This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

On the Connectivity Analysis over Large-Scale Hybrid Wireless Networks

Chi Yi and Wenye Wang
Department of Electrical and Computer Engineering
North Carolina State University

Abstract*—Many real systems are hybrid networks which include infrastructure nodes in multi-hop wireless networks, such as sinks in sensor networks and mesh routers in mesh networks. However, we have very little understanding of network connectivity in such networks. Therefore, in this paper, we consider hybrid networks denoted by $H(\alpha, \beta)$ with ad hoc nodes and base stations and prove how base stations can improve the connectivity of ad hoc nodes in *subcritical phase*, that is, the ad hoc node density, λ_α is lower than the critical density λ_α^c. We find that with the existence of a positive density of base stations, i.e., the density of base stations $\lambda_\beta > 0$ which have the same transmission range as ad hoc nodes, the number of connected ad hoc nodes is $\Theta(n)$ with probability nearly 1, where n is the number of ad hoc nodes. However, the size of connected ad hoc component scales linearly with λ_β when it is lower than $c_1(\lambda_\alpha)$ with probability nearly 1, which demonstrates a tremendous benefit of using base stations to enhance the connectivity of ad hoc nodes. Further, we study a hybrid network architecture that makes a significant connectivity improvement with transmission range r_β larger than r_α for ad hoc nodes. Therefore, our results provide a theoretical understanding of to what extent ad hoc nodes can benefit from base stations in multi-hop wireless networks.

I. Introduction

Connectivity is an essential issue in multi-hop wireless networks in that a wireless node must establish a communication path to other nodes in order to successfully transmit or receive information. To this end, there have been extensive studies on the connectivity of wireless ad hoc networks and sensor networks, in which a large number of wireless nodes, either mobile or static, communicate with each other in a multi-hop fashion [1]–[6]. For example, Gupta and Kumar in [1] provide the critical transmission range for asymptotic connectivity in a *dense network*, where n nodes are uniformly and independently placed in a unit disc. Further, a fundamental result in continuum percolation theory [7], [8] shows that in an *extended network*, where n nodes are randomly deployed in a square region by use of Poisson distribution with density λ_α, there exists a critical value, which is called *critical density* λ_α^c. When node density λ_α is above λ_α^c (*supercritical*), there exists a unique *giant component* containing $\Theta(n)$ nodes asymptotically almost surely (a.a.s.); and if λ_α falls below λ_α^c (*subcritical*), the network will be partitioned into a number of finite-size connected components almost surely (a.s.).

However, large-scale pure ad hoc networks rarely exist in real word, because the per-node throughput can diminish to zero [9]–[11]. The most widely used systems are in fact hybrid networks which include infrastructure nodes in multi-hop wireless networks, such as sinks in sensor networks and mesh routers in mesh networks. In recent years, there have been considerable studies of such networks [12]–[15], which have mainly focused on the *capacity* of hybrid networks. For example, Liu *et al.* in [12] investigate the capacity of hybrid networks, in which n ad hoc nodes are uniformly and independently located in a unit disk, but m base stations are placed on a regular grid within the area of the network. Their results show that the benefit of adding base stations on capacity is insignificant if m grows asymptotically slower than $\sqrt{n}$. But if m grows asymptotically faster than $\sqrt{n}$, the capacity increases linearly with the number of base stations, which provides an effective improvement over pure ad hoc networks.

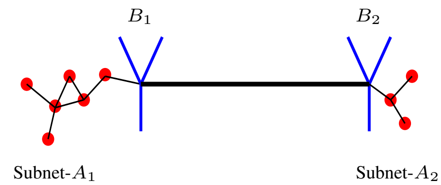


Fig. 1. A hybrid network with 2 base stations and ad hoc nodes.

One interesting problem, yet remains unsolved, is to what extent the base stations improve the network connectivity, even though it appears to be intuitive that base stations can surely improve network connectivity. As an example, Fig. 1 shows that two connected components of ad hoc nodes, Subnet-A_1 and Subnet-A_2 with size 6 and 3 respectively, are separate before two base stations, B_1 and B_2, are added in the network. Nevertheless, after B_1 and B_2 are added, Subnet-A_1 and Subnet-A_2 are connected together and thus form a larger connected component.

Therefore, in this paper, our objective is to study the connectivity in hybrid networks, and we focus specifically on the case that ad hoc node density, λ_α is lower than the critical density λ_α^c, under which the connectivity in pure ad hoc network is "bad". The challenge for our study is, in the existing literature, there are rarely analytical results on the connectivity in such *inhomogeneous* networks. Hence, our results can provide a fundamental understanding and analytical support of the connectivity in hybrid networks.

First, we study a pure ad hoc network which is in subcritical phase, and examine the upper bound of complementary cumulative distribution function (ccdf) of *origin stretch*, which is defined as the largest Euclidean distance between the origin and the nodes in a connected component containing the origin. Then we use a mapping approach called "*box connection mapping*" to obtain the upper bound of ccdf of origin stretch, and hereby find the value $c(\lambda_\alpha)$, by which origin stretch is bounded with probability nearly 1.

Then we proceed to study a hybrid network $H(\alpha, \beta)$ in which the transmission ranges of ad hoc (α-)nodes r_α is normalized to unit 1 and base stations (β-nodes) are multiples of this unit, denoted by r_β. We prove that when the Euclidean distance between adjacent base stations is larger than $c_2(\lambda_\alpha) = c(\lambda_\alpha)+1$ for $r_\beta = r_\alpha = 1$, α-nodes connected "directly" by any two base stations, do not intersect with probability nearly 1. Here, we say α-nodes are connected "directly" by a base station "directly", iff these α-nodes can be connected by the base station even if we remove all other base stations. For example, in Fig. 1, the α-nodes in Subnet-A_1 are connected "directly" by B_1, but the α-nodes in Subnet-A_2 are not connected "directly" by B_1. Based on this result, we find that with the existence of a positive density of base stations $\lambda_\beta > 0$, the number of connected ad hoc nodes is $\Theta(n)$ with probability nearly 1, where n is the number of ad hoc nodes. However, the size of connected ad hoc component scales linearly with λ_β that is lower than $c_1(\lambda_\alpha)$ with probability nearly 1, which demonstrates a tremendous benefit of using base stations to enhance the connectivity of ad hoc nodes.

Further, we study a hybrid network architecture which can utilize large $r_\beta (> 1)$ to obtain an additional benefit on the connectivity of ad hoc nodes. Our results indicate that by increasing r_β from unit to $k_0 c_2(\lambda_\alpha)$ ($k_0 \in \mathbb{N}$), with proper placement base stations and additional ad hoc nodes in the network, the connectivity of ad hoc nodes can be significantly improved, which provide a guideline on design of a new hybrid network.

The rest of the paper is organized as follows. In Section II, we introduce network models and formulate the problem. In Section III, we provide the value $c(\lambda_\alpha)$ that origin stretch is lower than with probability nearly 1 in pure ad hoc networks. In Section IV, we present theoretical proof of the connectivity improvement due to base stations, followed by the conclusions in Section V.

II. Network Models and Problem Statement

A. Hybrid Network Model

There are two types of nodes in the hybrid network $H(\alpha, \beta)$. The first set consists of ad hoc nodes which are called α-nodes. A set of n α-nodes $\mathcal{A} = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ are randomly deployed at locations $A = \{a_1, a_2, \ldots, a_n\}$ in a square region D with side length d, by use of Poisson distribution with density λ_α. This square region D is also called a continuum box [8]. Every α-node has the same transmission power and signal receiving capability, and thus has the same transmission range denoted as r_α. That is, for any two α-nodes $\alpha_i, \alpha_j \in \mathcal{A}$, there is an undirected link between them iff $\|a_i - a_j\| < r_\alpha$, where $\|\cdot\|$ denotes the Euclidean distance. Without loss of generality, we normalize r_α to unit throughout the paper. Then we denote *random geometric graph* (RGG) $G(\mathcal{H}_{\lambda_\alpha}, 1)$ as the ensemble of all α-nodes and the links between them. It is well known from continuum percolation theory [7], [8], that there exists a critical density $0 < \lambda_\alpha^c < \infty$ in $G(\mathcal{H}_{\lambda_\alpha}, 1)$ such that: (1)for $\lambda_\alpha > \lambda_\alpha^c$, there exists a unique connected component containing $\Theta(n)$ nodes a.a.s. (supercritical phase); and (2) for $\lambda_\alpha < \lambda_\alpha^c$, the connected components contain a finite number of nodes a.s. (subcritical phase). Until recently, the exact value of λ_α^c is still unknown although some bounds have been given in the literatures. The analytical bound $0.696 < \lambda_\alpha^c < 3.372$ is given in [7], while the simulation bound $1.434 < \lambda_\alpha^c < 1.438$ is given in [16]. In this paper, we assume the density of α-nodes, λ_α, is lower than λ_α^c.

