Heterogeneity Increases Multicast Capacity in Clustered Network

Qiuyu Peng, Xinbing Wang, Huan Tang
Dept. of Electronic Engineering
Shanghai Jiao Tong University, China

***Abstract*—In this paper, we investigate the multicast capacity for static network with heterogeneous clusters. We study the effect of heterogeneities on the achievable capacity from two aspects, including *heterogeneous cluster traffic* (HCT) and *heterogeneous cluster size* (HCS). HCT means cluster clients are more likely to appear near the cluster head, instead of being uniformly distributed across the network and HCS means each cluster is also not equal in size as most prior literatures assume. Both of these two properties are commonly found in realistic networks. For this class of networks, we find that HCT increases network capacity for all the clusters and HCS does not influence the network capacity. Our work can generalize various results obtained under non-heterogeneous networks in the literature.**

I. Introduction

Wireless network is modeled as a set of nodes that send and receive messages over a common wireless channel. Since the seminal work done by P. Gupta, P. R. Kumar [1], there is significant interest toward the asymptotic capacity of the network when the number of nodes n grows. The authors of [1] prove that the per-node capacity[1] is $\Theta(W/\sqrt{n\log n})$ in static network. Then M. Franceschetti, *et al*. [2] design an optimal routing protocol with capacity achieving $\Theta(W/\sqrt{n})$ via percolation theory and the scaling laws of broadband network are investigated by P. Li, *et al*. [3].

Multicast traffic, which generalizes the above unicast traffic, receives more attention recently and the estimation of the achievable multicast capacity is required in many applications like sensor network, TV streaming. Li, *et al*. [4] study the achievable capacity in multicast network. In their work, there are n multicast sessions, each comprises of 1 source and k destinations and they find that the capacity scales as $\Theta(1/\sqrt{kn\log n})$ based on the *Manhattan Routing* scheme. Their results generalize both unicast and broadcast [5] capacity. In [6], Shakkottai, *et al*. study a different multicast framework where there are n^ϵ multicast sources and $n^{1-\epsilon}$ destinations per flow. Their network can support a rate of $\Theta(\frac{1}{\sqrt{n^\epsilon \log n}})$ for each flow. Later on, the multicast capacity under Gaussian channel is obtained in [7], [8]. The achievable capacity in mobile multicast (motioncast) is explored in [9] and optimail mobile multicast capacity is presented in [10], which is a generalization of [11], [12]. And in [13], [14], MIMO cooperations are introduced to improve multicast capacity.

Since nodes in the same multicast session can be treated as members of a cluster, network supporting multicast traffic can also be viewed as clustered network. However, there are few work concerning the cluster behavior of multicast traffic. Uniformly distributed cluster (muticast session) traffics and sizes are assumed and cluster heterogeneities are rarely involved in previous works. Actually, most realistic networks are characterized by various clustered heterogeneities and some aspects have already been investigated in network with unicast traffic, which includes

Spatial heterogeneity: Wireless nodes are not likely to be uniformly distributed across the deployed region in realistic networks e.g., wireless users may cluster in urban areas so there are less users in suburban areas. Due to spontaneous grouping of the nodes around a few attraction points are common in wireless network, G. Alfano, *et al*. [15], [16] concern the capacity scaling with inhomogeneous node density. In their work, nodes are generated according to a specified point process and they show that the bottleneck is in node spares region because the network capacity is related to the minimal node intensity.

Pattern heterogeneity: It is likely that there exists more than one type of traffic patterns in the network and nodes of the same traffic patterns constitutes a cluster. In [17], Wang, *et al*. study a unified modeling framework composed of unicast, multicast, broadcast traffics. Later in [18], Ji, *et al*. explore the network composed of both unicast and converge-cast traffics and show that MIMO cooperation can be applied to increase capacity for both traffics. Li, *et al*. [19] deal with network containing some helping nodes for packets delivery, therefore the normal nodes and helping nodes can be viewed as two clusters.

The network heterogeneities investigated in prior works are inadequate for exploring the clustering behavior of multicast network. For instance, in military battlefield, commanders from different places must send requirements through a common wireless channel to their respective soldiers around them. In sensor network, local schedulers also need to send packets to their adjacent client sensors. Since clients of the same data flow are both non-uniformly distributed and size varied, now the open question is:

- What are the impacts of heterogeneous traffic and cluster size on multicast capacity in clustered network?

[1]Given two functions $f(n) > 0$ and $g(n) > 0$: $f(n) = o(g(n))$ means $\lim_{n\to\infty} f(n)/g(n) = 0$; $f(n) = O(g(n))$ means $\lim_{n\to\infty} \sup f(n)/g(n) < \infty$; $f(n) = \omega(g(n))$ is equivalent to $g(n) = o(f(n))$; $f(n) = \Omega(g(n))$ is equivalent to $g(n) = O(f(n))$; $f(n) = \Theta(g(n))$ means $f(n) = O(g(n))$ and $g(n) = O(f(n))$.

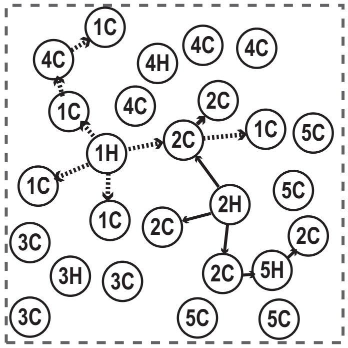

Fig. 1: Demonstration of Network topology. Nodes in the same cluster are labeled with the same number. H and D represent head (source) and clients (destination), respectively.

In this paper, nodes in each multicast session comprises of a cluster and the network heterogeneities include:

Heterogeneous Cluster Traffic (HCT): Clients of the same cluster (data flow) are likely to be deployed around a cluster head specified by an inhomogeneous poison process (IPP). We describe this clustering behavior with a variable $\sigma_{\mathcal{O}}$, which depicts the extent of HCT. We find that the network capacity increases in this case because the total transmission length is shortened and we offer a quantitative relationship between the network capacity and $\sigma_{\mathcal{O}}$.

Heterogeneous Cluster Size (HCS): Clusters may have different size (cardinality) and HCS is employed to describe the population variation for each multicast data flow. We show that the network capacity is inherently related to the total number of clusters and not affected by HCS.

The rest of the paper is organized as follows. In section II, we outline some preliminaries of the network and our main results. In section III and IV, a close form of the upper bound and maximized per-cluster capacity are derived respectively. In section V, we provide a routing scheme for the achievable capacity for *uniform random cluster model*. A discussion of the results is presented in section VI. Finally, we conclude this paper in section VII.

II. Preliminaries and Main Results

A. Network Topology

We consider networks composed of $n_s = n^{\alpha}$ $(0 \leq \alpha \leq 1)$ clusters distributed over a 2-dimensional torus region $\mathcal{O}$ of edge $L = n^{\beta}$. We specify a homogeneous poisson process (HPP) to generate cluster head v_j, whose position is denoted by k_j for cluster $C_j (1 \leq j \leq n_s)$. Then, v_j generates its cluster members according to an IPP whose intensity at ξ is given by $|\mathcal{C}_j|\phi(k_j, \xi)$, where $|\mathcal{C}_j| \leq p = n^{1-\alpha}$ is the expected size of the cluster and $\phi(\cdot)$ is the dispersion density function. Cluster dense regime is assumed such that $0 \leq \beta \leq \alpha/2$. In [15], they prove that the minimum and maximum node density are of the same order in this case.

These n_s clusters may have different size but we assume that $|\mathcal{C}_j| \leq p$ $(1 \leq i \leq n_s)$ and there are at least $c_0 n_s$ $(0 < c_0 < 1)$ clusters with size $\Theta(p)$. It indicates that the number of clusters of size $\Theta(p)$ can be viewed as backbone for packet delivery, which is useful for constructing our routing policy. As to the dispersion density function $\phi(\cdot)$, the following properties are satisfied:

1) $\phi(k_j, \xi)$ is invariant under both translation and rotation with respect to k_j, therefore $\phi(k_j, \xi)$ can be rewritten as $\phi(|k_j - \xi|)$ and it is a non-increasing function with respect to the Euclidean distance $|k_j - \xi|$.
2) Integrate $\phi(k_j, \xi)$ of ξ over the whole torus $\mathcal{O}$ equals 1, $\int_{\mathcal{O}} \phi(k_j, \xi) d\xi = 1$.

