exponential(c_4)$ for some constant c_4. If we can show that when $\mathcal{T}_h(v_i, v_j)$ is distributed as $exponential(c_4)$, the tail of the resulting dissemination latency $\mathcal{T}'_d$ is stochastically dominated by a *Gamma distribution*, then by Lemma 1 and *coupling*, which shows that $\mathbb{P}(\mathcal{T}_d > t) < \mathbb{P}(\mathcal{T}'_d > t)$ for large t, we complete the proof.

Assume $\mathcal{T}_h(v_i, v_j) \sim exponential(c_4)$ for any v_i and v_j. Denote ζ as the number of secondary users, which carry the message sent by source v_s before this message is successfully delivered to the destination v_d. The proof is based on modeling ζ as an absorbing finite-state Markov chain. The Markov chain consists of states $k = 0, 1, 2, \ldots, n-1$. The state $k > 0$ denotes $\zeta = k$ and the state 0 denotes the absorbing state that v_d successfully receives the message as shown in Fig. 6.

When secondary users hit each other, messages will be forwarded to the ones without a copy of the message. Therefore, when there are k secondary users carrying the message, the message is sent to another secondary user at rate $c_4 k(n-1-k)$ with the transition from k to $k+1$, and to destination v_d at rate $c_4 k$ with transition from k to 0, as shown in Fig. 6. The chain jumps from state k to $k+1$ with probability $\frac{n-1-k}{n-k}$ and transits from k to 0 with probability $\frac{1}{n-k}$. The sojourn time S_k in state k is exponentially distributed with intensity $c_4 k(n-k)$. Since $S_1, S_2, \ldots, S_{n-1}$ are mutually independent. Thus $\mathbb{P}(\zeta = k) = \frac{1}{n-k}\prod_{j=1}^{k-1}\frac{n-1-j}{n-j} = \frac{1}{n-1}$. Conditioning $\mathcal{T}'_d$ on ζ, we have $\mathcal{T}'_d|(\zeta = k) = \sum_{j=1}^{k} S_j$, which is clearly exponentially bounded and therefore, $\mathbb{P}(\mathcal{T}'_d > t) = \sum_{k=1}^{n-1}\mathbb{P}(\zeta = k)\mathbb{P}(\mathcal{T}'_d|(\zeta = k) > t) = \frac{1}{n}\sum_{k=1}^{n-1}\mathbb{P}(\mathcal{T}'_d|(\zeta = k) > t)$ is exponentially bounded (note that the sum of exponential variables is still exponentially distributed). That is, the tail of $\mathcal{T}'_d$ is bounded by $exponential(c_5) = Gamma(1, c_5)$ for some positive constant c_5. This completes the proof. ∎

Remark 7: Unlike EIHP and PIHP mobility, given $\alpha > \frac{\sqrt{\frac{n}{2\lambda}}-r}{2}$, home points under HHP mobility are homogeneously distributed and secondary users move around the whole network according to some stationary process. Thus $\mathcal{T}_d$ is homogeneous for any pair of secondary users and the *distance* between v_s and v_d $\mathcal{D}$ has no obvious impact on $\mathcal{T}_d$. Therefore here we only study the distribution of the dissemination latency $\mathcal{T}_d$.

We further study the expected delay in HHP mobility model. Based on the proof of Proposition 3, we have

$$\begin{aligned} E(\mathcal{T}_d) &= E(\mathcal{T}'_d) = E(\mathcal{T}'_d|(\zeta = k)) \\ &= \sum_{k=1}^{n-1} E(\mathcal{T}'_d|(\zeta = k))\mathbb{P}(\zeta = k) \\ &= \frac{1}{n-1}\sum_{k=1}^{n-1} E(\mathcal{T}'_d|(\zeta = k)) \\ &= \frac{1}{n-1}\sum_{k=1}^{n-1}\sum_{j=1}^{k} E(S_j) \\ &= \frac{c_4}{n-1}\sum_{k=1}^{n-1}\sum_{j=1}^{k} j(n-j) \\ &= \frac{c_4 \cdot n^2 \cdot (n+1)}{12}, \end{aligned} \quad (7)$$

where $c_4 = E(\mathcal{T}_h(v_i, v_j))$, as shown in the proof of Proposition 3. Note that this result is for generic mobility and given specific mobility models, we can derive more accurate c_4 and thus more accurate expected delay $E(\mathcal{T}_d)$. For example, we have the following corollary for the discrete models used in the Corollary 1:

Corollary 2: Given $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_H, \Psi, \alpha), (\lambda, \lambda_p)]$ with finite users, if $\alpha > \frac{\sqrt{\frac{n}{2\lambda}}-r}{2}$, the expected delay $E(\mathcal{T}_d) < \frac{n^3(n+1)}{12(1-\pi R_I^2/(n/\lambda))^m \lambda\pi r^2}$ for discrete models, where λ, r and n are the density, communication range and the number of secondary users, R_I is the interference range of primary users and m is the number of channels, as defined in Section II.

Proof: Given the discrete models in Corollary 1, there exists a link between v_i and v_j at any time slot t when given v_i, v_j is located within r, and no primary users within distance R_I. Note that when $\alpha > \frac{\sqrt{\frac{n}{2\lambda}}-r}{2}$, nodes are uniformly distributed over the whole network, given the network is in a torus surface. Thus the probability that v_i can communicate with v_j at any time slot is given by $\frac{\lambda\pi r^2}{n}(1-\pi R_I^2/(n/\lambda))^m$ and $\mathcal{T}_h(v_i, v_j)$ is a geometric distribution over time slots. Therefore, $c_4 = E(\mathcal{T}_h(v_i, v_j)) = \frac{n}{\lambda\pi r^2(1-\pi R_I^2/(n/\lambda))^m}$. Thus the result by putting $c_4 = E(\mathcal{T}_h(v_i, v_j))$ into Eq. (7). This completes the proof. ∎

Remark 8: From Propositions 1, 2 and 3, it seems that the dissemination latency $\mathcal{T}_d$ is independent of the density of primary users.

