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 only increases capacity for small clusters. Our work can generalize various results obtained under non-heterogeneous networks in the literature.**

***Index Terms*—Static Clustered Network, Capacity, Spatial Inhomogeneity, Heterogeneous Cluster.**

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. Later in [2], the capacity result is analyzed under more general fading channel and a similar result is given. Then M. Franceschetti, *et al*. [3] design an optimal routing protocol with capacity achieving $\Theta(W/\sqrt{n})$ via percolation theory.

Multicast network, which generalizes the above unicast network, 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, multicast network can also be viewed as clustered network. However, there are few work concerning the cluster behavior of multicast networks. 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 unicast network, 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

[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 $g(n) = O(f(n))$.

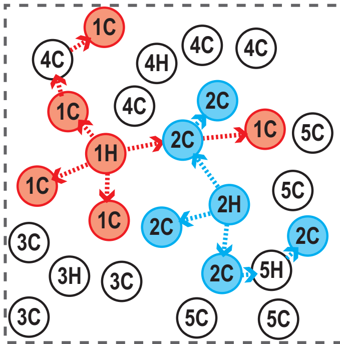

Fig. 1: Demonstration of Network topology. Nodes in the same cluster are labeled with the same number. H and D represent head (source) and clients (destination), respectively.

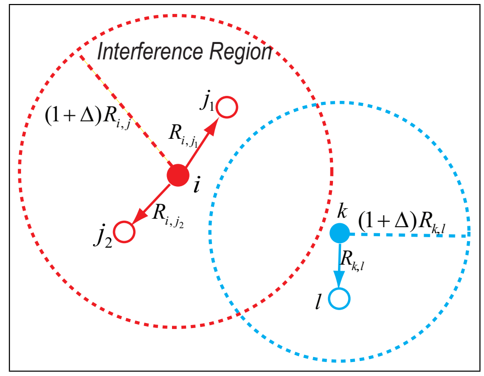

Fig. 2: Demonstration of two successful transmissions.

size on multicast capacity in clustered network?

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. However, smaller clusters can achieve a relative larger capacity because larger clusters can be virtually regarded as helping clusters for smaller ones. This generous help does not influence their own capacity in order sense and can be viewed as "tradeoff-less benefit".

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 $n_s = n^\alpha$ $(0 \le \alpha \le 1)$ clusters distributed over a 2-dimensional torus region $\mathcal{O}$ of edge[2] $L = n^\beta (0 \le \beta \le \alpha/2)$. We first specify a homogeneous poisson process (HPP) to generate cluster head v_j, whose position is denoted by k_j for cluster $C_j (1 \le j \le n_s)$. 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 $|\mathcal{C}_j| \le p = n^{1-\alpha}$ is the expected size of the cluster and $\phi(\cdot)$ is the dispersion density function. In order to explain the effect of cluster size, we classify these n_s clusters into k super clusters ($\mathcal{SC}$) based on their cluster size. For each cluster $\mathcal{C}_j \in \mathcal{SC}_i$ $(1 \le i \le k)$, its cluster size $|\mathcal{C}_j| = \Theta(n^{1-\alpha_i})$, where α_i is an increasing sequence over i and the clusters in $\mathcal{SC}_1$ possess the largest size. In addition, some further assumptions are assumed below.

1) $\alpha = \min\{\alpha_1, \alpha_2, \dots \alpha_{k-1}, \alpha_k\} = \alpha_1$.
2) There are $c_0 n_s$ $(0 < c_0 < 1)$ clusters in $\mathcal{SC}_1$, and the other $(1-c_0)n_s$ clusters are randomly allocated to $\mathcal{SC}_i$ $(2 \le i \le k)$. It indicates that the number of clusters with size $\Theta(p)$ is the same order with the total number of clusters.