Under the above assumptions, cluster clients are likely to be distributed near cluster head and the cluster size conforms to a poisson distribution with rate $|\mathcal{C}_j|$. A standard application of Chernoff bound reveals that the size of $\mathcal{C}_j$ is $\Theta(|\mathcal{C}_j|)$ and we take $|\mathcal{C}_j|$ as the cluster size to simplify our analysis.

Now we need to offer a quantitative value for a given dispersion density function $\phi(\cdot)$ to depict its degree of heterogeneity. As we know, the expectation can describe the average value of a function and variance can describe the fluctuation level of a signal around its expectation. The expectation of $\phi(\cdot)$ is the average node density of a cluster and

$$\mathbf{E}[\phi(x)] = \int_{\mathcal{O}} \frac{\phi(\xi)}{L^2} dx = \frac{1}{L^2}.$$

The variance of $\phi(\cdot)$, also defined as *distribution variance* $\sigma_{\mathcal{O}}$, can be utilized to depict HCT:

$$\sigma_{\mathcal{O}}^2 = \int_{\mathcal{O}} \left(\phi(\xi) - \mathbf{E}[\phi(\xi)]\right)^2 d\xi = \int_{\mathcal{O}} \phi^2(\xi) d\xi - \frac{1}{L^2}.$$

We omit the term k_j due to wrap around property of torus and it can liberate us from the border effect. In case of uniform traffic, $\phi(\xi) \equiv \frac{1}{L^2}$, $\sigma_{\mathcal{O}} = 0$ and larger heterogeneity results in larger $\sigma_{\mathcal{O}}$. Figure 1 is an example of the network topology with both HCT and HCS. Finally, we specify a special point process as *uniform cluster random model* (UCRM) whose dispersion density function is as follows:

$$\phi^u(\xi) = \begin{cases} \frac{1}{\pi R^2} & |\xi| \leq R \\ 0 & \text{otherwise} \end{cases}$$

where $R = \frac{L}{\sqrt{\pi(1+(L\sigma_{\mathcal{O}})^2)}}$ is defined as cluster radius. It means clients of each cluster are randomly and uniformly distributed in a disk of radius R centered at its cluster head. We prove that this topology leads to the maximized capacity when $\sigma_{\mathcal{O}}$ is fixed in section IV.

B. Transmission Protocol

All wireless transceivers can communicate over a common channel of limited bandwidth W. We adopt the protocol model for interference proposed in [1]. In each time slot, a sender i can successfully transmit at W bit/second to a destination j when the Euclidean distance between any other active transmitters and j is larger than $(1 + \Delta)R_{i,j}$, where $R_{i,j}$ is the Euclidean distance between i and j; Δ is a positive constant independent of the position of i, j, k and it specifies a guard zone for a successful transmission.

C. Traffic Model

Multicast traffic pattern is assumed where each cluster head generates data flows to their respect clients, e.g. in Figure 1. The *one to many* data flow in $\mathcal{C}_j$ can be modeled as a multicast tree $\mathcal{T}_j$ spanning 1 head and $|\mathcal{C}_j|$ clients. In [4], *Euclidean minimal spanning tree* (EMST) is employed to bound the length of transmission for each multicast session in non-clustered network with uniform node distributions. We employ new techniques to study heterogenous clustered network, which generalizes uniform cases. Let $EMST(\mathcal{C}_j)$ denote the EMST for $\mathcal{C}_j$. Noth that the communication between any SD pairs can also go through multiple relays from other clusters. Now we give the definition of capacity.

Definition of Asymptotic Capacity: Let $\lambda_j (1 \leq j \leq n_s)$ denote the sustainable rate of data flow for cluster $\mathcal{C}_j$. Assume that $\lambda = \min\{\lambda_1, \lambda_2, \ldots, \lambda_{n_s-1}, \lambda_{n_s}\}$. Then $\lambda = \Theta(f(n))$ is defined as the asymptotic network capacity if there exist constants $c > c' > 0$, such that

$$\lim_{n\to\infty} \mathbf{Pr}(\lambda = cf(n) \text{ is achievable}) < 1,$$
$$\lim_{n\to\infty} \mathbf{Pr}(\lambda = c'f(n) \text{ is achievable}) = 1.$$

Therefore λ is the minimal achievable data rate in these clusters.

D. Mathematical Notations

Throughout our paper, we denote $h^b_{\mathcal{C}_j}$ as the number of hops required for transmitting bit b to all clients in $\mathcal{C}_j$. ℓ^h_b is the length of transmission of bit b in its $h_{th} (1 \leq h \leq h^b_{\mathcal{C}_j})$ hop. $D(\xi, R)$ is the circular region centered at ξ with radius R. $\mathcal{R}$ is the radius of the influential region centered at the head, nodes outside the influential region are impossible to act as relays for that cluster. $|\mathcal{T}_j|$ is the total Euclidean length of a multicast tree $\mathcal{T}_j$.

E. Main Results

Our main results are summarized as follows:

- Given the dispersion density function $\phi(\cdot)$, the upper bound of capacity is given as follows:
$$\lambda \leq \min\left\{W, \frac{cWL}{\sqrt{n_s}\int_{\mathcal{O}}\sqrt{\phi(\xi)}d\xi}\right\},$$
where c is some constant.
- HCT increases the maximized network capacity λ and a universal relationship between λ and $\sigma_{\mathcal{O}}$ is obtained as:
$$\lambda \leq \min\left\{O(W), O\left(\frac{\max\{1, L\sigma_{\mathcal{O}}\}W}{\sqrt{n_s}}\right)\right\}.$$
- HCS does not affect the network capacity λ.

III. Derivation of Upper Bound to Multicast Capacity

In this section, we will discuss several restrictions inherent in the proposed network model irrespective of the routing policy. There are some tradeoffs that must be tolerated among number of hops, transmission range, limited radio resources and so force. Therefore, a thorough comprehension of the implicit relationships among them is constructive for deriving the upper bound of the achievable capacity.

Since it consumes radio resources to forward a bit b to relays or destinations. The following lemma captures the tradeoffs among number of hops, transmission range, limited radio resources.

Lemma 3.1: Constraint of Protocol model: Under the protocol model, the following inequality must be held for any routing scheme when the simulation time T is sufficient large.

$$\sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T}\sum_{h=1}^{h^b_{\mathcal{C}_j}} \frac{\pi}{16}\Delta^2(\ell^h_b)^2 \leq WTL^2. \tag{1}$$

Proof: Here, we briefly outline the intuition behind this inequality. The proof resembles the proof in [11]. Based on protocol model, each receiver should be at a distance of $(1+\Delta)$ times the transmission range from other active transmitter. This will impose some restrictions on the minimum distance between two active transmitters as a result. Therefore there is a guard region for each transmitter that other active transmitters cannot reside in. ■

The above inequality is a basic requirement for multihop transmission fashion and the cooperative MIMO as in [14] is not considered here. In static network, there is a minimal transmission range r to guarantee full network connectivity. In network with uniform node distribution, it is proved that the transmission range $r = \Omega(\sqrt{\log n/n})$ is sufficient for network connectivity *w.h.p.* in [1]. Now we need to characterize a feasible transmission range r, below which the transmission is not possible in our framework. In other words, we want to derive r such that the number of nodes in $D(\xi, r)$ should not be zero all the time.

Lemma 3.2: Let $\mathcal{N}(r)$ represents the number of nodes within the transmission range r when a node wants to transmit its packets, then $N(r) = \Theta(\frac{n_s pr^2}{L^2})$ when $r = \Omega(\frac{L}{\sqrt{n_s p}})$.