We need to clarify here that this "independency" is related to the *stochastic anatomy*. However, the density of primary users may have a negative impact on dissemination latency $\mathcal{T}_d$ and thus the speed $\mathcal{S}_d$ (see simulation results in Fig. 7). Particularly, as shown in the proof of Theorem 2, the expected first hitting time $E(\mathcal{T}_h(v_i, v_j))$ and thus the expected dissemination latency $E(\mathcal{T}_d)$ (Corollaries 1 and 2) obviously increase as the number of primary users m increases (see Figs. 7). On the other hand, in terms of the statistically macroscopic structure (i.e., light-tailed distribution or heavy-tailed distribution), the proofs for Proposition 1, 2 and 3 demonstrate that $\mathcal{T}_d$ is independent of the primary users, a *dichotomy* structure on the *mobility radius* α in mobile CRNs under general mobility.

The results in Propositions 1, 2 and 3 collectively suggest that when secondary users are mobile under any subclass of the general mobility $\mathcal{M}(\Phi, \Psi, \alpha)$, there exists a *dichotomy structure* on the *mobility radius* α. Above the *cutoff point* the dissemination latency $\mathcal{T}_d$ is bounded by some *Gamma* distribution (light-tailed), and below which $\mathcal{T}_d$ has a heavy tail and $\mathbb{P}(\mathcal{T}_d = \infty) > 0$. The cutoff value also increases as the *spatial homogeneity* of home points increases (i.e., increases in the order EIHP, PIHP and HHP). This is equivalent to saying that the *cutoff point* exists in CRNs under the general mobility $\mathcal{M}(\Phi, \Psi, \alpha)$, which completes the proof of Theorem 1. We next further explain and validate the theoretical *cutoff anatomy* through simulations.

D. Simulation Results and Discussions

We perform a series of simulations to verify our theoretical analysis on the distributions of the first hitting time $\mathcal{T}_h$ and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2015.2398420, IEEE Transactions on Mobile Computing

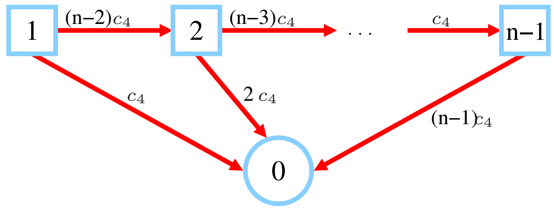


Fig. 6. Illustration of the Markov chain. Each state represents the number of secondary users carrying the message.

the dissemination latency $\mathcal{T}_d$. Because there is no real CRN trace that contains both information of primary user activities and secondary user mobility, we perform our simulations based on uniform distribution for primary users and generic mobility model for secondary users. More specifically, we here consider a finite CRN with $n = 16$ mobile secondary users within an area 2×2 square meters (i.e., $\lambda = 4$). The time is partitioned into unit slots. In each time slot, primary users are uniformly distributed at random within the network area and secondary users are uniformly distributed around their home points (i.e., Ψ is uniform). Furthermore, home points are uniformly distributed around the center points under PIHP mobility (i.e., Φ_P is uniform). The transmission range r of secondary users and the interference range R_I of primary users set $r = 0.1$ meter and $R_I = 0.3$ meter, respectively. The probability is calculated based on the average of 1000 independent simulations. Fig. 7 shows the complementary distribution (CCDF) of the dissemination latency $\mathbb{P}(\mathcal{T}_d > t)$ on a log-log scale for EIHP, PIHP and HHP models with different values of *mobility radius* α and the density of primary users λ_p. It is observed in Fig. 7 that as the density of primary users λ_p increases, the curves move right-ward, which indicates increasing expected dissemination latency. However, regardless of the value of λ_p, when $\alpha = 0.4$ meter, which is larger than the *cutoff point* under *EIHP* but smaller than those under PIHP and HHP, the dissemination latency $\mathcal{T}_d$ has a light tail under EIHP but *heavy tails* under PIHP. As α increases to 0.8 meter, which is larger than the *cutoff point* in PIHP, but still less than that in HHP, the *heavy tail* of $\mathcal{T}_d$ in PIHP disappears, but $\mathcal{T}_d$ in HHP presents a *heavy* tail. The complementary distribution (CCDF) of the first hitting time $\mathbb{P}(\mathcal{T}_h > t)$ on a log-log scale for EIHP, PIHP and HHP are shown in Fig. 8, which demonstrates a similar quantitative behavior to Fig. 7. In particular, a similar dichotomy property of the tails of $\mathcal{T}_h$ on mobility radius is observed for all three mobility models in Fig. 8. These results are in good agreement with our theoretical analysis in Propositions 1, 2 and 3, and arguments in Remark 8.

Note that the major objective of this paper is to build analytical models for node mobility and investigate the stochastic properties of information delivery latency, in finite as well as large-scale CRNs. Therefore, in this simulation for finite networks, we focus on illustrating the "dichotomy" property of the first hitting time $\mathcal{T}_h$ and the dissemination latency $\mathcal{T}_d$ over the node mobility.

V. The Scalability of $\mathcal{T}_d$ in Large CRNs

We next prove Theorem 2, which states that the dissemination latency $\mathcal{T}_d$ asymptotically scales linearly with the dissemination *distance* in large mobile CRNs. We studied the distribution of $\mathcal{T}_d$ in Section IV and demonstrated that $E(\mathcal{T}_d) \to \infty$ as the network size grows large, which implies that the distribution of $\mathcal{T}_d$ cannot be used to measure *how fast* information is disseminated in *large* CRNs. Therefore in large CRNs, instead of the distribution, we will investigate the *scalability* of $\mathcal{T}_d$, i.e., the scaling behavior of $\mathcal{T}_d$ with respect to the *distance* $\mathcal{D}$ between the source v_s and destination v_d, which can be characterized by the dissemination *speed* $\mathcal{S}_d = \frac{\mathcal{T}_d}{\mathcal{D}}$.