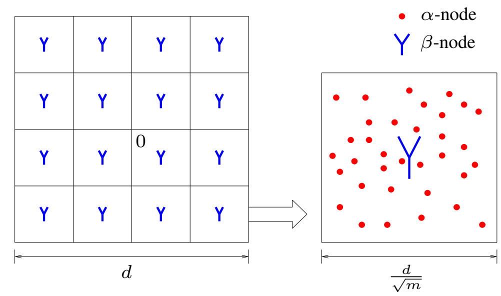


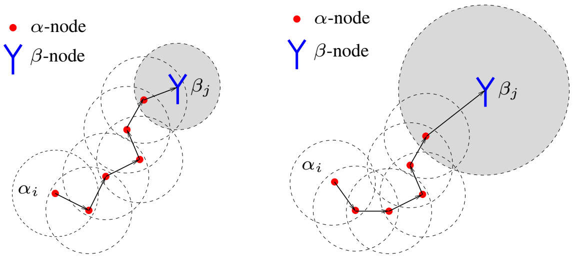
Fig. 2. The layout of a hybrid network.

The second set consists of infrastructure nodes which are called β-nodes. A set of m β-nodes $\mathcal{B} = \{\beta_1, \beta_2, \ldots, \beta_m\}$ (with density $\lambda_\beta = \frac{m}{d^2}$) are placed regularly at the locations $B = \{b_1, b_2, \ldots, b_m\}$ $([-\frac{d}{2} + \frac{d}{2\sqrt{m}} + \frac{d}{\sqrt{m}} i, -\frac{d}{2} + \frac{d}{2\sqrt{m}} + \frac{d}{\sqrt{m}} j], 0 \le i, j \le \sqrt{m} - 1)$ in the region D as depicted in Fig. 2. The β-nodes are connected by high bandwidth wirelines and act only as relay nodes for routing and forwarding the packets from and to the α-nodes. We assume that β-nodes, like base stations, are much more powerful than α-nodes that are power-constrained [9] and β-nodes have a transmission range of $r_\beta \ge r_\alpha = 1$). That is, for any pair of α-node $\alpha_i \in \mathcal{A}$ and β-node $\beta_j \in \mathcal{B}$, there is an undirected link between them iff $\|a_i - b_j\| < r_\beta$. Figs. 3(a) and 3(b) demonstrate an example of data flow from an α-node α_i to a β-node β_j in the case $r_\beta = 1$ and $r_\beta = 2$. In both cases, there exists one α-node, which is in the transmission range of β_j, and can communicate with α_i in a multi-hop fashion. Therefore, the data sent from α_i can reach β_j through the relay of α-nodes. As a result, the hybrid network $H(\alpha, \beta)$, is defined as the ensemble of all α-nodes, β-nodes, and links (wireline and wireless) between them.

B. Definitions

Since continuum percolation theory studies the connectivity properties in *homogeneous* networks, there is no definition in the theory for hybrid networks. Therefore, we define *β-component* and *β-giant component* in hybrid network $H(\alpha, \beta)$ as follows:

Definition 1 (β-Component): In hybrid network $H(\alpha, \beta)$, the *β-component* denoted by $\mathcal{C}_{(\alpha,\beta)}$ is defined as the connected component containing all β-nodes, and a subset of α-nodes



(a) $r_\beta = 1$ (b) $r_\beta = 2$

Fig. 3. Data flow from an α-node to a β-node.

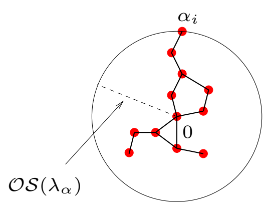


Fig. 4. Origin stretch in pure ad hoc network.

that are connected by the β-nodes. Note that, we assume all β-nodes are connected via broadband wireline, that is, if a connected component contains one β-node, it will contain all β-nodes, of which the number are constant. Thus, we define *β-component size* $|\mathcal{C}_{(\alpha,\beta)}|$ as the number of α-nodes in $\mathcal{C}_{(\alpha,\beta)}$. For example, as depicted in Fig. 1, there are two β-nodes in hybrid network, connecting two separate subsets together to form a larger connected component. The connected component including B_1, B_2, Subset-A_1 and Subset-A_2 is the β-component.

Definition 2 (β-Giant Component): In hybrid network $H(\alpha, \beta)$, the *β-giant component*, is the β-component with size $|\mathcal{C}_{(\alpha,\beta)}| = \Theta(n)$.

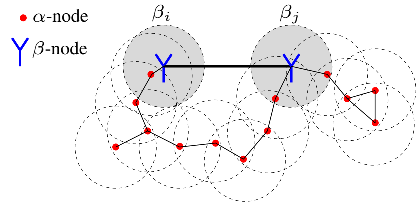
Next, we define a new metric *origin stretch* in pure ad hoc networks.

Definition 3 (Origin Stretch): Let $G(\mathcal{H}_{\lambda_\alpha,0}, 1) = G(\mathcal{H}_{\lambda_\alpha}, 1) \cup \{0\}$, and $\mathcal{C}_0$ be the connected component in $G(\mathcal{H}_{\lambda_\alpha,0}, 1)$ containing the origin 0. The *origin stretch* of $G(\mathcal{H}_{\lambda_\alpha,0}, 1)$ is defined as $\mathcal{OS}(\lambda_\alpha) = \max_{\alpha_i \in \mathcal{C}_0} \|\alpha_i - 0\|$, where $\|\alpha_i - 0\|$ is the Euclidean distance between α_i and the origin 0. For example, as depicted in Fig. 4, origin stretch $\mathcal{OS}(\lambda_\alpha)$ is the Euclidean distance between α_i and the origin 0.

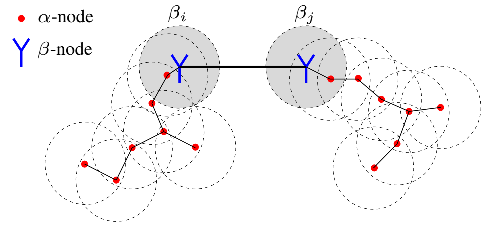
C. Problems and Main Results

1) Problem Formulation: Our objective in this paper, is to study how β-node can improve the connectivity of α-nodes, and investigate what are steps. Now, we formulate the problems addressed in this paper.

- First, we start from a special case, that is, α-nodes only and find the upper bound of origin stretch for any given α-node density, to understand the coverage of α-nodes.
- Then, we study how to use β-nodes efficiently. From previous discussions, we have observed that β-nodes can improve network connectivity. An interesting question is that how many β-nodes need to be added to a network. Similar questions, like the placement of base stations in sensor networks have been studied in [17], [18]. While this question is beyond the scope of this study, it is relevant to the evaluation of *β-nodes* utilization in network connectivity. For example, in Fig. 5(a), two β-nodes, β_i and β_j, are connected through the relay of a number of α-nodes. Let I be the number of α-nodes connected "directly" by β_i. Here, as we explain in Section I, α-nodes are connected "directly" by a β-node, iff these α-nodes can be connected by the β-node even if we remove all other β-nodes. Similarly, let J be the number of α-nodes connected "directly" by β_j. For the particular case shown in the figure, we can easily find that $I = 9$ and $J = 13$. However, the total number of α-nodes, which are connected "directly" by β_i and β_j, is $13 < I + J$. In Fig. 5(b), β_i and β_j cannot be connected through the relay of α-nodes. Then the total number of α-nodes, which are connected "directly" by β_i and β_j, is $14 = I + J$. Apparently, in the later case, the two β-nodes are better used. Thus, in this case, we call the two β-nodes are sufficiently used. Further, we define all β-nodes in hybrid network $H(\alpha, \beta)$ are sufficiently used, iff any two β-nodes cannot be connected by α-nodes.