Proof: In our cluster dense regime, when $r = \omega(\frac{1}{\sqrt{n_s p/L}})$, $\mathcal{N}(r) \leq \frac{2\pi\sum_{j=1}^{n_s}|\mathcal{C}_j|r^2}{L^2} \leq \frac{2\pi n_s pr^2}{L^2}$ based on Chernoff bound and Riemann sum (one can refer to Theorem 1 in [15]). Similarly, the lower bound of $\mathcal{N}(r)$ is as $\mathcal{N}(r) \geq \frac{c_0\pi n_s pr^2}{2L^2}$.

Then we come to $r = \Theta(\frac{L}{\sqrt{n_s p}})$ case. Note that the node density is upper bound by a HPP with rate $\mu = \frac{\pi r^2 n_s p}{L^2}$, the probability that the number of nodes inside the disk exceeding n_0 is

$$\mathbf{Pr}(\mathcal{N}(r) > n_0) \leq e^{-\mu}\sum_{i=n_0+1}^{\infty}\frac{\mu^i}{i!} \leq \frac{e^{(\theta-1)\mu}\mu^{n_0+1}}{(n_0+1)!}.$$

During the above derivation, we use Lagrange form of the remainder term and $0 \leq \theta \leq 1$. Therefore when $n_0 = \omega(\frac{n_s pr^2}{L^2})$, $\mathbf{Pr}(\mathcal{N}(r) > n_0) = 0$ w.h.p. Note that $\mathcal{N}(r) \geq 1$ is a prerequisite for any transmissions and we complete our proof. ■

Therefore we know a necessary condition for the transmission range r is $r = \Omega(\sqrt{\frac{L}{n_s p}})$. In our network, each node

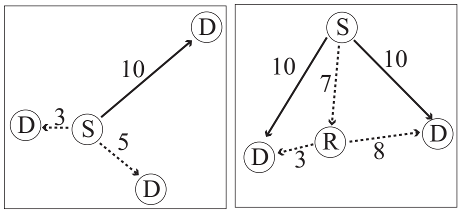

Fig. 2: Demonstration of the two conditions for MTTL.

adopts the same transmission range r in order sense, substitute $\ell_b^h = r$ into eqn. (1), we obtain

$$\sum_{j=1}^{n_s} \sum_{b=1}^{\lambda_j T} h_{\mathcal{C}_j}^b \leq \frac{16WTL^2}{\pi \Delta^2 r^2}. \quad (2)$$

Then we need to calculate the number of hops (transmission) required of bit b $h_{\mathcal{C}_j}^b$. In [4], they prove that nearest neighbor transmission can also achieve optimal capacity in uniform traffic distribution. They obtain that the number of hops required is $\sqrt{kn}$ for k destinations. The situation is complicated here since traffic is not uniformly distributed across the network. Clients in the same multicast session is more clustered around the head and a larger transmission range can cover more clients in a cluster when the transmission happens near the cluster head (source) than other places. Therefore, it is unknown whether much more destinations are involved in a larger transmission range can compensate for the sacrifice of the radio resources. Unfortunately, the following analysis reveals that nearest neighbor transmission is also optimal in HCT.

In [4], EMST is investigated in multicast traffic and it can help us obtain the capacity upper bound in non-heterogeneous networks. To obtain the results under heterogeneous cluster traffic and size, we first introduce the following lemma. Theorem 1 in [21] as follows.

Lemma 3.3: If f is the density of the probability function for picking points, then for large n and $d \neq 1$, the size of the EMST is approximately $c(d) n^{\frac{d-1}{d}} \int_{\mathbb{R}^d} f(x)^{\frac{d-1}{d}} dx$, where $c(d)$ is a constant depending only on the dimension d.

The proof can be found in Theorem 1, [21]. Our case corresponds to $d = 2$ and such that

$$|EMST(\mathcal{C}_j)| = \Theta\left(\sqrt{|\mathcal{C}_j|} \int_{\mathcal{O}} \sqrt{\phi(\xi)} d\xi\right).$$

It reveals that the length of the Euclidean minimum multicast spanning tree constitutes two terms, the square root of the number of nodes $\sqrt{|\mathcal{C}_j|}$ and $\int_{\mathcal{O}} \sqrt{\phi(\xi)} d\xi$. Note that we eliminate the constants $c(d)$ for simplicity. Intuitively, the minimum length for connecting all the nodes is the minimum total transmission length (MTTL). However, the length of EMST is far more than MTTL for the following reasons.

- Larger transmission range can cover more than one nodes, but only the length of the longest SD pair is counted. For instance, if a node broadcast its message to all the other nodes in one time, the MTTL is at most L. Therefore MTTL is related to the transmission range r.
- Nodes from other clusters can act as relays to help forward information. We must consider the impact of relay to the MTTL.

Figure 2 illustrates two examples of the above two questions, respectively. In order to answer the first question, in other words, derive the relationship between the transmission range r and MTTL, we construct a new concept *ρ-simplified cluster* $\rho - \mathcal{C}_j$. It is used to generate a less dense nodes distributions given $\mathcal{C}_j$ and $\phi(\cdot)$. Algorithm 1 illustrates how to generate $\rho - \mathcal{C}_j$ given $\mathcal{C}_j$ and the dispersion density function $\phi(\cdot)$.

Algorithm 1 Generation of ρ-simplified EMST from EMST.
Input: $\mathcal{C}_j$, $\phi(\cdot)$ Output: $\rho - \mathcal{C}_j$

1: Specify two point sets $S \leftarrow \emptyset$, $S' \leftarrow \mathcal{C}_j$.
2: Random label each nodes in S' with number $1, 2, \ldots, |\mathcal{C}_j| - 1, |\mathcal{C}_j|$.
3: Choose nodes with the smallest labeled number in S'.
4: Add the chosen nodes to S and discard all the nodes within $D(\xi', \rho)$ in S', where ξ' is the position of the chosen node.
5: Back to step 3 until no node is left in S', $(\rho - \mathcal{C}_j) \leftarrow S'$.

Let $\rho - \mathcal{T}_j$ denote the multicast tree spanning $\rho - \mathcal{C}_j$. The length of each branch in $\rho - \mathcal{T}_j$ is larger than ρ and all the abandoned nodes are within a distance ρ from the nodes in $\rho - \mathcal{C}_j$. $|\rho - \mathcal{T}_j| \geq |EMST(\rho - \mathcal{C}_j)|$ according to the definition of EMST. Because the point intensity after the thinning process determines $|EMST(\rho - \mathcal{C}_j)|$ according to lemma 3.3, we specify two regions according to $\phi(\cdot)$. Let $\phi'(\xi)$ denote the point intensity after Algorithm 1.

- **Dense Region** ($S_{1,j}$) Nodes in this region are populous and we specify a radius
$$\widetilde{\rho_j} = \sup\{\rho_j, \phi(\rho_j) \geq \frac{1}{\pi \rho^2 |\mathcal{C}_j|}\}$$
for this circular region $S_{1,j}$ because $\phi(\cdot)$ is invariant under rotations. After the thinning process, $\phi'(\xi) \geq \Theta(\frac{1}{\pi \rho^2 |\mathcal{C}_j|})$ on the basis of Chernoff bound.
- **Sparse Region** ($S_{2,j} = \mathcal{O}/S_{1,j}$) Nodes in this region are relative sparse such that there are at most a constant number of nodes in $D(\xi, \rho)$ if $\xi \in S_{2,j}$. After the thinning process, the node density is at least a constant fraction of the original density. Therefore $\phi'(\xi) \geq \Theta(\phi(\xi))$.

Then the following lemma is used to answer the first question.