The results in finite networks have presented an implication that the tail of $\mathcal{S}_d$ may *disappear* as the network size increases (see Remark 5 in Section IV), which will be validated in this section. We must emphasize that the derivation in Section IV is based on the assumption that the number of nodes n is finite. However, n may approach to infinity in large CRNs. Therefore, the proofs and results in finite CRNs may not be applied here. Instead, we will use the *large number theory* to demonstrate that $\mathcal{T}_d$ asymptotically scales linearly with $\mathcal{D}$, i.e., $\lim_{\mathcal{D}\to\infty} \mathcal{S}_d$ converges to some positive constant. Specifically, the main tool used in this section is the following theorem:

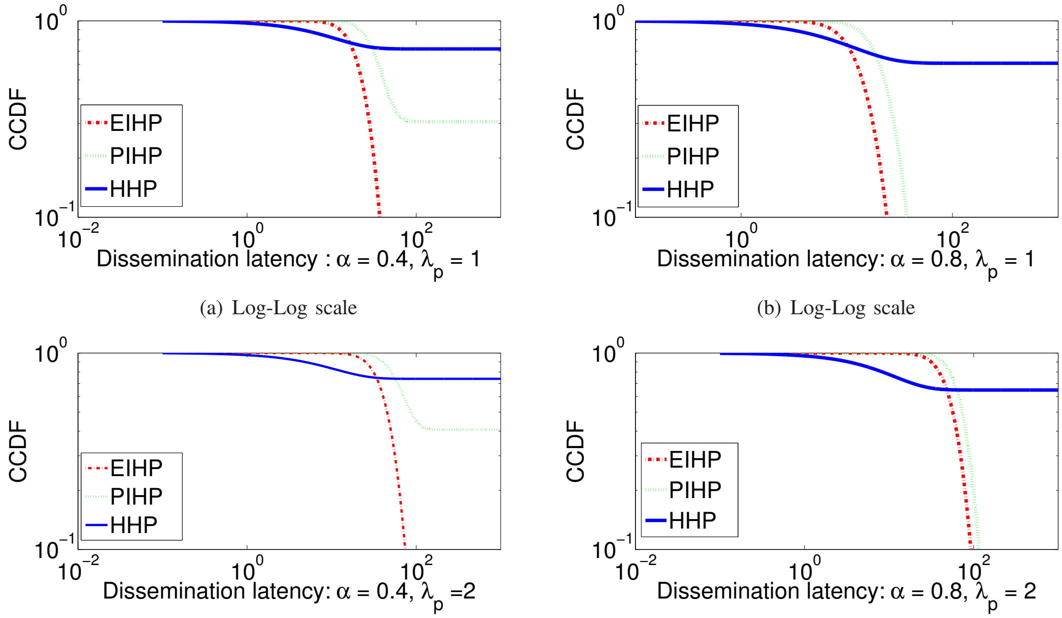
Theorem 3 (Liggett's subadditive ergodic theorem, [28]): Let $\{\mathcal{Z}_{h,q}\}$ be a collection of random variables indexed by integers satisfying $0 \leq h < q$. Suppose $\{\mathcal{Z}_{h,q}\}$ has the following properties: (i) $\mathcal{Z}_{0,q} \leq \mathcal{Z}_{0,h} + \mathcal{Z}_{h,q}$, (ii) For each q, $\mathbb{E}(|\mathcal{Z}_{0,q}|) < \infty$ and $\mathbb{E}(\mathcal{Z}_{0,q}) \geq cq$ for some constant $c > -\infty$. (iii) The distribution of $\{\mathcal{Z}_{h,h+k:k\geq 1}\}$ does not depend on h. (iv) For each $k \geq 1$, $\{\mathcal{Z}_{qk,(q+1)k} : q \geq 0\}$ is a stationary sequence. Then: (a) $\zeta = \lim_{q\to\infty} \mathbb{E}(\mathcal{Z}_{0,q})/q = inf_{q\geq 1} E(\mathcal{Z}_{0,q})/q$. (b) $\mathcal{Z} = \lim_{q\to\infty} \mathcal{Z}_{0,q}/q$ exists almost surely. (c) $\mathbb{E}(\mathcal{Z}) = \zeta$. Furthermore, (v) If $k \geq 1$, $\{\mathcal{Z}_{qk,(q+1)k} : q \geq 0\}$ are ergodic, then (d) $\mathcal{Z} = \zeta$ almost surely.

Liggett's theorem provides a method to study the *limiting behavior* of a large random process, which will be used to study the *limit* of the dissemination speed $\mathcal{S}_d$ in this section. Before we proceed, we first make some clarification about our models. To study the dissemination latency $\mathcal{T}_d$ and speed $\mathcal{S}_d$ in large CRNs, we progressively increase the number of secondary users n in Ω_n. The homogeneously distributed secondary users are asymptotically distributed as a Poisson point process with density λ. Similarly, to facilitate the analysis, we will prove the scalability of $\mathcal{T}_d$ under EIHP, PIHP and HHP models respectively.

A. $\mathcal{S}_d$ under EIHP and PIHP Mobility in Large CRNs

Since the study of the first hitting time $\mathcal{T}_h(v_i, v_j)$ between neighboring secondary users v_i and v_j under EIHP or PIHP models in Section IV is independent of the network size, these results still hold for large CRNs. That is, when the *mobility radius* $\alpha > \frac{\sqrt{\frac{1}{\lambda}} - r}{2}$ for EIHP or $\alpha > \frac{\sqrt{\frac{5}{\lambda}} - r}{2}$ for PIHP mobility, the first hitting time $\mathcal{T}_h(v_i, v_j)$ is exponentially bounded. Otherwise, $\mathcal{T}_h(v_i, v_j)$ and thus $\mathcal{T}_d$ have heavy tails independent of the transmission distance $\mathcal{D}$(i.e., $\mathcal{T}_d$ is unscalable with $\mathcal{D}$).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2015.2398420, IEEE Transactions on Mobile Computing