(a) Insufficient use of β-nodes



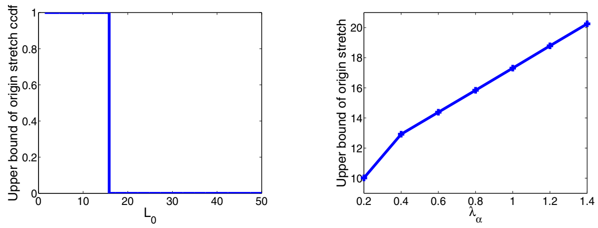
(b) Sufficient use of β-nodes

Fig. 5. Insufficient use vs. sufficient use of β-nodes ($r_\beta = 1$).

- Third, we evaluate the impact of β-nodes on the connectivity in a hybrid network. This problem is two-fold: the effect of node density λ_β and the effect of transmission range r_β on network connectivity, that is, the conditions for the existence of β-giant component in a hybrid network $H(\alpha, \beta)$.

2) Main Results: In order to answer the first question, our approach is to find the upper bound of origin stretch, that is, the value that origin stretch is lower than with probability nearly 1. Then, we use "*box connection mapping*" to investigate the upper bound of complementary cumulative distribution function (ccdf) of origin stretch in the following theorem. Based on this result, we will be able to find the answer regarding the first question.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.



(a) Upper bound of the ccdf of $\mathcal{OS}(\lambda_\alpha)$ for $\lambda_\alpha = 0.8$

(b) Upper bound of origin stretch for different λ_α

Fig. 6. Upper bound of origin stretch.

Theorem 1: For a randomly given value $L_0 \in (\sqrt{2}, \infty)$, the probability that the origin stretch $\mathcal{OS}(\lambda_\alpha)$ is larger than L_0 will be upper bounded by the following inequality

$$\begin{aligned} &\text{Prob}(\mathcal{OS}(\lambda_\alpha) > L_0) \\ &< 1 - (1 - (1 - e^{-\lambda_\alpha (2i_0+2) d_l^2})^{j_0})^{4 \cdot 3^{(2i_0+2)j_0 - 1}}. \end{aligned} \quad (1)$$

in which j_0 and i_0 can be obtained from Eq. (14) and (15), and the side length of the discrete lattice d_l satisfies

$$d_l = \frac{L_0}{\sqrt{((2i_0+2)j_0 + \frac{1}{2})^2 + (\frac{1}{2})^2}}. \quad (2)$$

Remark 1: The above theorem characterizes the upper bound of the ccdf of origin stretch. We plot the numerical results in above theorem and show it in Fig. 6(a). In the numerical computation, we set $\lambda_\alpha = 0.8$, which indicates $\lambda_\alpha < \lambda_\alpha^c$. From the figure, we know the probability that origin stretch $\mathcal{OS}(\lambda_\alpha)$ is larger than 15.84, is approximately equal to 0. That means, in the case $\lambda_\alpha = 0.8$, with probability nearly 1, origin stretch $\mathcal{OS}(\lambda_\alpha)$ is less or equal to 15.84. Then, from our definition, 15.84 is the upper bound of origin stretch $\mathcal{OS}(\lambda_\alpha)$ for $\lambda_\alpha = 0.8$. In Fig. 6(b), we show the upper bound of origin stretch for different α-node density. From the figure, we find that when $G(\mathcal{H}_{\lambda_\alpha,0}, 1)$ is in subcritical phase, the upper bound of origin stretch almost scales linearly with the α-node density.

For the second question, we first denote $U_{\mathcal{OS}}(\lambda_\alpha)$ as the upper bound of origin stretch $\mathcal{OS}(\lambda_\alpha)$, which can be found by using Theorem 1. Then we have the following result.

Theorem 2: For unit transmission range of β-nodes, i.e., $r_b = 1$, if the β-node density satisfies $\lambda_\beta < c_1(\lambda_\alpha)$, all m β-nodes in hybrid network $H(\alpha, \beta)$ are sufficiently used with probability nearly 1, where $c_1(\lambda_\alpha) = \frac{1}{c_2^2(\lambda_\alpha)}$ and $c_2(\lambda_\alpha) = U_{\mathcal{OS}}(\lambda_\alpha) + 1$.

Remark 2: The above theorem characterizes a sufficient condition that all β-nodes in hybrid network $H(\alpha, \beta)$ are sufficiently used, with probability nearly 1. From the results in Fig. 6(b), we know that for $\lambda_\alpha = 1.4$ the sufficient condition is $\lambda_\beta < 2.21 \times 10^{-3}$, while for $\lambda_\alpha = 0.8$, the sufficient condition becomes $\lambda_\beta < 3.53 \times 10^{-3}$.

As we discuss before, if all β-nodes in hybrid network $H(\alpha, \beta)$ are sufficiently used, any two β-nodes in the network cannot be connected through the relay of α-nodes. Then, β-component size $|\mathcal{C}_{(\alpha,\beta)}|$ is the summation of the number of α-nodes, which are connected "directly" by each β-node. Therefore, from Theorem 2, for unit transmission range of β-nodes, if the β-node density satisfies $\lambda_\beta < c_1(\lambda_\alpha)$, we can compute β-component size $|\mathcal{C}_{(\alpha,\beta)}|$, by summing up the number of α-nodes connected "directly" by each β-node. And then, we prove if $0 < \lambda_\beta < c_1(\lambda_\alpha)$, β-component size scales linearly with λ_β, as we show in Theorem 3. Our results provide a fundamental understanding and analytical support of the connectivity in hybrid networks.

In addition, we investigate the effect of transmission range r_β on the connectivity of α-nodes in hybrid network $H(\alpha, \beta)$. Specifically, we propose a hybrid network architecture, and study the connectivity improvement in this network. Our results show that through increasing r_β from unit to $k_0 c_2(\lambda_\alpha)$ ($k_0 \in \mathbb{N}$), additional benefit on the connectivity of α-nodes can be obtained. Hence, our results provide a guideline on design of a new hybrid network, which contains low density of ad hoc nodes.

III. Upper Bound of Origin Stretch $\mathcal{OS}(\lambda_\alpha)$

As we discuss in Section II-C1, if any two β-nodes in hybrid network $H(\alpha, \beta)$ cannot be connected through the relay of α-nodes, all β-nodes are sufficiently used. To find out the sufficient condition that all β-nodes are sufficiently used, we will first explore the upper bound of origin stretch $\mathcal{OS}(\lambda_\alpha)$ in $G(\mathcal{H}_{\lambda_\alpha,0}, 1)$, to understand the coverage of α-nodes. Here, the upper bound is from the probabilistic perspective, i.e., the value that $\mathcal{OS}(\lambda_\alpha)$ is lower than with probability nearly 1. In order to find the upper bound of origin stretch $\mathcal{OS}(\lambda_\alpha)$, in this section, we will investigate the upper bound of the complementary cumulative distribution function of $\mathcal{OS}(\lambda_\alpha)$, that is, if the probability that $\mathcal{OS}(\lambda_\alpha)$ is larger than S, is nearly 0, we know that $\mathcal{OS}(\lambda_\alpha)$ is bounded by S with probability nearly 1.

Because of the randomness of the node positions in $G(\mathcal{H}_{\lambda_\alpha,0}, 1)$, the shape of the connected component containing the origin 0 is quite difficult to describe. Thus, it is hard to derive the upper bound of the complementary cumulative distribution function of $\mathcal{OS}(\lambda_\alpha)$ directly in the continuous two-dimensional plane. Our strategy here, is to first map the pure ad hoc network onto a discrete lattice.

A. Box-Connection Mapping

In this section, we introduce our mapping approach called "*box-connection mapping*". First, we consider a discrete lattice $\mathcal{L}$ with side length d_l. The coordinates of the vertices of $\mathcal{L}$ are $(d_l \times i, d_l \times j)$ for $(i, j) \in \mathbb{Z}^2$. Then, we construct its dual lattice $\mathcal{L}'$ by moving $\mathcal{L}$ horizontally and vertically by $\frac{d_l}{2}$, as depicted in Fig. 7.

Let S_a be the square box centered at vertex a of $\mathcal{L}$. Now, we consider a randomly chosen path in $\mathcal{L}$ with length Ld_l, starting from the origin 0 through $v_1, v_2, \ldots, v_{L-1}$ to v_L. Then we define the event E_{opath} that this path is open iff the following two conditions hold.