Lemma 3.4: There exists a constant $c_2 > 0$ such that $|EMST(\rho - \mathcal{C}_j)|$ is lower bounded as

$$|EMST(\rho - \mathcal{C}_j)| \geq c_2 \sqrt{|\mathcal{C}_j|} \left(\frac{\sqrt{\pi} \widetilde{\rho_j}^2}{\rho} + \int_{S_{2,j}} \sqrt{\phi'(\xi)} d\xi\right).$$

Proof: The length of EMST is determined by the point intensity according to lemma 3.3, thus there exists a constant

$c' > 0$, such that

$$|EMST(\rho - \mathcal{C}_j)| \geq c' \int_{\mathcal{O}} \sqrt{|\mathcal{C}_j| \phi'(\xi)} d\xi$$
$$= c' \sqrt{|\mathcal{C}_j|} \left(\int_{S_{1,j}} \sqrt{\phi'(\xi)} d\xi + \int_{S_{2,i}} \sqrt{\phi'(\xi)} d\xi \right)$$
$$\geq c_2 \sqrt{|\mathcal{C}_j|} \left(\frac{\sqrt{\pi} \tilde{\rho}_j^2}{\rho} + \int_{S_{2,j}} \sqrt{\phi(\xi)} d\xi \right),$$

where c_2 is also a constant. Note that our result holds even when $S_{1,j}$ or $S_{2,j}$ is empty. ■

Now we come to the second question. The following lemma tells us that $|EMST(\rho - \mathcal{C}_j)|$ is at most $\frac{\sqrt{3}}{2}$ times larger than MTTL when relays are utilized, which is proved in [20].

Lemma 3.5: Given k nodes U, any multicast tree spanning these k nodes (may be using some additional relay nodes) will have an Euclidean length at least $\varrho ||EMST(U)||$, where $\varrho = \sqrt{3}/2$ and $\varrho ||EMST(U)||$ is the EMST spanning U.

Then we obtain the lower bound of MTTL for $\mathcal{C}_j$ as

$$MTTL(\mathcal{C}_j) \geq \Theta(|EMST(r - \mathcal{C}_j)|).$$

Based on the above inequality, we obtain a lower bound of $h^b_{\mathcal{C}_j}$ as:

$$h^b_{\mathcal{C}_j} \geq \frac{MTTL(\mathcal{C}_j)}{r} \geq \frac{\Theta(|EMST(r - \mathcal{C}_j)|)}{r} \quad (3)$$

Note that what we are interested in is the order sense of the result, therefore we directly regard r as the transmission length. Now we will investigate the upper bound of network capacity λ based on the previous analysis of the restrictions imposed by the network. The results obtained here are some fundamental limits that cannot be violated by any routing protocols.

Theorem 3.1: Under the assumptions of the proposed wireless network, the following tradeoffs must be satisfied by all scheduling policy.

$$\sum_{j=1}^{n_s} \lambda_j \sqrt{|\mathcal{C}_j|} \leq c \frac{\sqrt{n_s p} WL}{\int_{\mathcal{O}} \sqrt{\phi(\xi)} d\xi}, \quad (4)$$

where $\lambda_j = O(W)$ and c is a constant.

Proof: The proof is presented in Appendix A. ■

Theorem 3.2: The achievable capacity λ in our network is upper bounded as follows:

$$\lambda \leq \min \left\{ W, \frac{cWL}{\sqrt{n_s} \int_{\mathcal{O}} \sqrt{\phi(\xi)} d\xi} \right\}. \quad (5)$$

Proof: This theorem is a trivial extension of Theorem 3.1 just by substituting $\lambda \leq \frac{\sum_{j=1}^{n_s} \lambda_j}{n_s}$ into Eqn. (4). ■

IV. Maximized Capacity with Distribution Variance Constrained

In this section, we study the relationship between maximized capacity and *distribution variance* $\sigma_{\mathcal{O}}$. And we can predict the maximized achievable capacity just by knowing the *distribution variance* irrespective of the exact points process. Recall the definition of $\sigma_{\mathcal{O}}$, there are various dispersion functions $\phi(\cdot)$ satisfied given a fixed $\sigma_{\mathcal{O}}$. Now we will discuss what is the maximized capacity in this set of dispersion functions.

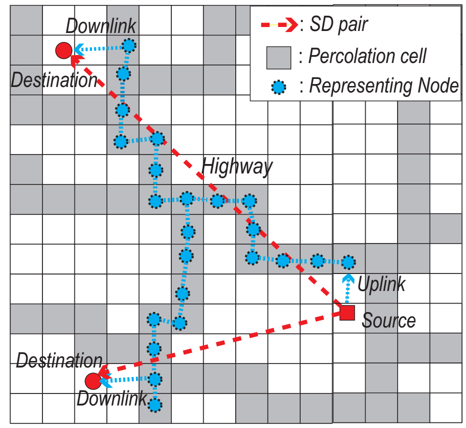

Fig. 3: Demonstration of routing protocol.

Theorem 4.1: Given the *distribution variance* $\sigma_{\mathcal{O}}$, the maximized network capacity λ is bounded as follows:

$$\lambda \leq \min \left\{ O(W), O\left(\frac{\max\{1, L\sigma_{\mathcal{O}}\} W}{\sqrt{n_s}} \right) \right\}. \quad (6)$$

The respect dispersion function is identical to $\phi^u(\xi)$.

According to Theorem 3.1, a smaller $|EMST(\mathcal{C}_j)|$ results in a larger capacity. Therefore we will derive the minimal $|EMST(\mathcal{C}_j)|$ given a fixed $\sigma_{\mathcal{O}}$.

Theorem 4.2: Define a real variable function $\Upsilon(\phi(\cdot)) = \int_{\mathcal{O}} \sqrt{\phi(\xi)} d\xi$, then we can prove that $\phi^u(\cdot)$ can minimize $\Upsilon(\phi(\cdot))$ among all the $\phi(\cdot)$ with distribution variance $\sigma_{\mathcal{O}}$.

Proof: The proof is presented in Appendix B. ■

Recall Theorem 3.2, we complete the proof of Theorem 4.1. In this case, the node distribution just conforms to the proposed UCRM. In this model, cluster members are uniformly and randomly distributed in a disk of radius R centered at the cluster head. In next section, we will provide the routing scheme to approach this bound and verify the maximized capacity is achievable in order sense.

V. Capacity Achieving Scheme of Uniform Cluster Random Model

In this section, we provide a routing scheme for the proposed UCRM based on percolation theory. We find that only when the length of the multicast spanning tree $|\mathcal{T}_j|$ is on the same order with $|EMST(\mathcal{C}_j)|$, the per-cluster capacity can approach the theoretical upper bound in order sense.

A. When $\sigma_{\mathcal{O}} = \Omega(\frac{\sqrt{n_s}}{L})$

In this case, $R = \frac{L}{\sqrt{\pi(1+L^2(\sigma_{\mathcal{O}})^2)}} = O(\frac{L}{\sqrt{n_s}})$, there are at most a constant number of clusters inside $D(\xi, R)$ for $\xi \in \mathcal{O}$ and a simple TDMA scheme can achieve $\Theta(W)$ capacity for each cluster.

B. When $\sigma_{\mathcal{O}} = o(\frac{\sqrt{n_s}}{L})$

In this case, $R = \Theta(\frac{1}{\sigma_{\mathcal{O}}}) = \omega(\frac{L}{\sqrt{n_s}})$ and the traffics in each cluster is not so aggregated because $\sigma_{\mathcal{O}}$ is relative smaller. In [2], *information highway* is proposed to approach the capacity upper bound for unicast non-clustered network based on percolation theory and we apply this concept to our routing scheme. One can refer to [2] for detailed properties about *information highway*. Briefly, information highway is an infinite large component such that each node in the deployed region can connect it within a hop of length $O(\frac{L\log(n_sp)}{\sqrt{n_sp}})$. Based on Lemma 3.2, we obtain

$$\mathbf{Pr}(\mathcal{N} \geq 1) \geq 1 - 2\exp(-\frac{\pi\tau^2}{4}),$$

where $\mathcal{N}$ represents the number of nodes in a cell. Therefore we can choose a τ to make $\mathbf{Pr}(\mathcal{N} \geq 1) \geq \frac{5}{6}$.

Now we need to construct a multicast spanning tree $\mathcal{T}_j$ for data routing of $\mathcal{C}_j$. We show that the length of the tree is of the same order with EMST, which is crucial to prove that our scheme can achieve the maximized capacity in order sense.