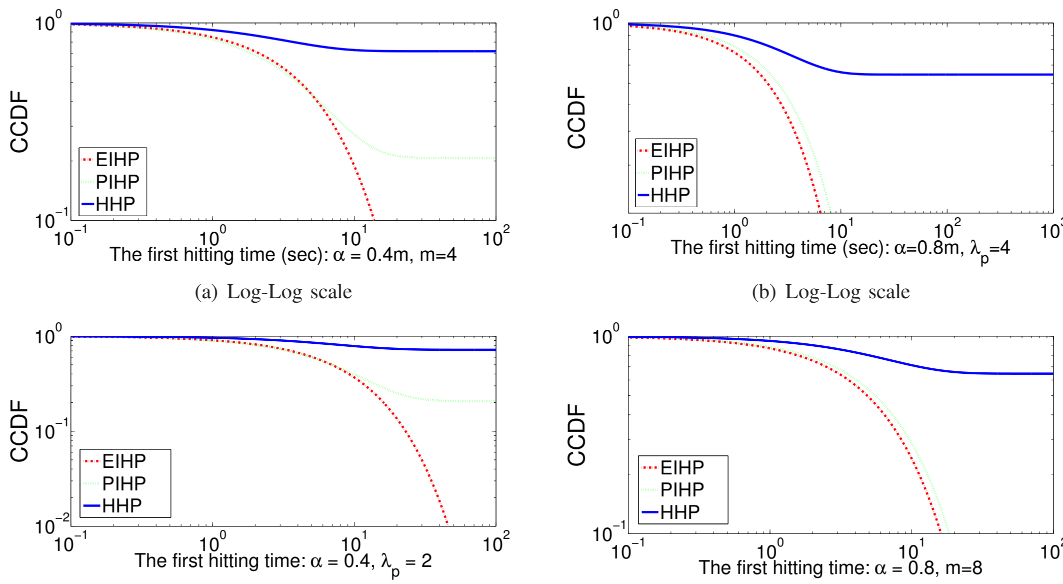
(a) Log-Log scale

(b) Log-Log scale

(c) Log-Log scale

(d) Log-Log scale

Fig. 7. CCDF of the dissemination latency $\mathcal{T}_d$ under general mobility.



(a) Log-Log scale

(b) Log-Log scale

(c) Log-Log scale

(d) Log-Log scale

Fig. 8. CCDF of the first hitting time $\mathcal{T}_h(v_i, v_j)$ between neighboring secondary users v_i and v_j.

Therefore, we only need to study the scalability of $\mathcal{T}_d$ with *exponentially bounded* $\mathcal{T}_h(v_i, v_j)$. In addition, when it comes to the *distance* $\mathcal{D}$ between the source v_s and destination v_d, it can be any *p-norm* metric function and we consider two of the most popular metrics *transmission hops* and *Euclidean distance*. As analyzed in Section IV-A and IV-B, hop by hop communication is necessary for **EIHP** and **PIHP** mobility, which indicates that *transmission hops* can describe "how far" more accurately than the Euclidean distance in these two models. Therefore, $\mathcal{D}$ here denotes the *Manhattan distance* between v_s^c and v_d^c by which the maximum number of transmission hops between v_s and v_d can be expressed as $\sqrt{\lambda}\mathcal{D}$. To proceed, we make the following assumption about information dissemination between secondary users v_i and v_j under EIHP and PIHP models, by ignoring border effects.

Assumption 1: Information may be delivered directly between secondary users v_i and v_j only when $d_c(v_i, v_j) = \sqrt{\frac{1}{\lambda}}$.

Remark 9: According to the interference model in Section II-A1, v_i may communicate with v_j directly when $2\alpha + r > d_h(v_i, v_j)$. Given $\alpha > \frac{\sqrt{\frac{1}{\lambda}}-r}{2}$ for **EIHP** (or $\alpha > \frac{\sqrt{\frac{5}{\lambda}}-r}{2}$ for **PIHP**), v_i and v_j with $d_c(v_i, v_j) = \sqrt{\frac{1}{\lambda}}$ clearly satisfy

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2015.2398420, IEEE Transactions on Mobile Computing

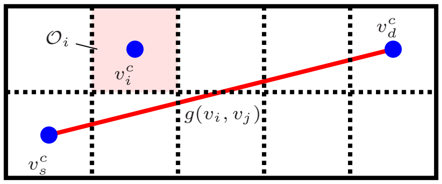


Fig. 9. Illustration of the dissemination direction.

the requirement $2\alpha + r > d_h(v_i, v_j)$, which implies that Assumption 1 conforms to the interference constraint. However, Assumption 1 is clearly more restrictive than the interference constraint since the former ignores some *border effect*. For example, given v_i and v_k under **EIHP** with $d_h(v_i, v_k) = d_c(v_i, v_k) = \sqrt{\frac{2}{\lambda}}$ and $\alpha > \frac{\sqrt{\frac{2}{\lambda}} - r}{2}$, v_i may communicate directly with v_k by the interference constraint but this possible communication link is ignored by Assumption 1. Assumption 1 simplifies the notations and description of our proof, and thus highlights the fundamental properties of the dissemination latency $\mathcal{T}_d$ in large CRNs. However, as shown later, our proof techniques do not require Assumption 1 and hence the derived results can be easily extended to the scenario which takes the *border effect* into account. We next present our main results.

Proposition 4: Given $\alpha > \frac{\sqrt{\frac{1}{\lambda}} - r}{2}$ for a large network $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_E, \Psi, \alpha), (\lambda, \lambda_p)]$ (or $\alpha > \frac{\sqrt{\frac{5}{\lambda}} - r}{2}$ for $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_P, \Psi, \alpha), (\lambda, \lambda_p)]$), there exists some finite constant κ such that $\mathbb{P}(\lim_{\mathcal{D}\to\infty} \mathcal{S}_d = \lim_{\mathcal{D}\to\infty} \frac{\mathcal{T}_d}{\mathcal{D}} = \kappa) = 1$.

Because of the page limit, we only prove the result for EIHP mobility. The proof for PIHP mobility is similar. To initiate the proof of Proposition 4, we first define the following notations. In Section IV, we argued that $\mathcal{D}$ should be specified as *Manhattan distance* between the *center* points of the source and destination under EIHP and PIHP mobility models. Here we further denote $d_c^{(1)}(v_i, v_j)$ as the *Manhattan distance* between center points v_i^c and v_j^c for any v_i and v_j. Let N_h be the set of secondary users with h hops away from the source v_s, i.e., $N_h \triangleq \{v_i : d_c^{(1)}(v_s, v_i) = h\sqrt{\frac{1}{\lambda}}\}$ (note that the length of each hop is $\sqrt{\frac{1}{\lambda}}$). Denote the straight line joining v_s^c and v_d^c $g(v_s, v_d)$ as the dissemination direction between v_s and v_d (as shown in Fig. 9) and $v(h)$ as the secondary user which is h hops away from v_s on the dissemination direction $g(v_s, v_d)$, i.e., $v(h) \triangleq \{v_i : v_i \in N_h \text{ and } g(v_s, v_d) \cap \mathcal{O}_i \neq \emptyset\}$ ($\mathcal{O}_i$ is the cell associated with v_i^c, see the shaded region in Fig. 9). We next define the collection of indexed variables by $\mathcal{T}_{h,q}$ as the *dissemination latency* between $v(h)$ and $v(q)$ (thus $\mathcal{T}_d = \mathcal{T}_{0,\mathcal{D}\sqrt{\lambda}}$). Therefore, Proposition 4 is equivalent to showing $\mathbb{P}(\lim_{q\to\infty} \frac{\mathcal{T}_{0,q}}{q} = \kappa\sqrt{\lambda}) = 1$, which will be proved by Liggett's theorem next.