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 $|\mathcal{C}_j| = \Theta(n^{1-\alpha_i})$ if $\mathcal{C}_j \in \mathcal{SC}_i$. In addition, there are $\Theta(n_s/k)$ clusters in $\mathcal{SC}_i$ for $2 \le i \le k$ applying Chernoff bound. In order to simplify our analysis, we take both $|\mathcal{C}_j|$ and $n^{1-\alpha_i}$ as the cluster size and n_s/k as the number of clusters in $\mathcal{SC}_i$ for $2 \le i \le k$. Such simplification does not influence our results in order sense. Figure 1 is an example of the network topology.

Variance is usually utilized to describe the fluctuation level of a signal around its average value. We employ ac variable *distribution variance* $\sigma_{\mathcal{O}}$ to depict HCT, which is defined as:

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

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

[2]A cluster dense regime is assumed, which means the distance between adjacent clusters $\frac{L}{\sqrt{n_s}}$ tends to 0 when n approaches infinity.

traffic, $\phi(\xi) \equiv \frac{1}{L^2}$ therefore $\sigma_{\mathcal{O}} = 0$ and larger heterogeneity results in larger $\sigma_{\mathcal{O}}$. 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], which follows:

$$\text{Capacity} = W \log_2(1 + \text{SINR}).$$

The protocol model is identical to this model when analyzing the scaling laws, which is verified 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. Note that in broadcast cases, the interference radius is defined as $(1+\Delta)$ times the length between the furthest nodes to the source. Figure 2 illustrates two concurrent transmissions and $R_{i,j} = \max\{R_{i,j_1}, R_{i,j_2}\}$.

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. Let $B_j(t)$ denote the number of data units already generated in $\mathcal{C}_j$ which has not yet been delivered to all of its members at time t. Then λ must guarantee a non-backlog network, which means $\lim_{t\to\infty} \sup_{1\leq j\leq n_s} B_j(t) < \infty$.

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. $\delta_{h,b}$ is the number of nodes that can overhear a packet during a transmission of bit b in its h_{th} 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. However, cluster $\mathcal{C}_j$ can be allocated $\Omega(\sqrt{p/|\mathcal{C}_j|}\lambda)$ data rate in UCRM therefore smaller cluster can be assigned larger capacity than λ. And such property is not limited to this special model.

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. First, we will investigate some restrictions from global aspects.

A. Restrictions From Global Aspects

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.

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

Proof: When T is sufficient large, the total number of bits communicated from head to its clients in cluster $\mathcal{C}_j$ is $\lambda_j T$. Assume two SD pairs $X_i \to X_k$ and $X_j \to X_l$ are active in the given time slot, then according to the transmission protocol model,

$$|X_k - X_l| \geq \frac{\Delta}{2}\left(|X_i - X_k| + |X_j - X_l|\right),$$

which is derived in [1]. Thus disks of radius $\frac{\Delta}{2}$ times the transmission radius centered at the transmitter can be viewed as a "exclusion region" that rejects transmitters from other data flows. Such property also holds for broadcast that the transmission range is defined as the furthest node that can receive the packets. Let S_b^h be the overlapped area between the "exclusion region" of bit b's h_{th} hop and the deployed region $\mathcal{O}$, then

$$S_b^h \geq \frac{\pi}{4}(\frac{\Delta \ell_b^h}{2})^2 = \frac{\pi \Delta^2 (\ell_b^h)^2}{16}.$$

Therefore, radio resources can be viewed as the limited bandwidth W times the simulation time T and the network area L^2, such that the following inequality is satisfied:

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

■

The above inequality is a basic requirement for multihop transmission fashion and the cooperative MIMO as in [14] is not considered here. For each wireless nodes cannot send and receive concurrently, the number of wireless transceivers can also be regarded as wireless resources. The following lemma illustrates this property, which is also suitable for broadcast.

Lemma 3.2: **Constraint of Half Duplex**: The following trivial inequality must be satisfied for sufficient large simulation time T.

$$\sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T}\sum_{h=1}^{h_{\mathcal{C}_j}^b} \delta_{h,b} \leq \sum_{j=1}^{n_s} |\mathcal{C}_j| TW \leq n_s p TW. \quad (2)$$

Lemma 3.3: Given the transmission range $r = \Omega(\frac{L}{\sqrt{n_s p}})$, $\delta_{h,b}$ is tight bounded as $\delta_{h,b} = \Theta(\frac{n_s p r^2}{L^2})$.