Lemma 5.1: The Prim's algorithm is utilized to construct an $\mathcal{T}_j$ for cluster $\mathcal{C}_j$ and we prove that $|\mathcal{T}_j| \leq 5\sqrt{2p}R$.

Proof: Prim's algorithm: To begin with, each node is a separate part, we iteratively find a shortest edge to compose a lager part until one part is left. Each member in $\mathcal{C}_j$ is confined in a disk of radius R, we utilize a square of edge $2R$ to cover the whole circle. At each i_{th} $(1 \leq j \leq |\mathcal{C}_j|)$ step, there are $|\mathcal{C}_j| + 1 - i$ parts remained. We equally partition the square into $\lfloor\sqrt{|\mathcal{C}_j| + 1 - i}\rfloor^2$ cell with edge length $\frac{2R}{\lfloor\sqrt{|\mathcal{C}_j|+1-i}\rfloor}$, and there exists at least one cell which contains more than 2 parts, which means the shortest edge connecting two parts in i_{th} step is at most $\frac{2\sqrt{2}R}{\lfloor\sqrt{|\mathcal{C}_j|+1-i}\rfloor}$. Therefore, the upper bound of $|\mathcal{T}_j|$ is:

$$|\mathcal{T}_j| \leq \sum_{i=1}^{|\mathcal{C}_j|} \frac{2\sqrt{2}R}{\lfloor\sqrt{|\mathcal{C}_j| + 1 - i}\rfloor} \leq 5\sqrt{2|\mathcal{C}_j|}R \leq 5\sqrt{2p}R.$$

■

Based on the above analysis, a capacity routing scheme is provided in Algorithm 2.

According to Algorithm 2, the average number of nodes a percolated cell has to serve for is $\frac{\kappa}{\delta}$ and the next lemma illustrates the minimal achievable data rate in the uplink and downlink phases.

Lemma 5.2: In the uplink and downlink phase, a data rate of $\Omega(\log^{-2}(n_sp))$ can be sustained for each transmission.

Proof: One can refer to [22] for the technical proof. ■

Then we begin to analyze the second phase, although the highway can be virtually considered a wired network with bandwidth $\Theta(W)$, each path must help relay data for considerable clusters. Therefore the allocated radio resources for a cluster is limited. In the following part, we will study the maximum number of clusters a percolated cell serves for.

Lemma 5.3: Each percolated cell cannot relay data for cluster with head at a distance of $\sqrt{2}(R + \frac{2\kappa\log(n_sp)\tau L}{\sqrt{n_sp}})$ away from the cell, therefore $\mathcal{R} \leq \sqrt{2}(R + \frac{2\kappa\log(n_sp))\tau L}{\sqrt{n_sp}})$.

Algorithm 2 Capacity Achieving Scheme for UCRM

1: **Access Point Mapping**: Establish mappings $\mathcal{F}_h(X)$, $\mathcal{F}_v(X)$ for each node X. Horizontally divide the $L \times L$ square into slices of size $L \times (\kappa\log(n_sp)\frac{L}{\sqrt{n_sp}} - \epsilon_L)$. Then there are at least $\delta\log(n_sp)$ paths in each slice. Denote each path in the slice as $path_i(1 \leq i \leq \delta\log(n_sp))$. We further divide each slice into $\delta\log(n_sp)$ sub-slice of size $L \times (\frac{\kappa L}{\delta\sqrt{n_s}})$ each. If node X is in the i_{th} sub-slice, $\mathcal{F}_h(X)$ denotes the percolated cell on $path_i$ with the same ordinate. And mapping $\mathcal{F}_v(X)$ is the dual of previous one just by applying the above algorithm into vertical path.

2: **Medium Access in Highway**: Therefore there exists a spatial and temporal accessing policy such that each representing node can deliver $O(W)$ bits to its adjacent representing node. Such a property is proved in [2] when TDMA is utilized.

3: **Routing Protocol**: A multicast spanning tree $\mathcal{T}_j$ is constructed in Lemma 5.1. Each time slot is divided into 3 mini-slot and corresponds to 3 phases as in Figure 3. For each branch in $\mathcal{T}_j$ linking nodes X_1 and X_2, the 3 phases are as follows:
- **Uplink**: X_1 drains its data to $\mathcal{F}_h(X_1)$.
- **Highway**: This phase corresponds to step 2 and utilize multihop transmission along the horizontal path from $\mathcal{F}_h(X_1)$ to the intersection with the vertical path in which $\mathcal{F}_v(X_2)$ resides then forward to $\mathcal{F}_v(X_2)$.
- **Downlink**: X_2 downloads the data from $\mathcal{F}_v(X_2)$.

Proof: Each path is constrained within a strip of width $\kappa\log(n_sp)\frac{L}{\sqrt{n_sp}}$. Recall that the radius of the disk is R, then we know the farthest cell that could be used is the cell of distance $\mathcal{R} \leq \sqrt{2}(R + \frac{(1+\kappa\log(n_sp))\tau L}{\sqrt{n_sp}}) \leq \sqrt{2}(R + \frac{2\kappa\log(n_sp))\tau L}{\sqrt{n_sp}})$ away from the kernel. ■

For each branch in the spanning tree $\mathcal{T}_j$, it is regarded as a SD pair linking two nodes. The length of it determines how many hops should be used on the highway.

Lemma 5.4: Assume that the length of a SD pair is ℓ, then the number of hops required is $\frac{c_3}{\tau}\left(\frac{\sqrt{2n_sp}\ell}{L} + 4\kappa\log(n_sp)\right)$, where c_3 is a constant.

Proof: One can refer to [22] for the technical proof. ■

Lemma 5.5: Assume that cell s is within a distance $\mathcal{R}$ from $\mathcal{C}_j$'s head, then if $R \geq \frac{2\kappa\log(n_sp)L}{5\sqrt{n_s}}$, the probability $\mathcal{P}$ that s is utilized to transmit for cluster $\mathcal{C}_j$ is upper bounded as:

$$\mathcal{P} \leq \frac{5c_3\delta\tau}{\kappa}\frac{L\sqrt{|\mathcal{C}_j|}}{\sqrt{n_sp}R}. \quad (7)$$

Proof: One can refer to [22] for the technical proof. ■

For each cluster $\mathcal{C}_j$, $|\mathcal{C}_j|$ clients are randomly and independently distributed in a disk of radius R. And each SD pair generated in the Prim's Algorithm is also a random process. Then we will apply Vapnik-Chervonenkis theorem to prove our results and the VC-dimension of a multicast spanning tree is $O(\log p)$ according to [4].

Theorem 5.1: (VC-Theorem): If $\mathcal{S}$ is a set of finite VC-dimension VC-d($\mathcal{S}$), and $\{X_i | i = 1, 2 \cdots, N\}$ is a sequence of i.i.d. random variables with common probability distribution P, then for every $\epsilon, \delta > 0$,

$$\mathbf{Pr}\left(\sup_{A \in \mathcal{S}} |\frac{\sum_{i=1}^{N} I(X_i \in A)}{N} - P(A)| \le \epsilon\right) > 1 - \delta$$

$$\text{when } N \ge \max\left\{\frac{8\text{VC-}d(\mathcal{S})}{\epsilon}\log\frac{13}{\epsilon}, \frac{4}{\epsilon}\log\frac{2}{\delta}\right\}.$$

Theorem 5.2: As to the n_s clusters, we can prove that there exists a sequence $\delta(n) \to 0$, such that the following inequality should be satisfied based on *VC-Theorem*.

$$\mathbf{Pr}\left(\sup_{s \in \mathcal{O}}\left(\mathcal{FL}(s) \le \frac{20\pi c_3 \delta\tau}{\kappa}\frac{R\sqrt{n_s}}{L}\right)\right) \ge 1 - \delta(n),$$

where $\mathcal{FL}(s)$ is the number of flows using s.