1) E_A: There is at least one α-node in S_0 and one node in S_{v_L};

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

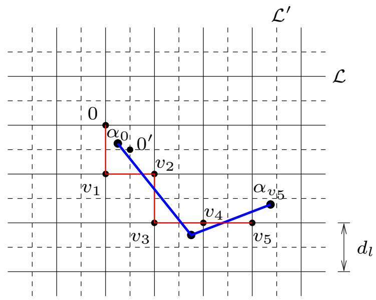


Fig. 7. The lattice $\mathcal{L}$ (real), and its dual $\mathcal{L}'$ (dashed).

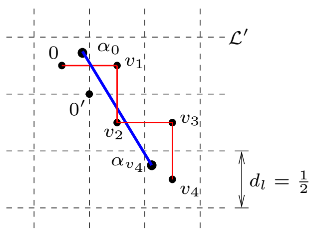


Fig. 8. A single link can go across 5 square boxes at most for $d_l = \frac{1}{2}$.

2) E_B: Among all α-nodes in S_0 and S_{v_L}, there exists at least one pair of α-nodes, one in S_0 and the other in S_{v_L}, that can be connected by several links. And, these links are all contained in $S_0, S_{v_1}, \ldots, S_{v_L}$ and go through all these square boxes.

Otherwise, this path is defined to be closed. Fig. 7 gives an example for an open path in $\mathcal{L}$ with length $5d_l$. The path begins from 0 to v_5. There is one α-node in S_0, α_0, and another in S_{v_5}, α_{v_5}. The two links connecting α_0 and α_{v_5} are all contained in $S_0, S_{v_1}, \ldots, S_{v_5}$ and go through all these 6 square boxes. Thus, from the above definition, the path is open.

B. Upper Bound of Origin Stretch $\mathcal{OS}(\lambda_\alpha)$

In this section, we study how to use "box-connection mapping" to derive the upper bound of origin stretch $\mathcal{OS}(\lambda_\alpha)$. First, we will study the relation between the largest number of square boxes that a single link can go across and the box side length d_l. Here, we only consider $d_l < 1$, to make sure a single link can go across multiple square boxes. As an example for illustration, we take $d_l = \frac{1}{2}$ as depicted in Fig. 8. From the figure, it is not difficult to find out a single link can go across at most 5 square boxes. Then, we may ask, for any given d_l, what is the largest number of square boxes that a single link can go across? Based on this question, we give the following lemma.

Lemma 1: If d_l satisfies

$$\frac{1}{\sqrt{(i+1)^2 + i^2}} < d_l < \frac{1}{\sqrt{i^2 + i^2}}, \qquad i \in \mathbb{N}. \tag{3}$$

a single link can go across at most $(2i+3)$ square boxes.

Proof: We first define the coordinates of square boxes in $\mathcal{L}'$. The square box containing the origin 0 is assigned coordinate $\overline{(0,0)}$. Then the box at the right side of box $\overline{(0,0)}$ is assigned coordinate $\overline{(1,0)}$, and the box at the right side of box $\overline{(1,0)}$ is assigned coordinate $\overline{(2,0)}$, and so on. Similarly, the square box below box $\overline{(0,0)}$ is assigned coordinate $\overline{(0,1)}$, and the box below box $\overline{(0,1)}$ is assigned coordinate $\overline{(0,2)}$, and so on. From this definition, we know if a single link reaches box $\overline{(m,n)}$ (without loss of generality, we assume $m \geq n$), it can go across $(m+n+1)$ square boxes. Thus, the problem of finding the largest number of square boxes a single link can go across, is switched to the problem of finding the maximum of $(m+n+1)$ under the constraint that box (m,n) has non-empty intersection with the disk centered at $0'$ with radius 1. That is, this problem can be described as the following constrained optimization problem.

$$\max_{i,j \in \mathbb{N}} \quad m+n+1 \tag{4}$$

$$\text{subject to} \quad \sqrt{((m-1)d_l)^2 + ((n-1)d_l)^2} < 1 \tag{5}$$

If the condition (3) holds, the above optimization problem can be solved as: when $m = i+1, n = i+1$ or $m = i+2, n = i$, then $m+n+1$ reaches its maximum value $2i+3$. □

Lemma 1 provides the largest number of square boxes that a single link can go across, if the side length d_l satisfies Eq. (3). Based on this result, we have the following lemma.

Lemma 2: If the side length of each square box d_l satisfies Eq. (3), the probability $P_{(2i+2)j}$ that there exists an open path in $\mathcal{L}$ with length $(2i+2)jd_l$ $(i, j \in \mathbb{N})$ will satisfy

$$P_{(2i+2)j} < 1 - (1 - (1 - e^{-\lambda_\alpha (2i+2) d_l^2})^j)^{4 \cdot 3^{(2i+2)j-1}}. \tag{6}$$

Proof: For a randomly chosen open path in $\mathcal{L}$ with length $(2i+2)jd_l$, from Lemma 1, there exists at least one α-node in every $(2i+2)$ consecutive square boxes. Let P' be the probability that there exists at least one α-node in $(2i+2)$ consecutive square boxes, and P'' be the probability that a randomly chosen path in $\mathcal{L}$ with length $(2i+2)jd_l$ is open. Then, we have

$$P'' < (P')^j. \tag{7}$$

We denote $\theta(k)$ $(k \in \mathbb{N})$ as the number of paths starting from 0 with length kd_l. Further, let $P_{(2i+2)j}$ be the probability that there exists an open path in $\mathcal{L}$ with length $(2i+2)jd_l$, and $\overline{P_{(2i+2)j}}$ be the probability that any open path in $\mathcal{L}$ with length $(2i+2)jd_l$ is closed. Using FKG's inequality [7], [19], we know $\overline{P_{(2i+2)j}}$ will satisfy the following condition.

$$\overline{P_{(2i+2)j}} \geq (1 - P'')^{\theta((2i+2)j)}. \tag{8}$$

Based on the relation $P_{(2i+2)j} = 1 - \overline{P_{(2i+2)j}}$, we have

$$P_{(2i+2)j} \leq 1 - (1 - P'')^{\theta((2i+2)j)}. \tag{9}$$

We know $\theta(k) < 4 \cdot 3^{k-1}$ from many literatures (e.g., [7]). And, since the α-nodes obey Poisson point process, we can easily know $P' = 1 - e^{-\lambda_\alpha (2i+2) d_l^2}$. Hence, Eq. (9) further satisfy

$$P_{(2i+2)j} < 1 - (1 - (1 - e^{-\lambda_\alpha (2i+2) d_l^2})^j)^{4 \cdot 3^{(2i+2)j-1}}. \tag{10}$$

□

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

Lemma 2 provides the upper bound of the probability of existence of open path with a given length. From this result, we can derive the upper bound of the ccdf of graph stretch $\mathcal{OS}(\lambda_\alpha)$ by the following theorem.

Lemma 3: If the side length of lattice $\mathcal{L}$ d_l satisfies Eq. (3), the probability that $\mathcal{OS}(\lambda_\alpha)$ is larger than $\sqrt{\left((2i+2)j+\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^2}d_l$ $(i,j\in\mathbb{N})$ will satisfy

$$\begin{aligned}&\text{Prob}\left(\mathcal{OS}(\lambda_\alpha)>\sqrt{((2i+2)j+\frac{1}{2})^2+(\frac{1}{2})^2}d_l\right)\\&\quad<1-(1-(1-e^{-\lambda_\alpha(2i+2)d_l^2})^{j+1})^{4\cdot3^{(2i+2)j-1}}.\end{aligned}\tag{11}$$

Proof: We assume the length of the open path in $\mathcal{L}$ is $(2i+2)jd_l$. If the coordinate of the square box containing the end point of the open path is (m',n') (without loss of generality, we assume $m'\geq n'$). It is obvious that the largest Euclidean distance between the origin 0 and the possible α-node in this box will be less or equal to $\sqrt{(m'+\frac{1}{2})^2+(n'+\frac{1}{2})^2}d_l$. Then we maximize $\sqrt{(m'+\frac{1}{2})^2+(n'+\frac{1}{2})^2}$ for $m',n'\in\mathbb{N}$ under the constraint $m'+n'\leq(2i+2)j$, and solve problem as: when $m'=(2i+2)j$ and $n'=0$, $\sqrt{(m'+\frac{1}{2})^2+(n'+\frac{1}{2})^2}$ reaches its maximum value $\sqrt{\left((2i+2)j+\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^2}$.