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

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

$$\mathbf{Pr}(\delta_{h,b} > n_0) \leq e^{-\mu} \sum_{i=n_0+1}^{\infty} \frac{\mu^i}{i!} \leq \frac{e^{(\theta-1)\mu}\mu^{n_0+1}}{(n_0+1)!}.$$

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

We have completed our analysis of the network from a global aspect. However, some local perspectives inside a cluster are also deterministic of the capacity upper bound and we will study them in the following part.

B. Restrictions from Local Aspects

HCT and multicast traffic pattern are inner properties inside a cluster. And we will investigate them from a clustered perspective in this section.

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 Theorem 1 in [21] as follows.

Lemma 3.4: 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.

Our case corresponds to $d = 2$ and such that

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

The order of $|EMST(\mathcal{C}_j)|$ constitutes two terms, $\sqrt{|\mathcal{C}_j|}$ and $\int_{\mathcal{O}} \sqrt{\phi(\xi)}d\xi$. Since larger transmission radius covers more than one destinations a time, we must derive the relationship between the transmission range r and the minimal overall multihop transmission length. *ρ-simplified cluster* $\rho - \mathcal{C}_j$ is constructive for deriving the above relationships. 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. For the point intensity after the thinning process determines $|EMST(\rho - \mathcal{C}_j)|$ according to lemma 3.4, 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 $\phi'(\xi) \geq \Theta(\phi(\xi))$.

Lemma 3.5: There exists a constant $c_2 > 0$ such that $|EMST(\rho - \mathcal{C}_j)|$ can be 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 3.4, thus there exists a constant $c' > 0$, such that

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

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

Lemma 3.5 is a cornerstone for us to derive the lower bound of the overall multihop length for packet delivery in cluster $\mathcal{C}_j$. However in multihop scenario, nodes from other clusters can act as relays to help forward information. In this case, $EMST(\rho - \mathcal{C}_j)$ sometimes is not the minimal transmission length. However, the following lemma in [20] tells us that $|EMST(\rho - \mathcal{C}_j)|$ is at most $\frac{\sqrt{3}}{2}$ times larger than the minimal value of $\sum_{h=1}^{h^b_{\mathcal{C}_j}} \ell^h_b$ when relays are utilized.

Lemma 3.6: 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 a lower bound of $\sum_{h=1}^{h^b_{\mathcal{C}_j}} \ell^h_b$ is trivial given the transmission range r.

Lemma 3.7: The overall transmission length in cluster $\mathcal{C}_j$ is no smaller than a constant times $|EMST(r - \mathcal{C}_j)|$ when the transmission range r is given, such that

$$\sum_{h=1}^{h^b_{\mathcal{C}_j}} \ell^h_b \geq \Theta(|EMST(r - \mathcal{C}_j)|).$$

Proof: The transmission range of each hop is $\Theta(r)$, then

$$\sum_{h=1}^{h^b_{\mathcal{C}_j}} \ell^h_b \geq \frac{\sqrt{3}}{2}|EMST(r - \mathcal{C}_j)| = \Theta(|EMST(r - \mathcal{C}_j)|),$$

based on Lemma 3.6. ■

C. Upper Bound of Multicast Capacity

In this section, 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.

$$\sum_{j=1}^{n_s} \lambda_j\sqrt{|\mathcal{C}_j|} \leq c\frac{\sqrt{n_s p}WL}{\int_{\mathcal{O}}\sqrt{\phi(\xi)}d\xi}, \tag{3}$$

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:

$$\lambda \leq \min\left\{W, \frac{cWL}{\sqrt{n_s}\int_{\mathcal{O}}\sqrt{\phi(\xi)}d\xi}\right\}. \tag{4}$$

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. (3). ■

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 vsrious 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.

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

$$\lambda \leq \min\left\{O(W), O\left(\frac{\max\{1, L\sigma_{\mathcal{O}}\}W}{\sqrt{n_s}}\right)\right\}. \tag{5}$$

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 E. ■

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 [3], *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. Before we outline what is *information highway*, some useful lemmas are provided.