Proof: Based on VC-theorem, we can obtain that for each cell s and the whole set of the cell $\mathcal{O}$:

$$\mathbf{Pr}\left(\sup_{s \in \mathcal{O}} |\frac{\mathcal{FL}(s)}{N} - \mathcal{P}| \le \epsilon(n)\right) > 1 - \delta(n)$$

$$\text{when } N \ge \max\left\{\frac{8d}{\epsilon(n)}\log\frac{13}{\epsilon(n)}, \frac{4}{\epsilon(n)}\log\frac{2}{\delta(n)}\right\},$$

where $d = O(\log p)$ is the VC-dimension. Then substitute Eqn. (7) into it, we obtain

$$\mathbf{Pr}\left(\sup_{s \in \mathcal{O}} \frac{\mathcal{FL}(s)}{N} \le \frac{10c_3\delta\tau L}{\kappa\sqrt{n_s}R} + \epsilon(n)\right) > 1 - \delta(n). \quad (8)$$

Now let $\epsilon(n) = \frac{10c_3\delta\tau L}{\kappa\sqrt{n_s}R}$ and $\delta(n) = \frac{2}{n}$ and when

$$N \ge \max\left\{\frac{8d}{\epsilon(n)}\log\frac{13}{\epsilon(n)}, \frac{4}{\epsilon(n)}\log\frac{2}{\delta(n)}\right\} = \frac{4\kappa R\sqrt{n_s}\log n}{5c_3\delta\tau L}, \quad (9)$$

Eqn. (8) is satisfied. Applying the same technique as Lemma 3.2, the number of clusters within the disk $N \in (\frac{\pi n_s R^2}{2L^2}, \frac{2\pi n_s R^2}{L^2})$ if $R = \omega(\frac{L}{\sqrt{n_s}})$. Therefore when $N \ge \frac{\pi n_s R^2}{2L^2} \ge \frac{4\kappa R\sqrt{n_s}\log n}{5c_3\delta\tau L}$, namely $R \ge \frac{8\kappa k L\log n}{5\pi c_3\delta\tau\sqrt{n_s}}$, Eqn. (9) is satisfied. Substitute $N \le \frac{2\pi n_s R^2}{L^2}$ into Eqn. (8), we complete the proof. ■

The above constraint of R means that only when R is sufficient large, the number of flows over a certain cell s can be upper bounded. Theorem 5.2 gives us an upper bound of the number of flows utilizing a cell, and the achievable data rate λ is just W over $\mathcal{FL}(s)$, such that

$$\lambda^h \ge \frac{\kappa L W}{40\pi c_3\delta\tau\sqrt{n_s}R} = \Omega(\frac{L}{\sqrt{n_s}R}W).$$

Recall Lemma 5.2, we know the bottleneck is due to data delivery on the highway when $R \ge \Omega(\frac{L\log n}{\sqrt{n_s}})$ and let λ^u, λ^d denote the data rate of the uplink and downlink, respectively. And $\lambda = \min\{\lambda^u, \lambda^d, \lambda^h\} = \lambda^h$.

Theorem 5.3: The achievable per-cluster capacity in uniform cluster random model $\lambda = \Omega(\frac{L}{\sqrt{n_s}R}W)$. Therefore we know the maximized capacity is achievable in order sense.

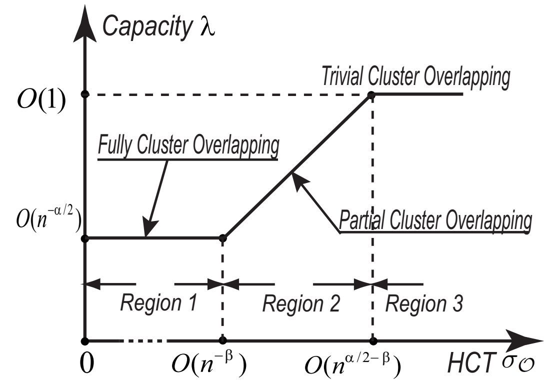

Fig. 4: Relationship between capacity λ and $\sigma_{\mathcal{O}}$ (log scale). In region 2, the network capacity increases with $\sigma_{\mathcal{O}}$. $\sigma_{\mathcal{O}} = 0$ means a non-heterogeneous network with the same achievable capacity as prior works.

VI. Discussion

Our results can generalize various results obtained in non-heterogeneous network like [2], [4], [6]–[8]. Here we highlight two features in our clustered network, *heterogeneous cluster traffic* and *heterogeneous cluster size*, both of which are characterized in realistic multicast network. Because only HCT increases network capacity, we only discuss the impact of HCT for space limitation here.

HCT characterizes a network property that clients of each multicast session are not uniformly distributed across the network. In many cases, the packets are more likely to be delivered to its adjacent nodes and previous works are insufficient to estimate the achievable capacity. Indeed, Eqn. (4) is a precise formula for the achievable capacity. However, we cannot decide whether HCT increases the network capacity because there is no criteria to judge the level of heterogeneity for a dispersion density function. By introducing the variable *distribution variance* $\sigma_{\mathcal{O}}$, we offer a quantitative description of the extent of heterogeneous traffics. We find that if the traffic of a cluster is not uniformly disseminated across the region as is assumed in prior works, the network capacity λ increases.

Figure 4 illustrates the relationship between λ and $\sigma_{\mathcal{O}}$ derived from Eqn. (6). In region 1, HCT is relative slight and each cluster is fully overlapped with other clusters. Fully overlapping indicates that in each small region $\mathcal{O}' \subseteq \mathcal{O}$ with area $\Theta(L^2/n)$, there are $\Theta(n_s) = \Theta(n^{\alpha})$ clusters whose members have approximately equal probability to reside in $\mathcal{O}'$. Therefore it resembles much alike the uniform distributed cases and the network capacity is not improved. In region 2, HCT begins to influence the network performance by increasing the maximized capacity. Here clusters are partially overlapped, it indicates that in each small region $\mathcal{O}' \subseteq \mathcal{O}$ with area $\Theta(L^2/n)$, there are nearly $\Theta(n^{\theta})$ clusters whose members have approximately equal probability to reside in $\mathcal{O}'$. In this case, $\theta < \alpha$ and θ is close related to σ. And the implicit reason for an increased capacity is that each relay only needs to deliver packets for smaller portion of the clusters compared to previous one. In region 3, trivial overlapping means in each small region $\mathcal{O}' \subseteq \mathcal{O}$ with area $\Theta(L^2/n)$, there are at most $\Theta(1)$ clusters whose members have approximately equal probability to reside in $\mathcal{O}'$. Therefore each relay only needs to deliver packets for a constant number of clusters and the achievable capacity tends to be $O(W)$.

VII. Conclusion

In this paper, we study the effect of heterogeneity on the asymptotic multicast capacity in clustered network. Our contributions are mainly divided into two parts. First, we find that heterogeneous cluster traffic increases the achievable capacity for all the clusters. Through analyzing the fundamental constraints of wireless network from global and local aspects, a quantitative prospective is provided between network capacity λ and *distribution variance* $\sigma_{\mathcal{O}}$, which is first utilized for describing heterogeneity in the literature. As for heterogeneous cluster size, we analyze its effect in *uniform cluster random model*, which is the optimal network layout given a fixed $\sigma_{\mathcal{O}}$. we find it cannot increase λ but increase the achievable capacity for small clusters nonetheless under our framework.

VIII. Acknowledgment

This work is supported by National Fundamental research grant (2011CB302701, 2010CB731803, 2009CB3020402), NSF China (No. 60702046, 60832005, 60972050, 60632040); China Ministry of Education (No. 20070248095); China Ministry of Education Fok Ying Tung Fund (No. 122002); Qualcomm Research Grant; China International Science and Technology Cooperation Programm (No. 2008DFA11630); PUJIANG Talents (08PJ14067); Shanghai Innovation Key Project (08511500400); National Key Project of China (2009ZX03003-006-03, 2009ZX03002-003, 2009ZX03002-005,2010ZX03003-001-01); National High tech grant of China (2009AA01Z248, 2009AA1Z249, 2009AA011802); Shanghai Jiao Tong University Undergraduate Participation in Research Program.