Subsequently, we define two events as follows.

1) E'_A: There exists an open path in $\mathcal{L}$ with length $(2i+2)jd_l$;
2) E'_B: The origin stretch $\mathcal{OS}(\lambda_\alpha)$ is larger than $\sqrt{((2i+2)j+\frac{1}{2})^2+(\frac{1}{2})^2}d_l$.

Using *Law of Total Probability*, we have $\text{Prob}(E'_A)=\text{Prob}(E'_A|E'_B)\cdot\text{Prob}(E'_B)+\text{Prob}(E'_A|\overline{E'_B})\cdot\text{Prob}(\overline{E'_B})\geq\text{Prob}(E'_A|E'_B)\cdot\text{Prob}(E'_B)$. Since when $\mathcal{OS}(\lambda_\alpha)$ is larger than $\sqrt{((2i+2)j+\frac{1}{2})^2+(\frac{1}{2})^2}d_l$, there must exist an open path in $\mathcal{L}$ with length $(2i+2)jd_l$, we have $\text{Prob}(E'_A|E'_B)=1$, and thus further obtain $\text{Prob}(E'_A)\geq\text{Prob}(E'_B)$. From Lemma 2, we know $\text{Prob}(E'_A)<1-(1-(1-e^{-\lambda_\alpha(2i+2)d_l^2})^{j+1})^{4\cdot3^{(2i+2)j-1}}$. Then, we prove the lemma. □

To simplify the long expression, we introduce a new term $c(i,j)$ denoted as $c(i,j)=\sqrt{\left((2i+2)j+\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^2}$. From Lemma 3, for a given d_l satisfying Eq. (3), we can only find the upper bound of the probability that $\mathcal{OS}(\lambda_\alpha)$ is larger than the discrete values $c(i,j)d_l$ $(i,j\in\mathbb{N})$. Then, we will ask, for any given value L_0, what is the upper bound of $\text{Prob}(\mathcal{OS}(\lambda_\alpha)>L_0)$? Then, we will investigate if we can find appropriate d_l, i_0 and j_0, such that d_l satisfies Eq. (3) and $L_0=c(i_0,j_0)d_l$. If this is feasible, we can directly use Eq. (11) to obtain the upper bound of $\text{Prob}(\mathcal{OS}(\lambda_\alpha)>L_0)$.

Notice that when d_l satisfies Eq. (3), the following condition holds: $\frac{c(i,j)}{\sqrt{(i+1)^2+i^2}}<c(i,j)d_l<\frac{c(i,j)}{\sqrt{i^2+i^2}}$. Then, we have the following lemma.

Lemma 4: The union of all such intervals $(\frac{c(i,j)}{\sqrt{(i+1)^2+i^2}},\frac{c(i,j)}{\sqrt{i^2+i^2}})$ $(i,j\in\mathbb{N})$ is $(\sqrt{2},\infty)$, i.e., $\bigcup_{i,j\in\mathbb{N}}(\frac{c(i,j)}{\sqrt{(i+1)^2+i^2}},\frac{c(i,j)}{\sqrt{i^2+i^2}})=(\sqrt{2},\infty)$.

Proof: To prove the lemma, we first let j fixed and investigate $\bigcup_{i\in\mathbb{N}}(\frac{c(i,j)}{\sqrt{(i+1)^2+i^2}},\frac{c(i,j)}{\sqrt{i^2+i^2}})$. To find this result, we then consider the union of two following consecutive intervals: $(\frac{c(i,j)}{\sqrt{(i+1)^2+i^2}},\frac{c(i,j)}{\sqrt{i^2+i^2}})$ and $(\frac{c(i+1,j)}{\sqrt{(i+2)^2+(i+1)^2}},\frac{c(i+1,j)}{\sqrt{(i+1)^2+(i+1)^2}})$.

It is easy to prove the union of above two consecutive intervals is $(\frac{c(i+1,j)}{\sqrt{(i+2)^2+(i+1)^2}},\frac{c(i,j)}{\sqrt{i^2+i^2}})$. From this result, we know

$$\bigcup_{i\in\mathbb{N}}(\frac{c(i,j)}{\sqrt{(i+1)^2+i^2}},\frac{c(i,j)}{\sqrt{i^2+i^2}})=(\sqrt{2}j,\sqrt{8j^2+2j+\frac{1}{4}}).\tag{12}$$

Then, we consider the union of two following consecutive intervals: $(\sqrt{2}j,\sqrt{8j^2+2j+\frac{1}{4}})$ and $(\sqrt{2}(j+1),\sqrt{8(j+1)^2+2(j+1)+\frac{1}{4}})$. Following the similar steps as above, we can prove

$$\bigcup_{j\in\mathbb{N}}\left(\sqrt{2}j,\sqrt{8j^2+2j+\frac{1}{4}}\right)=\left(\sqrt{2},\infty\right).\tag{13}$$

□

Lemma 4 validates that for any value $L_0\in(\sqrt{2},\infty)$, we can find appropriate d_l, i_0 and j_0, such that d_l satisfies Eq. (3) and $L_0=c(i_0,j_0)d_l$. In order to find the upper bound of the ccdf of origin stretch $\mathcal{OS}(\lambda_\alpha)$, we will investigate how to find the appropriate d_l, i_0 and j_0 to satisfy the above condition.

Lemma 5: For any given value $L_0\in(\sqrt{2},\infty)$, we can find the appropriate j_0, i_0 and d_l by the following equations, such that d_l satisfies Eq. (3) and $L_0=c(i_0,j_0)d_l$.

$$j_0=\text{ceil}(\frac{L_0}{\sqrt{2}})-1.\tag{14}$$

where ceil(X) rounds X towards positive infinity;

$$i_0=\text{floor}(\frac{-L_0^2+4j_0^2+j_0+\sqrt{\Delta}}{2L_0^2-4j_0^2})+1.\tag{15}$$

where $\Delta=-L_0^4+(4j_0^2+2j_0+1)L_0^2-j_0^2$ and floor(X) rounds X towards negative infinity; and

$$d_l=\frac{L_0}{c(i_0,j_0)}.\tag{16}$$

Proof: We start from finding the appropriate j_0. From the proof of Lemma 4, we know for any given value $L_0\in(\sqrt{2},\infty)$, there must exist $\widetilde{j}\in\mathbb{N}$ such that $L_0\in(\sqrt{2}\widetilde{j},\sqrt{8\widetilde{j}^2+2\widetilde{j}+\frac{1}{4}})$. Then, it is not difficult to find that $j_0=\text{ceil}(\frac{L_0}{\sqrt{2}})-1$ satisfies this condition.

Then we will locate the appropriate i_0 such that

$$\frac{c(i_0,j_0)}{\sqrt{(i_0+1)^2+i_0^2}}<L_0<\frac{c(i_0,j_0)}{\sqrt{i_0^2+i_0^2}}.\tag{17}$$

From Eq. (17), we can find Eq. (15) is a solution of i_0. Apparently, for the j_0 and i_0 obtained from Eq. (14) and (15), $d_l=\frac{L_0}{c(i_0,j_0)}$ will satisfy Eq. (3). □

Lemma 5 provides an approach to find appropriate appropriate d_l, i_0 and j_0, such that d_l satisfies Eq. (3) and

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

$L_0 = c(i_0, j_0)d_l$. After combining the results in Lemma 3, 4 and 5, we obtain a main result of this section illustrated in Theorem 1.

IV. How Do β-Nodes Improve the Connectivity of α-nodes?

In this section, we will study how β-nodes improve the connectivity of α-nodes. First, we consider unity transmission range of β-nodes, i.e., $r_\beta = 1$. From the results in Section III, we investigate the sufficient condition that all β-nodes in hybrid network $H(\alpha, \beta)$ are sufficiently used with probability nearly 1. Then, under the condition that all β-nodes in hybrid network $H(\alpha, \beta)$ are sufficiently used, β-component size is the summation of the number of α-nodes, which are connected by each β-node "directly". Based on this study, we provide the impact of β-nodes density on the connectivity of α-nodes, that is, how β-component size scales with the β-node density λ_β. Further, we will explore how β-nodes transmission range, r_β affect the connectivity of α-nodes. Specifically, we study a hybrid network architecture which can utilize large r_β to obtain additional benefit on the connectivity of α-nodes.