Lemma 5.1: Let $\mathcal{N}(r)$ denote number of nodes within a disk of radius $r = \Theta(L/\sqrt{n_s p})$. Then if the cluster radius $R = \Omega(L/\sqrt{n_s})$, the following inequality is satisfied.

$$\mathbf{Pr}(\mathcal{N}(r) = k) \leq 2\exp(-\frac{\pi r^2 n_s p}{L^2})\frac{(\frac{\pi r^2 n_s p}{L^2})^k}{k!}.$$

Proof: The proof is presented in Appendix B. ■

Lemma 5.2: There exists a constant τ such that if we equally partition the square $\mathcal{O}$ into cells with side length $\tau L/\sqrt{n_s p}$, the probability that at least one node resides in a cell is larger than $1 - 2\exp(\frac{\pi\tau^2}{4})$.

Proof: Each cell of edge length $\tau L/\sqrt{n_s p}$ contains a disk of radius $\frac{\tau L}{2\sqrt{n_s p}}$ such that the probability that at least one node resides in a cell can be lower bounded as follows:

$$\mathbf{Pr}(\mathcal{N}(r) \geq 1) \geq 1 - 2\exp(-\frac{\pi r^2 n_s p}{L^2}) = 1 - 2\exp(-\frac{\pi\tau^2}{4}).$$

■

Now we outline what is *information highway*: Choose τ large enough so that $1 - 2\exp(-\frac{\pi\tau^2}{4}) > 5/6$ and we equally partition the square $\mathcal{O}$ into cells of edge $\tau L/\sqrt{n_s p}$. Thus there are $\lfloor\frac{\sqrt{n_s p}}{\tau}\rfloor \times \lfloor\frac{\sqrt{n_s p}}{\tau}\rfloor$ cells and each cell in i_{th} row and j_{th} column is denoted by $s_{i,j}$. $s_{i,j}$ is open if it contains at least one node. A horizontal (vertical) cross path is defined as a set of open cells that cross $\mathcal{O}$ from left to right (top to bottom). The gray cell in Figure 3 is example of percolation path. According to Theorem 5 in [3], if the probability that a cell contains at least 1 node is larger than $\frac{5}{6}$, there are w.h.p. $\Theta(n_s p)$ disjoint paths crossing from left to right (top to bottom). A set containing these $\Theta(n_s p)$ horizontal and vertical paths are called *information highway*. Note that the cell composed of the highway is called percolated cell, in which the node is called representing node. The following properties about information highway are proved in [3] via percolation theory.

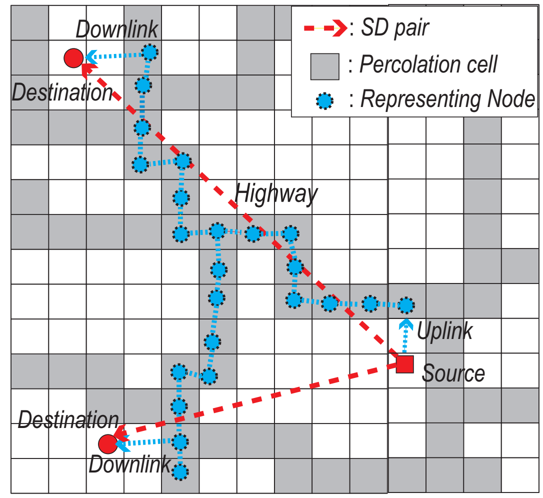

Fig. 3: Demonstration of routing protocol.