Appendix A
Proof of Theorem 3.2

First we divide n_s clusters into two sets $\mathbb{S}_1$, $\mathbb{S}_2$ as:

$$\mathbb{S}_1 = \{\mathcal{C}_j | \text{Dense Region } S_{1,j} \neq \emptyset\},$$
$$\mathbb{S}_2 = \{\mathcal{C}_j | \text{Dense Region } S_{1,j} = \emptyset\}.$$

Then based on eqn. (3), we obtain that if $\mathbb{S}_1 \neq \emptyset$, there exists a constant $c_4 > 0$, such that

$$\sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T} h^b_{\mathcal{C}_j} \geq \frac{c_4}{r}\sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T} |EMST(r-\mathcal{C}_j)| \geq \frac{c_4 T}{r}\left(\sum_{j=1}^{n_s}\lambda_j\sqrt{|\mathcal{C}_j|}\right)\psi\left(\phi(\cdot),\mathbb{S}_1\right), \quad (10)$$

where

$$\psi\left(\phi(\cdot),\mathbb{S}_1\right) = \min_{\mathcal{C}_j\in\mathbb{S}_1}\left\{\frac{|EMST(r-\mathcal{C}_j)|}{\sqrt{|\mathcal{C}_j|}}\right\}.$$

Then substitute Eqn. (2) into Eqn. (10), we can obtain

$$\sum_{j=1}^{n_s}\lambda_j\sqrt{|\mathcal{C}_j|} \leq \frac{r}{c_4 T\psi\left(\phi(\cdot),\mathbb{S}_1\right)}\sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T} h^b_{\mathcal{C}_j} \leq \frac{4\sqrt{2}WL^2}{\Delta\sqrt{c_0}c_4 r\psi\left(\phi(\cdot),\mathbb{S}_1\right)} \quad (11)$$

Else if $\mathbb{S}_1 = \emptyset$:

$$\sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T} h^b_{\mathcal{C}_j} \geq \frac{c_4 T}{r}\int_{\mathcal{O}}\sqrt{\phi(\xi)}d\xi\left(\sum_{j=1}^{n_s}\lambda_j\sqrt{|\mathcal{C}_j|}\right) \quad (12)$$

Substitute Eqn. (2) into (12), we obtain

$$\sum_{j=1}^{n_s}\lambda_j\sqrt{|\mathcal{C}_j|} \leq \frac{4WL\sqrt{2n_s p}}{\Delta c'\sqrt{c_0}c_4\int_{\mathcal{O}}\sqrt{\phi(\xi)}d\xi} \quad (13)$$

During the above derivation, $r \geq c'\frac{L}{\sqrt{n_s p}}$ is a necessary condition to guarantee network connectivity, where $c' > 0$ is a constant. Compare the results under the two cases, the only thing required to do is to prove that there exists a constant $c_5 > 0$ such that

$$\frac{L\sqrt{n_s p}}{c'\int_{\mathcal{O}}\sqrt{\phi(\xi)}d\xi} \geq \frac{c_5 L^2}{r\psi\left(\phi(\cdot),\mathbb{S}_1\right)}.$$

Recall Lemma 3.4, it is equivalent to prove

$$c'c_5\int_{\mathcal{O}}\sqrt{\frac{\phi(\xi)}{n_s p/L^2}}d\xi \leq c_2\widetilde{\rho}_j^{\,2} + r\int_{S_{2,j}}\sqrt{\phi(\xi)}d\xi.$$

Note that $r \geq c'L/\sqrt{n_s p}$ and $\phi(\xi) \leq \frac{2n_s p}{L^2} = \Theta(n_s p/L^2)$, such that there exists a constant c_5 that can meet the above inequality and c satisfies Eqn. (4).

Appendix B
Proof of Theorem 4.2

Let $\sigma_{\mathcal{O}} = \frac{1}{\pi R^2} - \frac{1}{L^2}$. To prove that $\phi'(\cdot)$ achieves the minimal value of $\Upsilon(\phi(\cdot))$, we first transform the conditions and objective functions into Riemann sum as follows.

$$\begin{cases} \sum_{i=1}^m \phi_i^m = \frac{1}{\Delta S} \\ \sum_{i=1}^m (\phi_i^m)^2 \leq \frac{1}{\pi R^2\Delta S} \\ \Upsilon(\phi(\cdot)) = \Delta S\left(\sum_{i=1}^m\sqrt{\phi_i^m}\right) \end{cases}$$

During the above derivation, we equally divide $\mathcal{O}$ into m cells of area $\Delta S = \frac{L^2}{m}$ each. ϕ_i^m is the chosen point in cell i to approximate the value of $\phi(\xi, k_i)$ in cell i. It can be verified that when $m \to \infty$, the Riemann sum equals the integrations. We further assume that $\phi_i^m \leq \phi_j^m$ iff $i \geq j$.

In the following part, we utilize $\Phi^m = (\phi_1^m, \phi_2^m, \ldots \phi_{m-1}^m, \phi_{m-1}^m)$ during the proof and the optimal $(\Phi^m)^u = ((\phi_1^m)^u, (\phi_2^m)^u, \ldots (\phi_{m-1}^m)^u, (\phi_{m-1}^m)^u)$ that can minimize $\Upsilon(\phi(\cdot))$ is proved to be

$$(\phi_i^m)^u = \begin{cases} \frac{1}{\pi R^2} & 1 \leq i \leq m_c \\ \frac{1}{\Delta S} - \frac{m_c}{\pi R^2} & i = m_c + 1 \\ 0 & \text{otherwise} \end{cases}$$

where $m_c = \lfloor \frac{\pi R^2}{\Delta S} \rfloor$. To obtain this result, the proposed Algorithm 3 can convert arbitrary Φ^m to $(\Phi^m)^u$ with finite step and in each step, it decreases $\Upsilon(\phi(\cdot))$.

Algorithm 3 Conversion from Arbitrary Φ^m to Optimal $(\Phi^m)^u$
Input: Φ^m Output: $\widetilde{\Phi^m} = (\Phi^m)^u$

Require: Φ^m satisfies all the conditions listed above
Ensure: $\widetilde{\Phi^m}$ can be generated
1: $\widetilde{\Phi^m} \leftarrow \Phi^m$
2: Find $\widetilde{\phi_i^m}$, $\widetilde{\phi_j^m}$, $\widetilde{\phi_l^m}$ in $\widetilde{\Phi^m}$ satisfying the conditions below:
- For $k \leq i$, $\phi_k^m > \frac{1}{\pi R^2}$ and for $k > i$, $\phi_k^m \leq \frac{1}{\pi R^2}$
- For $k < j$, $\phi_k^m \geq \frac{1}{\pi R^2}$ and for $k \geq j$, $\phi_k^m < \frac{1}{\pi R^2}$
- For $k \leq l$, $\phi_k^m > 0$ and for $i > l$, $\phi_k^m = 0$

If ϕ_i^m can not be found, finish the program.
3: Find $\max\{\Delta_i\}$, $\max\{\Delta_l\}$ satisfying the conditions below:

$$\phi_i^m \geq \frac{1}{\pi R^2} + \Delta_i \quad \phi_j^n \leq \frac{1}{\pi R^2} - \Delta_i - \Delta_l \quad \phi_l^m \geq \Delta_l \tag{14}$$

$$(\phi_i^m - \Delta_i)^2 + (\phi_j^m + \Delta_i + \Delta_l)^2 + (\phi_l^m - \Delta_l)^2 \leq (\phi_i^m)^2 + (\phi_j^m)^2 + (\phi_l^m)^2 \tag{15}$$

4: $\widetilde{\phi_i^m} = \widetilde{\phi_i^m} - \Delta_i$, $\widetilde{\phi_j^m} = \widetilde{\phi_j^m} + \Delta_i + \Delta_l$, $\widetilde{\phi_l^m} = \widetilde{\phi_l^m} - \Delta_l$. Go back to step 2.