A. Sufficient Conditions for Sufficient Use of β-Nodes

As we defined in Section II-C1, all β-nodes in hybrid network $H(\alpha, \beta)$ are sufficiently used iff any two β-nodes cannot be connected through the relay of β-nodes. Here, we consider unity β-nodes transmission range and derive the sufficient conditions for sufficient use of β-nodes.

Proof of Theorem 2: As we define before, $U_{\mathcal{OS}}(\lambda_\alpha)$ is the upper bound of origin stretch $\mathcal{OS}(\lambda_\alpha)$. From Section III, we know the event E_{bound} that α-nodes connected by a β-node β_i "directly", is bounded within the disk centered at β_i with radius $U_{\mathcal{OS}}(\lambda_\alpha)$ happens with probability nearly 1. Then, under the condition that event E_{bound} happens, if the Euclidean distance between the adjacent β-nodes is larger than $U_{\mathcal{OS}}(\lambda_\alpha) + 1$, all β-nodes in hybrid network $H(\alpha, \beta)$ are sufficiently used. Otherwise, if two adjacent β-nodes β_i and β_j can be connected through the relay of α-nodes, there must exists one α-node α_s that is within the unit disk centered at β_j and can be connected by β_i through the relay of α-nodes. This is equivalent to say, α_s is not within the disk centered at β_i with radius $U_{\mathcal{OS}}(\lambda_\alpha)$, and thus contradicts with the condition that event E_{bound} happens. Then from $\frac{d}{\sqrt{m}} > U_{\mathcal{OS}}(\lambda_\alpha) + 1$, we know $\lambda_\beta = \frac{m}{d^2} < \frac{1}{(U_{\mathcal{OS}}(\lambda_\alpha)+1)^2}$. □

B. Impact of β-Nodes Density, λ_β

In this part, we consider the case that all β-nodes in hybrid network $H(\alpha, \beta)$ are sufficiently used. If we ignore the wireline connection between the β-nodes, each connected component containing a β-node is an independent connected component, which can be viewed as $\mathcal{C}_0$ in $G(\mathcal{H}_{\lambda_\alpha,0}, 1)$. Then in hybrid network $H(\alpha, \beta)$, β-component size is the summation of the number of α-nodes connected by each β-node "directly". In other word, let k_i $(i \in \{1, 2, \cdots, m\})$ be the number of α-nodes that are connected by β_i "directly". Then in hybrid network $H(\alpha, \beta)$, β-component size is equal to $|\mathcal{C}_{(\alpha,\beta)}| = \Sigma_{i=1}^m k_i$. As the β-node density λ_β increases, β-component size will become larger. Then, how does $|\mathcal{C}_{(\alpha,\beta)}|$ scale with λ_β? To answer this question, we first give the following lemma.

Lemma 6: Let $|\mathcal{C}_0|$ be the number of α-nodes in $\mathcal{C}_0$ of $G(\mathcal{H}_{\lambda_\alpha,0}, 1)$. When $G(\mathcal{H}_{\lambda_\alpha,0}, 1)$ is in subcritical phase, the mean of $|\mathcal{C}_0|$: $E(|\mathcal{C}_0|)$, and the variance of $|\mathcal{C}_0|$: $\text{Var}(|\mathcal{C}_0|)$, are both finite.

Proof: From Theorem 9.7 in [8], when $G(\mathcal{H}_{\lambda_\alpha,0}, 1)$ is in subcritical phase, there exist constants $\mu > 0$, $g_0 > 0$ such that $\text{Prob}(|\mathcal{C}_0| \geq g) \leq e^{-\mu g}$ when $g \geq g_0$. Hence, we have $E(|\mathcal{C}_0|) = \Sigma_{g=1}^\infty g \cdot \text{Prob}(|\mathcal{C}_0| = g) = \Sigma_{g=1}^{g_0} g \cdot \text{Prob}(|\mathcal{C}_0| = g) + \Sigma_{g=g_0}^\infty g \cdot \text{Prob}(|\mathcal{C}_0| = g) \leq \Sigma_{g=1}^{g_0} g \cdot \text{Prob}(|\mathcal{C}_0| = g) + \Sigma_{g=g_0}^\infty g \cdot e^{-\mu g}$, and $E(|\mathcal{C}_0|^2) = \Sigma_{g=1}^\infty g^2 \cdot \text{Prob}(|\mathcal{C}_0| = g) \leq \Sigma_{g=1}^{g_0} g^2 \cdot \text{Prob}(|\mathcal{C}_0| = g) + \Sigma_{g=g_0}^\infty g^2 \cdot e^{-\mu g}$. Using the ratio test criteria, we know both $\Sigma_{g=g_0}^\infty g \cdot e^{-\mu g}$ and $\Sigma_{g=g_0}^\infty g^2 \cdot e^{-\mu g}$ converge. Thus, both $E(|\mathcal{C}_0|)$ and $E(|\mathcal{C}_0|^2)$ are finite, and further $\text{Var}(|\mathcal{C}_0|) = E(|\mathcal{C}_0|^2) - E^2(|\mathcal{C}_0|)$ is finite. □

Based on the results in Lemma 6, we will investigate the variance of the independent and identically distributed (i.i.d.) random variable k_i. In fact, k_i is 1 less than $|\mathcal{C}_0|$, since $|\mathcal{C}_0|$ accounts the origin 0 as an α-node. Thus, it is easy to know both the mean of k_i and the variance of k_i are also finite, with $E(k_i) = E(|\mathcal{C}_0|) - 1$ and $\text{Var}(k_i) = \text{Var}(|\mathcal{C}_0|)$. Then, let $\eta(\lambda_\alpha) = E(k_i)$ be the mean of k_i for a given λ_α, i.e., $\eta(\lambda_\alpha) = \text{E}(|\mathcal{C}_0|) - 1$. Then, we will have the following theorem.

Theorem 3: With a density of $0 < \lambda_\beta < c_1(\lambda_\alpha)$ β-nodes in hybrid network $H(\alpha, \beta)$, β-giant component will emerge with probability nearly 1. And, β-giant component size scales linearly with λ_β as $|\mathcal{C}_{(\alpha,\beta)}| = \frac{\eta(\lambda_\alpha)n}{\lambda_\alpha}\lambda_\beta + o(n)$.

Proof: From Theorem 2, when $\lambda_\beta < c_1(\lambda_\alpha)$, all β-nodes in the hybrid network are sufficiently used with probability nearly 1. Then, as we discuss before, if all β-nodes in the hybrid network are sufficiently used, β-component size can be computed by $|\mathcal{C}_{(\alpha,\beta)}| = \Sigma_{i=1}^m k_i$, in which k_i $(i \in \{1, 2, \cdots, m\})$ are i.i.d. random variables with finite mean $\eta(\lambda_\alpha)$ and finite variance. When $\lambda_\beta > 0$, $m \to \infty$, using the *Weak Law of Large Numbers* [20], we know $|\mathcal{C}_{(\alpha,\beta)}|$ converges to $m\eta(\lambda_\alpha) \pm o(m)$ almost surely. Since $m = \lambda_\beta d^2 = \frac{\lambda_\beta n}{\lambda_\alpha}$, we have $|\mathcal{C}_{(\alpha,\beta)}| = \frac{\eta(\lambda_\alpha)n}{\lambda_\alpha}\lambda_\beta + o(n) = \Theta(n)$, which indicates β-giant component emerges. □

Theorem 3 indicates that β-component size scales linearly with λ_β with probability nearly 1, when $0 < \lambda_\beta < c_1(\lambda_\alpha)$. Further, the result in Theorem 3 also shows that with a density of λ_β $(0 < \lambda_\beta < \lambda_\alpha^c - \lambda_\alpha)$ β-nodes in hybrid network, β-giant component, which contains $\Theta(n)$ α-nodes, will emerge with probability nearly 1. Thus, adding base stations in the pure ad hoc network will bring tremendous benefit on the connectivity of α-nodes.