Particularly, if we can verify that the sequence $\{\mathcal{T}_{h,q}, h \leq q\}$ satisfies the conditions $(i) - (v)$ of Liggett's theorem, we will finish our proof by the conclusions of Liggett's theorem. By definition, $\mathcal{T}_{0,q}$ is the shortest time that $v(q)$ will receive the message from $v(0)$, which is clearly at most $\mathcal{T}_{0,h} + \mathcal{T}_{h,q}$. Condition (i) is thus verified. We have shown that the first hitting time $\mathcal{T}_h(v_i, v_j)$ is exponentially bounded and thus $0 < E(\mathcal{T}_h(v_i, v_j)) < \infty$, which verifies condition (ii). Conditions (iii) and (iv) are clearly verified, as $\mathcal{T}_{h,q}$ is defined in a stationary way. The following lemma is to prove that the sequence $\mathcal{T}_{h,q}$ is ergodic, i.e., $\{\mathcal{T}_{h,q}, h \leq q\}$ satisfies condition (v). In fact, we show that $\mathcal{T}_{h,q}$ is mixing (i.e., roughly speaking, asymptotically independent), which is a stronger property than ergodicity.

Lemma 4: The sequence $\{\mathcal{T}_{q,q+1}, q \geq 0\}$ is mixing.

Proof: We compute $\mathcal{T}_{q,q+1}$ by the following construction: Denote $\mathcal{N}_{q,k}$ as the set of secondary users that are within k hops from $v(q)$, i.e., $\mathcal{N}_{q,k} \triangleq \{v_i : d_c^{(1)}(v(q), v_i) < k\sqrt{\frac{1}{\lambda}}\}$, and $\mathcal{T}_{q,q+1}^{(k)}$ as the transmission delay from $v(q)$ to $v(q+1)$ only using nodes $v \in \mathcal{N}_{q,k}$ as relays. Observe that

$$\lim_{k\to\infty} \mathbb{P}\Big(\mathcal{T}_{q,q+1}^{(k)} < t\Big) = \mathbb{P}\Big(\mathcal{T}_{q,q+1} < t\Big)$$

for all t. We can show that $\{\mathcal{T}_{q,q+1}\}$ is mixing by

$$\begin{aligned} &\lim_{k\to\infty} \mathbb{P}\Big((\mathcal{T}_{q,q+1} < t) \cap (\mathcal{T}_{q+2k,q+2k+1} < t')\Big) \\ &= \lim_{k\to\infty} \mathbb{P}\Big((\mathcal{T}_{q,q+1}^{(k)} < t) \cap (\mathcal{T}_{q+2k,q+2k+1}^{(k)} < t')\Big) \\ &= \lim_{k\to\infty} \mathbb{P}\Big(\mathcal{T}_{q,q+1} < t\Big)\mathbb{P}\Big(\mathcal{T}_{q+2k,q+2k+1}^{k} < t'\Big) \quad \forall t, t'. \end{aligned}$$

The second equality follows that $\mathcal{T}_{q,q+1}^{(k)}$ and $\mathcal{T}_{q+2k,q+2k+1}^{k}$ are independent, as they depend on non-intersected node sets $\mathcal{N}_{q,k}$ and $\mathcal{N}_{q+2k,k}$. ∎

We can see that $\{\mathcal{T}_{h,q}, h \leq q\}$ satisfies all the conditions of Liggett's theorem and thus prove Proposition 4.

Remark 10: Note that our proof for Proposition 4 does not depend on Assumption 1. Specifically, Assumption 1 simplifies the description and definition of the dissemination process $\{\mathcal{T}_{h,q}, h \leq q\}$, but the proof to show that $\{\mathcal{T}_{h,q}, h \leq q\}$ satisfy the conditions of Liggett's theorem is independent of Assumption 1. Without Assumption 1, we can still define dissemination process as $\{\mathcal{T}'_{h,q}, h \leq q\}$ similarly and then show that $\{\mathcal{T}'_{h,q}, h \leq q\}$ satisfy the conditions of Liggett's theorem by the same process. This completes the proof. Next we study the dissemination speed $\mathcal{S}_d$ under **HHP** mobility.

B. $\mathcal{S}_d$ under HHP Mobility in Large CRNs

In the previous analysis of information dissemination in large CRNs under EIHP or PIHP, a fundamental assumption is that the *mobility radius* α is large enough, which is actually used to ensure that the networks are *fully connected*. That is, there exists a communication path (may be dynamic over time) between any two secondary users with high probability. However, such a *full connectivity* requirement may be overly restrictive in large CRNs with HHP mobility. For example, it is well-known that to achieve the *full connectivity* in a large homogeneous networks with n static wireless nodes distributed as a Poisson process, the required transmission range is $\Theta(\sqrt{\log n})$ [24]. By ignoring the interference from primary users to secondary users, a large CRN under HHP with mobility radius α and transmission range r can be mapped as a static Poisson distributed homogeneous networks with transmission range $2\alpha + r$ [16]. This indicates that $2\alpha + r = \Theta(\sqrt{\log n})$ is required for *full connectivity*, which is impractical to be satisfied when the number of nodes n is large.

Therefore instead of *full connectivity*, we will investigate information dissemination in large CRNs under HHP mobility from a *percolation* perspective, which has been extensively studied recently [15], [16], [22], [29], [30]. The main result of percolation theory concerns a phase transition in the macroscopic behavior of large random networks [31], [32]. Specifically, it demonstrates that there exists a finite and positive value of the transmission range, or equivalently of the node spatial density, above which the network is percolated (supercritical) and under which the network is not percolated (subcritical). This is called *critical phenomenon* in percolation theory. When the network is *percolated*, there exists a *large connected component* (usually called *giant component*) of nodes spanning almost the entire network. When the network is not percolated, it consists only of small isolated components of nodes. Note that the ultimate goal of the *full connectivity* requirement is to ensure that information can be disseminated through the whole network (thus the network is considered functional). When the network is percolated, information can still be disseminated to the entire network (through the *giant component*). But the conditions to achieve the latter are much less restrictive (only requires *finite* transmission range r, or equivalently a finite α). Therefore, we will only study how fast information is disseminated in large *percolated* (instead of fully connected) CRNs under HHP.