Denote ρ as: $\rho = \frac{\phi_i^m - \phi_j^m}{\phi_j^m - \phi_l^m}$ to simplify our analysis. Now we will prove that step 4 in Algorithm 3 decreases $\Upsilon(\phi(\cdot))$, which is equivalent to prove

$$\sqrt{\phi_i^m - \Delta_i} + \sqrt{\phi_j^m + \Delta_i + \Delta_l} + \sqrt{\phi_l^m - \Delta_l} \leq \sqrt{\phi_i^m} + \sqrt{\phi_j^m} + \sqrt{\phi_l^m}. \tag{16}$$

According to step 3, we can obtain

$$\Delta_l = \rho\Delta_i - \frac{\Delta_i^2 + \Delta_l^2 + \Delta_i\Delta_l}{\phi_j^m - \phi_l^m}. \tag{17}$$

Rewrite Eqn. (16) and substitute Eqn. (17) into it, we can obtain

$$\begin{aligned}
&\sqrt{\phi_i^m - \Delta_i} + \sqrt{\phi_j^m + \Delta_i + \Delta_l} + \sqrt{\phi_l^m - \Delta_l} \\
&-(\sqrt{\phi_i^m} + \sqrt{\phi_j^m} + \sqrt{\phi_l^m}) \\
\leq& -\frac{\Delta_i}{2\sqrt{\phi_i^m}} + \frac{\Delta_i + \Delta_l}{2\sqrt{\phi_j^m}} - \frac{\Delta_l}{2\sqrt{\phi_l^m}} + \delta(\Delta_i, \Delta_l) \\
\leq& \frac{\Delta_i}{2\sqrt{\phi_j^m}} \left(\frac{\sqrt{\phi_i^m} - \sqrt{\phi_j^m}}{\sqrt{\phi_i^m}} - \frac{\sqrt{\phi_j^m} - \sqrt{\phi_l^m}}{\sqrt{\phi_l^m}} \frac{\Delta_l}{\Delta_i} \right) + \delta(\Delta_i, \Delta_l) \\
\leq& -\frac{\Delta_i(\sqrt{\phi_i^m} - \sqrt{\phi_j^m})(\sqrt{\phi_i^m} - \sqrt{\phi_l^m})}{2\sqrt{\phi_i^m\phi_j^m\phi_l^m}} + \\
&\frac{\Delta_i^2(1+\rho^2)(\phi_j^m + \phi_l^m)}{2\sqrt{\phi_j^m}\phi_l^m\phi_j^m} \\
\leq& 0.
\end{aligned}$$

During the above derivation, $\delta(\Delta_i, \Delta_l)$ is the sum of the reminder terms and we can prove that $\delta(\Delta_i, \Delta_l) \leq \frac{(\Delta_i + \Delta_l)^2}{2(\phi_j^m)^{1.5}}$. Therefore, Algorithm 3 can reduce $\Upsilon(\phi(\cdot))$ in every iteration until $\widetilde{\Phi^m} = (\Phi^m)^u$. When $m \to \infty$, the Riemann sum becomes integration and due to the non-increasing characteristic of Φ^m, the respect $\phi(\cdot)$ is the same as $\phi^u(\cdot)$.

REFERENCES

[1] P. Gupta, P. R. Kumar. "*The capacity of wireless networks*", in IEEE Trans. on Info. Theory, Vol. 46, No. 2, pp. 388-404.

[2] M. Franceschetti, O. Dousse, D.N.C. Tse, P. Thiran. "*Closing the gap in the capacity of random wireless networks via percolation theory*", in IEEE Trans. on Info. Theory, Vol. 53, No. 3, pp. 1009-1018, 2007.

[3] P. Li, C. Zhang, Y. Fang, "*Capacity and Delay of Hybrid Wireless Broadband Access Networks*," in IEEE JSAC, Vol. 27, No. 2, pp. 117-125, February 2009.

[4] X.-Y. Li, "*Multicast Capacity of Wireless Ad Hoc Networks*", in IEEE/ACM Transaction on Networking, vol. 17, p. 950-961, June 2009.

[5] A. Keshavarz-Haddad, V. Ribeiro, R. Riedi, "*Broadcast capacity in multihop wireless networks,*" in Proc. ACM Mobicom'06, Los Angeles, CA, USA, 2006.

[6] S. Shakkottai, X. Liu, and R. Srikant. "*The multicast capacity of large multihop ad hoc networks*", in Proc. ACM Mobihoc'07, Montreal, Quebec, Canada, 2007.

[7] X.-Y. Li, Y. Liu, S. Li, S. Tang, "*Multicast Capacity of Wireless Ad Hoc Networks Under Gaussian Channel Model*", IEEE/ACM Transactions on Networking (TON), Vol. 18, No. 4, August 2010, Pages 1145-1157.

[8] C. Wang, X.-Y Li, C. Jiang, S. Tang, Y. Liu, J. Zhao. "*Scaling Laws on Multicast Capacity of Large Scale Wireless Networks,*" in Proc. IEEE INFOCOM'09, Rio de Janeiro, Brazil, April 2009.

[9] C. Hu, X. Wang, F. Wu, "*MotionCast: On the Capacity and Delay Tradeoffs,*" in ACM MobiHoc'09, New Orleans, USA, May 2009.

[10] S. Zhou and L. Ying, "*On Delay Constrained Multicast Capacity of Large-Scale Mobile Ad-Hoc Networks*" in Proc. INFOCOM'10 mini-conference, San Diego, CA, USA, March 2010.

[11] X. Lin, Ness B. Shroff, "*The Fundamental Capacity-Delay Tradeoff in Large Mobile Wireless Networks,*" in Proc. IEEE MedHoc'04, Bodrum, Turkey, June 27-30, 2004.

[12] L. Ying, S. Yang R. Srikant. "*Optimal Delay-Throughput Trade-Offs in Mobile Ad Hoc Networks*" in IEEE Tran. on Info. Theory, Vol. 54, No. 9, September, 2008.

[13] U. Niesen, P. Gupta, D. Shah, "*The Multicast Capacity Region of Large Wireless Networks*" in Proc. of IEEE INFOCOM'09, Rio de Janeiro, Brazil, April 2009.

[14] C. Hu, X. Wang, D. Nie, J. Zhao, "*Multicast Scaling Laws with Hierarchical Cooperation,*" in Proc. of IEEE INFOCOM'10, San Deigo, CA, USA, 2010.

[15] G. Alfano, M. Garetto, E. Leonardi, "*Capacity Scaling of Wireless Networks with Inhomogeneous Node Density: Upper Bounds*", in IEEE JSAC, Vol. 27, No. 7, September 2009.

[16] G. Alfano, M. Garetto, E. Leonardi, "*Capacity Scaling of Wireless Networks with Inhomogeneous Node Density: Lower Bounds,*" in Proc. IEEE INFOCOM'09, Rio de Janeiro, Brazil, April 2009.

[17] Z. Wang , H. R. Sadjadpour and J. J. Garcia-Luna-Aceves, "*A Unifying Perspective on the Capacity of Wireless Ad Hoc Networks,*"in Proc. IEEE INFOCOM'08, Phoenix, AZ, USA, April 2008.

[18] M. Ji, Z. Wang, H. Sadjadpour, J. J. Garcia-Luna-Aceves, "*The Capacity of Ad Hoc Networks with Heterogeneous Traffic Using Cooperation*", in Proc. IEEE INFOCOM'10, San Diego, CA, USA, March, 2010.

[19] P. Li and Y. Fang, "*The Capacity of Heterogeneous Wireless Networks*,"in Proc. IEEE INFOCOM'10, San Diego, CA, USA, March, 2010.

[20] D.-Z. Du and F.-K. Hwang."*A proof of the gilbert-pollak conjecture on the steiner ratio*", in Algorithmica 7((2,3)):121-135, 1992.

[21] J. M. Steele. "*Growth rates of euclidean minimal spanning trees with power weighted edges.*" in The Annals of Probability, 16(4):1767-1787, Oct 1988.

[22] Q. Peng, X. Wang, H. Tang, "*Heterogeneity Increases Multicast Capacity in Clustered Network*", Technical report, 2010, available at http://iwct.sjtu.edu.cn/Personal/xwang8/paper/HCN.pdf