C. Impact of β-Nodes Transmission Range, r_β

As the transmission range of β-nodes, r_β increases from unit, more α-nodes will be inside the disk centered at a single β-node with radius r_β, that is, a single β-node can connect together more connected component of α-nodes, which are

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

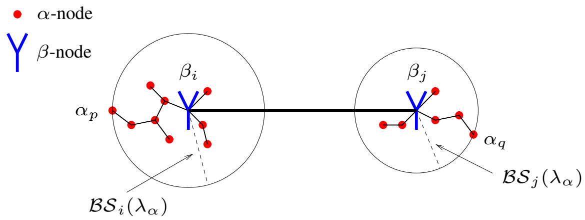


Fig. 9. β_i stretch and β_j stretch in hybrid network.

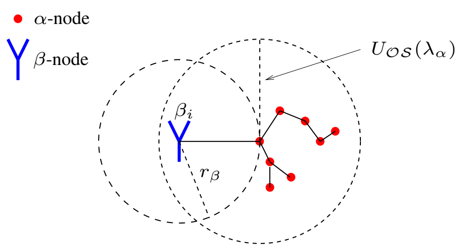


Fig. 10. The upper bound of $\mathcal{BS}_i(\lambda_\alpha)$.

separate without the help of β-nodes. In this section, we will study the impact of r_β on the connectivity of α-nodes. First, we give a new definition as follows.

Definition 4 (β_i Stretch): In hybrid network $H(\alpha, \beta)$, let $\mathcal{C}_{(\alpha,\beta_i)}$ be the connected component containing β_i, and a subset of α-nodes that are connected by β_i "directly". *β_i stretch* is defined as $\mathcal{BS}_i(\lambda_\alpha) = \max_{\alpha_j \in \mathcal{C}_{(\alpha,\beta_i)}} \|\alpha_j - \beta_i\|$. For example, as depicted in Fig. 9, β_i stretch is the Euclidean distance between α_p and β_i, and β_j stretch is the Euclidean distance between α_q and β_j.

Further, let $U_{\mathcal{BS}_i}(\lambda_\alpha)$ be the upper bound of $\mathcal{BS}_i(\lambda_\alpha)$. This definition of upper bound is also from the probabilistic perspective as before, i.e. $\mathcal{BS}_i(\lambda_\alpha) < U_{\mathcal{BS}_i}(\lambda_\alpha)$ with probability nearly 1. Then, we have the following theorem.

Lemma 7: In hybrid network $H(\alpha, \beta)$ for $r_\beta > 1$, if the β-node density satisfies $\lambda_\beta < c_3(\lambda_\alpha)$, all β-nodes in the network are sufficiently used with probability nearly 1, where $c_3(\lambda_\alpha) = \frac{1}{(r_\beta + c_2(\lambda_\alpha))^2}$.

Proof: This proof is very similar to the proof in Theorem 2. When the Euclidean distance between the adjacent β-nodes is large than $U_{\mathcal{BS}_i}(\lambda_\alpha) + 1$, all β-nodes in hybrid network $H(\alpha, \beta)$ are sufficiently used. From Fig. 10, it is easy to know $U_{\mathcal{BS}_i}(\lambda_\alpha) < r_\beta + U_{\mathcal{OS}}(\lambda_\alpha)$ with probability nearly 1. Therefore, when $\lambda_\beta < \frac{1}{(r_\beta + U_{\mathcal{OS}}(\lambda_\alpha)+1)^2}$, we have $\lambda_\beta < \frac{1}{(U_{\mathcal{BS}_i}(\lambda_\alpha)+1)^2}$, which indicates all β-nodes in the network are sufficiently used. □

Lemma 7 provides the sufficient condition that all β-nodes in hybrid network $H(\alpha, \beta)$ are sufficiently used with probability nearly 1, for the case $r_\beta > 1$. Then, in order to utilize the larger β-node transmission range r_β, we propose a new hybrid network architecture below.

First, we let $r_\beta = k_0 c_2(\lambda_\alpha)$ ($k_0 \in \mathbb{N}$). From Lemma 7, when β-node density satisfies $\lambda_\beta < \frac{c_1(\lambda_\alpha)}{(k_0+1)^2}$, all β-nodes in hybrid network $H(\alpha, \beta)$ are sufficiently used with probability nearly 1. Then, a number of additional α-nodes are placed at some specific locations on circles $Circ(\beta_i, kc_2(\lambda_\alpha))$ centered at β_i ($i \in \{1, 2, \cdots, m\}$) with radius $kc_2(\lambda_\alpha)$ ($k \in \{1, 2, \cdots, k_0\}$). The locations of the additional α-nodes are described as follows: For given i and k, let $\alpha_{a1}, \alpha_{a2}, \cdots, \alpha_{aq}$ be all additional α-nodes located on $Circ(\beta_i, kc_2(\lambda_\alpha))$. Without loss of generality, we assume $\alpha_{a1}, \alpha_{a2}, \cdots, \alpha_{aq}$ are in clockwise direction. Then the location of α_{aj} ($j \in \{1, 2, \cdots, q\}$) is at the intersection of $Circ(\beta_i, kc_2(\lambda_\alpha))$ and $Circ(\alpha_{a(j-1)}, c_2(\lambda_\alpha))$. For example, for $k_0 = 1$, i.e., $r_\beta = c_2(\lambda_\alpha)$, as depicted in Fig. 11, six additional α-nodes $\alpha_{a1}, \alpha_{a2}, \cdots, \alpha_{a6}$ are placed on the circle $Circ(\beta_i, c_2(\lambda_\alpha))$. The locations of α_{aj} ($j \in \{1, 2, \cdots, 6\}$) is at the intersection of $Circ(\beta_i, c_2(\lambda_\alpha))$ and $Circ(\alpha_{a(j-1)}, c_2(\lambda_\alpha))$.

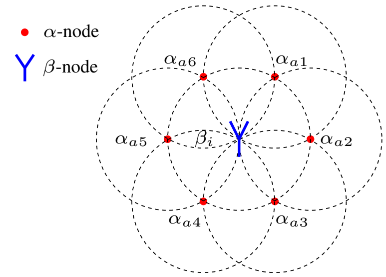


Fig. 11. The locations of additional α-nodes ($k_0 = 1$).

This hybrid network architecture is different from the previous one, because that a subset of α-nodes in the network are deployed at deterministic locations to improve the connectivity of other α-nodes. Therefore, we denote this hybrid network as $\overline{H}(\alpha, \beta)$, and then define a new metric *β-α-component* as follows.

Definition 5 (β-α-Component): In hybrid network $\overline{H}(\alpha, \beta)$, the *$\beta$-$\alpha$-component* denoted by $\overline{\mathcal{C}}_{(\alpha,\beta)}$ is defined as the connected component containing all β-nodes, all additional α-nodes, and a subset of other α-nodes that are connected by the β-nodes and additional α-nodes. Similar to the definition of β-component size, *β-α-component size* $|\overline{\mathcal{C}}_{(\alpha,\beta)}|$ is defined as the number of α-nodes, which are contained by $\overline{\mathcal{C}}_{(\alpha,\beta)}$ but not the additional α-nodes.

Then, we have the following theorem illustrating the benefit on connectivity of α-nodes by using this hybrid network architecture.

Theorem 4: In hybrid network $\overline{H}(\alpha, \beta)$, when β-node density satisfies $\widetilde{\lambda}_\beta < \frac{c_1(\lambda_\alpha)}{(k_0+1)^2}$, with a density of $\Delta\lambda_\alpha = (K_0 - 1)\widetilde{\lambda}_\beta$ additional α-nodes at the specific locations described above, β-α-component size satisfies $|\overline{\mathcal{C}}_{(\alpha,\beta)}| > K_0 \frac{\eta(\lambda_\alpha)n}{\lambda_\alpha} \widetilde{\lambda}_\beta$ with probability nearly 1. Here, $K_0 = 1 + \text{floor}(\frac{2\pi}{arccos(\frac{1}{2})}) + \cdots + \text{floor}(\frac{2\pi}{arccos(\frac{2k_0^2-1}{2k_0^2})})$.