Ren *et al.* [22] studied the percolation of a large static random CRN. Given the density and interference range of primary users, they identified a critical value on the transmission range of secondary users, above which the CRN is percolated. By mapping the CRN under HHP mobility to a large static random CRN (similar to [16]), we can derive the similar conditions for percolation of the former easily. Because of the page limit, we will skip this step and directly focus on the dissemination latency $\mathcal{T}_d$ and speed $\mathcal{S}_d$ in large *percolated* CRNs under HHP, by assuming that the mobility radius α or transmission range r is large enough to satisfy the percolation conditions. We have the following results on $\mathcal{T}_d$ and $\mathcal{S}_d$:

Proposition 5: Given $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_H, \Psi, \alpha), (\lambda, \lambda_p)]$, where the large CRN $\mathcal{F}_{m,n}$ is percolated [2], for any two nodes v_s and v_d in the *giant component* of $\mathcal{F}_{m,n}$, there is a finite and positive constant κ such that $\mathbb{P}(\lim_{\mathcal{D}\to\infty} \mathcal{S}_d = \lim_{\mathcal{D}\to\infty} \frac{\mathcal{T}_d}{\mathcal{D}} = \kappa) = 1$.

Proof: (Sketch.) Instead of proving our results by using *percolation theory* from the scratch, we will map our networks to the existing models [15], [16], [22], [29], [30] and thus use the existing results to save space. Particularly, Kong *et al.* [16] studied a percolated wireless homogeneous networks where *home points* $\{v_i^h\}$ are uniformly distributed and wireless nodes v_i is independently and uniformly mobile within the circular region $\mathcal{A}(v_i^h, \alpha)$ centered at v_i^h (called *initial positions* in [16]) with radius $\alpha > 0$. Kong *et al.* demonstrated that the limiting dissemination latency scales linearly with the dissemination distance $\mathcal{D}$ (*Euclidean distance* between home points) in such a percolated network. Note that the positions of *home points* and the mobility of secondary users in our network $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_H, \Psi, \alpha), (\lambda, \lambda_p)]$ are similar to those in [16], which motivates us to extend the result in [16] to our network. If we can validate the extension, we will finish the proof. We next verify the extension. There exist two main differences between our model $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_H, \Psi, \alpha), (\lambda, \lambda_p)]$ and the network in [16]. First, in our model, the secondary user v_i is independently mobile within $\mathcal{A}(v_i^h, \alpha)$ according to some stationary distribution Ψ (not necessarily the uniform distribution as in [16]). Furthermore, there exist primary users in our network, which interferes with the communications among secondary users. However, we find that the proofs in [16] require neither *uniform* distribution of v_i around v_i^h, nor non-interference from other nodes (e.g., primary users). Indeed, the fundamental requirement for proofs in [16] is that, given any two nodes v_i and v_j with $d_h(v_i, v_j) = \|v_i^h - v_j^h\| < 2\alpha + r$, the expected first hitting time $E(\mathcal{T}_h(v_i, v_j)) < \infty$. This requirement has been proven in Theorem 2, which indicates that the result of the dissemination speed in [16] can also be applied in $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_H, \Psi, \alpha), (\lambda, \lambda_p)]$. This completes the proof. ■

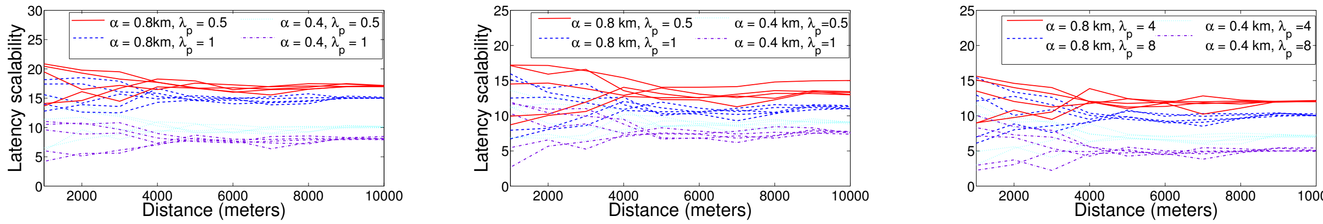
Remark 11: Propositions 4 and 5 together complete the proof for Theorem 2, which demonstrates that information spreads *linearly* in large *connected* (full connected or percolated) CRNs under general mobility and seem to be independent of primary users. Note that this *independency* refers to the *linearity*, not the *specific value*, of the dissemination speed. As shown in Propositions 4, 5 and Liggett's theorem, the specific value of the asymptotic speed is proportional to the expected first hitting time $E(T_h(v_i, v_j))$ between neighboring secondary users v_i and v_j, which adversely depends on the density of primary users (see Theorem 2). Furthermore, as shown in [22] and Fig. 11(a), the required transmission range r (or mobility radius α) for percolation increases as the density of the primary users increases. When the density of the primary users is larger than some threshold value, the CRN is not percolated *even with infinitely large r or α*. Therefore our dissemination speed analysis in Proposition 5 is based on the implicit assumption that the density of primary users is not high so that the network is percolated.

C. Simulation Results and Discussions

As simulation proves to be an effective tool in studying large-scale network, we conduct a series of simulations to further validate our theoretical results concerning the asymptotic latency. In these simulations, time is partitioned into unit slots. In each time slot, secondary users are independently and uniformly distributed around their home points. Fig. 10 shows the dissemination speed in CRNs under EIHP and PIHP models respectively, where the transmission range r of secondary users and the interference range R_I of primary users are set as $r = 0.1\ km$ and $R_I = 0.3\ km$, and the spatial densities of secondary users and primary users are set $\lambda = 4$ (per km^2) and $\lambda_p = 0.5$ (per km^2). As shown in Fig. 10(a) and 10(b), no matter how large the *mobility radius*

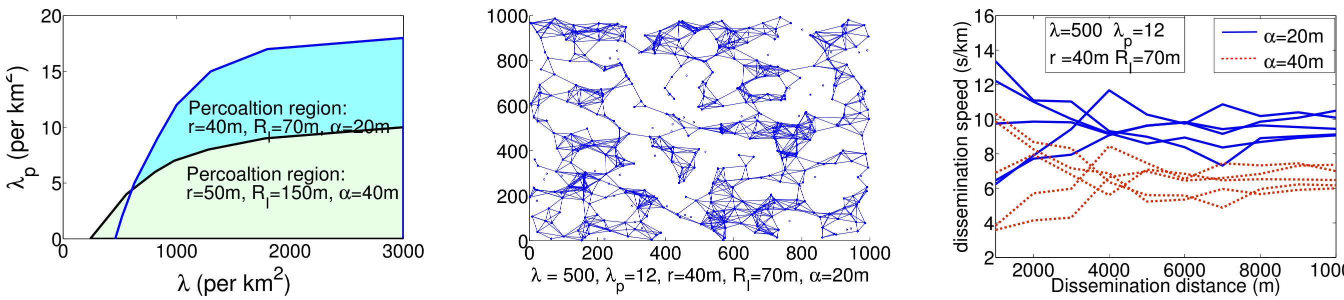
[2] [15] defines two types of percolation for mobile networks: permanent percolation and cumulative percolation. The former is that the network is percolated at any time instant; the latter is that the network is not percolated at any time instant, but it is percolated over time. When we mention percolation of mobile networks, we refer to the latter in this paper.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2015.2398420, IEEE Transactions on Mobile Computing



(a) EIHP (b) PIHP (c) With/without border effects.

Fig. 10. The dissemination speed $\mathcal{S}_d = \frac{\mathcal{T}_d}{\mathcal{D}}$ (s/km) based on five independent simulations.



(a) Percolation regions. (b) An example percolated CRN. (c) Dissemination speed in percolated CRNs.

Fig. 11. Percolation conditions and the dissemination speed in large percolated CRNs.

α is, the dissemination latency $\mathcal{T}_d$ scales linearly with the dissemination distance $\mathcal{D}$ (Manhattan distance) as $\mathcal{D}$ increases, which agrees with Proposition 4. Moreover, as the density of primary user λ_p increases, the curves in Fig. 10(a) and 10(b) move downward, which indicates decreasing dissemination speed. These observations also demonstrate that the density of primary adversely affects the dissemination speed, but has no influence on the linearity of dissemination speed in large networks. Fig. 10(c) compares the dissemination speed under EIHP mobility with and without taking Assumption 1 into account. The main difference is that the latency is smaller in the latter. However, we observe that their qualitative behaviors are similar, which verifies our arguments in Remarks 9 and 10. The dissemination speed in large percolated CRNs under HHP mobility is shown in Fig. 11. Particularly, Fig. 11(a) shows the simulated percolation conditions under different network parameters and Fig. 11(b) gives an example of such a percolated CRN. Fig. 11(c) shows that in percolated CRNs, the dissemination latency $\mathcal{T}_d$ scales linearly with the dissemination distance $\mathcal{D}$ (Euclidean distance) as $\mathcal{D}$ increases. Fig. 11 validates the Proposition 5 and arguments in Remark 11.

As we mentioned in Section IV-D, the major objective of this paper is to provide analytical models and study the stochastic properties of distribution latency $\mathcal{T}_d$, rather than obtaining the exact value of $\mathcal{T}_d$, which is not practical to achieve given so many random factors in the wireless networks. And for large CRNs, we investigate the scaling behavior of $\mathcal{T}_d$ with respect to the *distance* $\mathcal{D}$ between the source v_s and destination v_d. Therefore, as shown in Fig. 10 and Fig. 11, we focus on demonstrating the linear scalability of latency $\mathcal{T}_d$ for different mobility models in this simulation.

VI. Conclusions and Applications

We have studied in this paper the distribution of the information dissemination latency $\mathcal{T}_d$ in *finite* CRNs and the scalability of $\mathcal{T}_d$ in *large* CRNs under general mobility. We found that in finite networks, there exists a cutoff point on the *mobility radius* α of secondary users, above which the tail distribution of $\mathcal{T}_d$ is bounded by some Gamma distribution and below which $\mathcal{T}_d$ has a *heavy-tailed* distribution. When networks become large, the dissemination latency $\mathcal{T}_d$ is (linearly) scalable with respect to the dissemination distance. Our results demonstrate that when secondary users can move in a large region, a *Gamma* distributed (light-tailed) latency in finite networks, or a scalable latency in large networks, is achievable.

Our results encourage the deployment of CRNs for real-time and large applications. One possible application is CRN-based smart metering in the next generation smart grid. Smart meters can dynamically utilize unused spectrum to communicate with each other and with network gateway. The network gateway can connect with a spectrum database and determine which channels to use for smart meters based on their locations and bandwidth requirement. Upon CRN-based smart metering, accurate real time energy consumption information can be provided to user and utility companies in order to save energy and money. Another application is CRN-based vehicular communications for information dissemination. As radio frequencies allocated for vehicular networks are limited, vehicular users equipped with spectrum sensing technologies can use additional spectrum for emergency message dissemination. In an emergency situation, vehicles can coordinate with primary users around an accident area to opportunistically access

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2015.2398420, IEEE Transactions on Mobile Computing

spectrum. Under high vehicle mobility and large mobility radius, short dissemination latency is expected through CRN-based vehicle communications. In summary, CRN is promising in supporting large applications that require short latency, such as CRN-based smart metering and vehicular communication.

REFERENCES

[1] F. C. Commission, "Spectrum policy task force report, FCC 02-155." Nov 2002.

[2] W.-Y. Lee and I. Akyildiz, "Spectrum-Aware Mobility Management in Cognitive Radio Cellular Networks," in *IEEE Transactions on Mobile Computing*, vol. 11, no. 4, 2012, pp. 529–542.

[3] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey," *Computer Networks*, May 2006.

[4] Y. T. Hou, Y. Shi, and H. D. Sherali, "Optimal Spectrum Sharing for Multi-hop Software Defined Radio Networks," in *Proceedings of IEEE INFOCOM*, May 2007.