Proof: For $k_0 = 1$, as depicted in Fig. 11, we denote $(k_i)_1, (k_i)_2, \cdots, (k_i)_6$ respectively as the number of α-nodes connected by each additional α-node "directly". Further, let $(k_i)_7$ be the number of α-nodes connected by β_i "directly" when $r_\beta = 1$. From our discussion in Section IV-B, $(k_i)_j$ ($j \in \{1, 2, \cdots, 7\}$) are i.i.d. random variables with finite mean $\eta(\lambda_\alpha)$ and finite variance. Obviously, in hybrid network

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

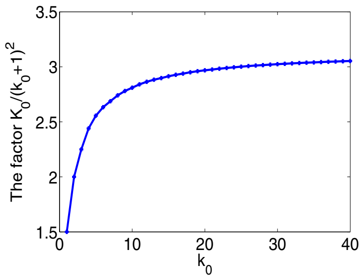


Fig. 12. The factor $\frac{K_0}{(k_0+1)^2}$ for different k_0.

$\overline{H}(\alpha,\beta)$, the number of α-nodes connected by β_i and additional six α-nodes "directly", is larger than $\Sigma_{j=1}^{7}(k_i)_j$. From the Weak Law of Large Numbers [20], β-α-component size $|\overline{\mathcal{C}}_{(\alpha,\beta)}|$ is larger than $7m\eta(\lambda_\alpha) \pm o(m)$ almost surely.

Then we consider the general case for $r_\beta = k_0 c_2(\lambda_\alpha)$. Following the same logic as above, it is not difficult to know β-α-component size $|\overline{\mathcal{C}}_{(\alpha,\beta)}|$ is larger than $K_0 m\eta(\lambda_\alpha) \pm o(m)$ almost surely, where $K_0 = 1 + \text{floor}(\frac{2\pi}{\arccos(\frac{1}{2})}) + \cdots + \text{floor}(\frac{2\pi}{\arccos(\frac{2k_0^2-1}{2k_0^2})})$. Then, β-α-component size satisfies $|\overline{\mathcal{C}}_{(\alpha,\beta)}| \geq K_0 \frac{\eta(\lambda_\alpha)n}{\lambda_\alpha}\widetilde{\lambda}_\beta$. At this time, $\widetilde{\lambda}_\beta < \frac{c_1(\lambda_\alpha)}{(k_0+1)^2}$ and $\Delta\lambda_\alpha = (K_0-1)\widetilde{\lambda}_\beta$. □

We plot the factor $\frac{K_0}{(k_0+1)^2}$ in Fig. 12. As we can see from the figure, when k_0 increases from 1 to 40, the factor $\frac{K_0}{(k_0+1)^2}$ increases from 1.5 to 3.05. From Theorem 4, we know $\widetilde{\lambda}_\beta + \Delta\lambda_\alpha = K_0\widetilde{\lambda}_\beta < \frac{K_0}{(k_0+1)^2}c_1(\lambda_\alpha)$, and $|\overline{\mathcal{C}}_{(\alpha,\beta)}| \geq K_0\frac{\eta(\lambda_\alpha)n}{\lambda_\alpha}\widetilde{\lambda}_\beta = \frac{\eta(\lambda_\alpha)n}{\lambda_\alpha}(\widetilde{\lambda}_\beta + \Delta\lambda_\alpha)$. Compared with the result in Section IV-B, $|\mathcal{C}_{(\alpha,\beta)}| = \frac{\eta(\lambda_\alpha)n}{\lambda_\alpha}\lambda_\beta$ when $\lambda_\beta < c_1(\lambda_\alpha)$, we find that through increasing r_β from unit to $k_0 c_2(\lambda_\alpha)$ ($k_0 \in \mathbb{N}$), we enlarge the linear scale region from $(0, c_1(\lambda_\alpha))$ to $(0, \frac{K_0}{(k_0+1)^2}c_1(\lambda_\alpha))$, while decrease the required β-node density, and thus obtain additional benefit on the connectivity of α-nodes. This result provides a guideline on constructing the hybrid network architecture. Further, from Fig. 12, as k_0 increases, the factor $\frac{K_0}{(k_0+1)^2}$ increases slower. Taking k_0 near 10 would be an optimal choice, though we still need to consider the receiver complexity and the transmission power etc. of the base stations. Therefore, the result also provides us with a reference on how to choose the value of r_β.

V. Conclusions

In this paper, we have studied how base stations in large-scale hybrid network improve the connectivity of ad hoc nodes in subcritical phase. This problem is two-fold: the effect of base station density λ_β and the effect of base station transmission range r_β on network connectivity. For r_β is unit, through investigating the upper bound of origin stretch in pure ad hoc networks, we find the sufficient condition, i.e., $\lambda_\beta < c_1(\lambda_\alpha)$, that any two base stations cannot be connected through the relay of ad hoc nodes with probability nearly 1. Based on this result, we find that under $0 < \lambda_\beta < c_1(\lambda_\alpha)$, the number of connected ad hoc nodes is $\Theta(n)$ and scales linearly with λ_β with probability nearly 1, which demonstrates a tremendous benefit of using base stations to enhance the connectivity of ad hoc nodes. Further, we study a hybrid network architecture which can utilize large base station transmission range r_β, to obtain an additional benefit on the connectivity of ad hoc nodes. The work in this paper provides fundamental understanding of the connectivity in hybrid networks, and also a provide guidelines on design of a new hybrid network that contains low density of ad hoc nodes.

References

[1] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," in *Proc. of IEEE Conference on Decision and Control*, December 1998, pp. 1106 – 1110.
[2] F. Xue and P. R. Kumar, "The number of neighbors needed for connectivity of wireless networks," *Wireless Networks*, vol. 10, no. 2, pp. 169–181, March 2004.
[3] P. Santi and D. M. Blough, "The critical transmitting range for connectivity in sparse wireless ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 2, no. 1, pp. 25–39, January 2003.
[4] P. Santi, "The critical transmitting range for connectivity in mobile ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 4, no. 3, pp. 310–317, June 2005.
[5] O. Dousse, F. Baccelli, and P. Thiran, "Impact of interferences on connectivity in ad hoc networks," *IEEE/ACM Transactions on Networking*, vol. 13, no. 2, pp. 425–436, April 2005.
[6] Z. Kong and E. M. Yeh, "Connectivity and latency in large-scale wireless networks with unreliable links," in *Proc. of IEEE Infocom*, April 2008, pp. 394 – 402.
[7] R. Meester and R. Roy, *Continuum Percolation*. Cambridge University Press, 1996.
[8] M. Penrose, *Random Geometric Graphs*. Oxford University Press, 2003.
[9] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, March 2000.
[10] O. Lévêque and I. E. Telatar, "Information-theoretic upper bounds on the capacity of large extended ad hoc wireless networks," *IEEE Transaction on Information Theory*, vol. 51, no. 3, pp. 858–865, March 2005.
[11] P. L, C. Zhang, and Y. Fang, "Capacity and delay of hybrid wireless broadband access networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 2, pp. 117–125, February 2009.
[12] B. Liu, Z. Liu, and D. Towsley, "On the capacity of hybrid wireless networks," in *Proc. of IEEE Infocom*, March 2003, pp. 1543 – 1552.
[13] U. C. Kozat and L. Tassiulas, "Throughput capacity of random ad hoc networks with infrastructure support," in *Proc. of ACM MobiCom*, September 2003, pp. 55 – 65.
[14] A. Agarwal and P. R. Kumar, "Capacity bounds for ad hoc and hybrid wireless networks," in *Proc. of ACM SIGCOMM*, July 2004, pp. 71 – 81.
[15] X. Mao, X.-Y. Li, and S. Tang, "Multicast capacity for hybrid wireless networks," in *Proc. of ACM MobiHoc*, May 2008, pp. 189 – 198.
[16] P. Balister, B. Bollobás, and M. Walters, "Continuum percolation with steps in the square or the disc," *Random Structures and Algorithms*, vol. 26, no. 4, pp. 392–403, April 2005.
[17] Y. Shi, Y. T. Hou, and A. Efrat, "Algorithm design for base station placement problems in sensor networks," in *Proc. of ACM QShine*, August 2006.
[18] Y. Shi and Y. T. Hou, "Approximation algorithm for base station placement in wireless sensor networks," in *Proc. of IEEE SECON*, June 2007, pp. 512–519.
[19] G. Grimmett, *Percolation*. Springer, 1999.
[20] R. Durrett, *Probability: Theory and Examples*. Duxbury Press, 1996.