[5] C. Han and Y. Yang, "Understanding the Information Propagation Speed in Multihop Cognitive Radio Networks ," in *IEEE Transactions on Mobile Computing*, vol. 12, no. 6, 2013, pp. 1242–1255.

[6] D. Xu, X. Liu, and Z. Han, "Decentralized Bargain: A Two-Tier Market for Efficient and Flexible Dynamic Spectrum Access ," in *IEEE Transactions on Mobile Computing*, vol. 12, no. 9, 2013, pp. 1697–1711.

[7] D. Kalathil and R. Jain, "Spectrum Sharing through Contracts for Cognitive Radios ," in *IEEE Transactions on Mobile Computing*, vol. 12, no. 10, 2013, pp. 1999–2011.

[8] W. S. Jeon, J. A. Han, and D. G. Jeong, "A Novel MAC Scheme for Multichannel Cognitive Radio Ad Hoc Networks," in *IEEE Transactions on Mobile Computing*, vol. 11, no. 6, 2012, pp. 922–934.

[9] M. J. Neely and E. Modiano, "Capacity and Delay Tradeoffs for Ad-Hoc Mobile Networks," in *IEEE Transactions on Information Theory*, vol. 51, no. 6, June 2005.

[10] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-Delay Trade-off in Wireless Networks," in *Proceedings of IEEE INFOCOM*, Apr 2004.

[11] S. Toumpis and A. Goldsmith, "Large Wireless Networks under Fading, Mobility, and Delay Constraints," in *Proceedings of IEEE INFOCOM*, Apr 2004.

[12] H. Zhang, Z. Zhang, and H. Dai, "On the Capacity Region of Cognitive Multiple Access over White Space Channels," in *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 11, September 2013, pp. 2517–2527.

[13] R. Zheng, "Information Dissemination in Power-Constrained Wireless Networks," in *Proceedings of IEEE INFOCOM*, April 2006.

[14] Y. Xu and W. Wang, "The Speed of Information Propagation in Large Wireless Networks," in *Proceedings of IEEE INFOCOM*, April 2008.

[15] O. Dousse, P. Mannersalo, and P. Thiran, "Latency of Wireless Sensor Networks with Uncoordinated Power Saving Mechanisms," in *Proceedings of ACM MOBIHOC*, May 2004.

[16] Z. Kong and E. M. Yeh, "On the Latency for Information Dissemination in Mobile Wireless Networks," in *Proceedings of ACM MOBIHOC*, May 2008.

[17] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad hoc wireless networks," in *IEEE Transactions on Networking*, vol. 10, no. 2, August 2002, pp. 477–486.

[18] S. Toumpis and A. Goldsmith, "Delay and Capacity Trade-offs in Mobile Ad Hoc Networks: A Global Perspective," in *Proceedings of IEEE INFOCOM*, Apr 2006.

[19] A. A. Hanbali, A.A.Kherani, R. Groenevelt, P. Nain, and E. Altman, "Impact of Mobility on the Performance of Relaying in Ad Hoc Networks," in *Proceedings of IEEE INFOCOM*, Apr 2006.

[20] R. Moraes, H. Sadjadpour, and J. G.-L. Aceves, "Mobility-Capacity-Delay Trade-off in Wireless Ad Hoc Networks," in *Elsevier Journal on ad hoc Networks*, July 2005.

[21] M. Garetto, P. Giaccone, and E. Leonardi, "Capacity Scaling in Delay Tolerant Networks with Heterogeneous Mobile Nodes," in *Proceedings of ACM MOBIHOC*, May 2007.

[22] W. Ren, Q. Zhao, and A. Swami, "Connectivity of Cognitive Radio Networks: Proximity vs. Opportunity," in *Proceedings of ACM MOBICOM Workshop on Cognitive Radio Networks*, September 2009.

[23] ——, "On the Connectivity and Multihop Delay of Ad Hoc Cognitive Radio Networks," in *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, April 2011.

[24] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," in *IEEE Transactions on Information Theory*, vol. 46, no. 2, March 2000, pp. 388–404.

[25] T. Lindavll, *Lectures on the Coupling Method*. John Wiley & Sons, Inc.

[26] H. Cai and D. Y. Eun, "Crossing Over the Bounded Domain: From Exponential to Power-law Inter-meeting Time in MANET," in *Proceedings of ACM MOBICOM*, August 2007.

[27] G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed. Wadsworth Group, 2002.

[28] T. Liggett, "An Improved Subadditive Ergodic Theorem," in *The Annals of Probability*, vol. 13, no. 4, 1985, pp. 1279–1285.

[29] Z. Kong and E. M. Yeh, "A Distributed Energy Management Algorithm for Large-Scale Wireless Sensor Networks," in *Proceedings of ACM MOBIHOC*, May 2007.

[30] O. Dousse, M. Franceschetti, N. Macris, R. Meester, and P. Thiran, "Percolation in Signal to Interference Ratio Graph," in *Journal of Applied Probability*, vol. 43, no. 2, 2006.

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

[32] M. Penrose, *Random Geometric Graphs*. Oxford University Press, 2003.

Lei Sun received his Ph.D. degree in the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh NC, in 2012. His research interests include network performance, topology, mobility and resilience.

Wenye Wang received the M.S.E.E. degree and Ph.D. degree in computer engineering from the Georgia Institute of Technology, Atlanta, in 1999 and 2002, respectively. She is an Associate Professor with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh NC. Her research interests include mobile and secure computing, modeling and analysis of wireless networks, network topology, and architecture design. Dr. Wang has been a Member of the Association for Computing Machinery (ACM) since 1998, and a Member of the Eta Kappa Nu and Gamma Beta Phi honorary societies since 2001. She is a recipient of the NSF CAREER Award 2006. She is the co-recipient of the 2006 IEEE GLOBECOM Best Student Paper Award - Communication Networks and the 2004 IEEE Conference on Computer Communications and Networks (ICCCN) Best Student Paper Award.

Yujin Li received her B.S. and M.S. degrees in Control Science and Engineering from Beijing Institute of Technology in Beijing, China. She received her Ph.D. in Electrical Engineering from the Department of Electrical and Computer Engineering at North Carolina State University in 2014. Her research interests are in Device-to-Device (D2D) wireless communication networks.