- In each horizontal (vertical) rectangular of size[3] $L \times (\kappa\log(n_s p)\frac{L}{\sqrt{n_s p}} - \epsilon_L)$, there are at least $\delta\log(n_s p)$ horizontal (vertical) highway paths w.h.p. It indicates there are $\Theta(\sqrt{n_s p})$ disjoint crossing paths from left to right and top to bottom, respectively. (Theorem 5)
- The length of each crossing path is bounded by $\Theta(L)$.
- The distance between two adjacent horizontal (vertical) path is at most $O(L\log(n_s p)/\sqrt{n_s p})$.
- There exists a spatial and temporal scheme that can achieve $\Theta(W)$ throughput on the highway. It means that each cell composed of the highway can be considered linked by optical wires to its neighbors with bandwidth $\Theta(W)$. (Theorem 3)

Therefore, information highway is an infinite large component such that each node in the deployed region can connect it within a hop of length $O(\frac{L\log(n_s p)}{\sqrt{n_s p}})$. Now we need to construct a multicast spanning tree $\mathcal{T}_j$ for $\mathcal{C}_j$.

Lemma 5.3: The Prim's algorithm is utilized to construct an $\mathcal{T}_j$ for cluster $\mathcal{C}_j$ and we prove that $|\mathcal{T}_j| \leq 5\sqrt{2|\mathcal{C}_j|}R$.

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:

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

■

[3] κ and δ is some constant, $\epsilon_L = o(\log(n_s p)\frac{L}{\sqrt{n_s p}})$ and is to make $\kappa\log(n_s p)\frac{L}{\sqrt{n_s p}} - \epsilon_L$ an integer.

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

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 $L \times (\kappa \log(n_s p)\frac{L}{\sqrt{n_s p}} - \epsilon_L)$. Then there are at least $\delta \log(n_s p)$ paths in each slice. Denote each path in the slice as $path_i (1 \leq i \leq \delta \log(n_s p))$. We further divide each slice into $\delta \log(n_s p)$ 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**: Each representing node can be active for a constant portion of time in a cell partitioned network based on Lemma 6 in [11]. Therefore there exists a spatial and temporal accessing policy such that each representing node can deliver $O(W)$ bits to its adjacent representing node as [3].

3: **Bandwidth Allocation in Highway**: Frequency division multiplexing (FDM) is utilized. Each link with bandwidth W between two percolation cell is equally divided into $\hbar$ (determined in Lemma 5.8) sub-channels and clusters of the same $\mathcal{SC}$ share the same channel.

4: **Routing Protocol**: A multicast spanning tree $\mathcal{T}_j$ is constructed in Lemma 5.3. 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 $\mathcal{F}_v(X_2)$.

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

Proof: According to Algorithm 2, for each node X, both $|\mathcal{F}_h(X) - X|$ and $|\mathcal{F}_v(X) - X|$ is upper bounded by $\kappa \log(n_s p)\frac{L}{\sqrt{n_s p}}$. It means that if we equally divide the region $\mathcal{O}$ into sub-square of size $2\kappa \log(n_s p)\frac{L}{\sqrt{n_s p}} \times 2\kappa \log(n_s p)\frac{L}{\sqrt{n_s p}}$, X and $\mathcal{F}_h(X)$ must reside in the same cell. The same is true for X and $\mathcal{F}_v(X)$. Then according to Lemma 6 in [11], each cell can be allocated a constant time to be active within the mini-slot. According to Lemma 3.3, there are at most $8\kappa^2 \log^2(n_s p)$ nodes in each cell, the same to say there are at most $8\kappa^2 \log^2(n_s p)$ SD pairs. Thus, each SD pairs can be allocated $\frac{1}{O(\log^2(n_s p))} = \Omega(\log^{-2}(n_s p))$ fraction

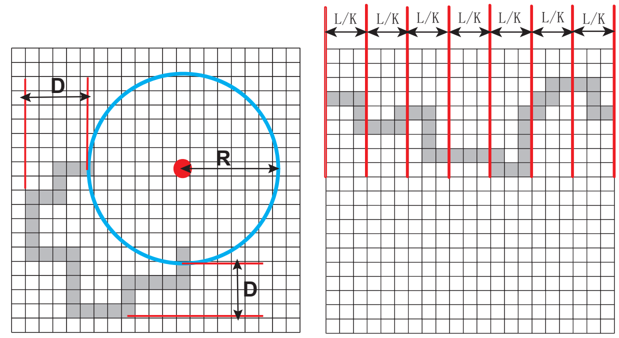

Fig. 4: Demonstration of influential range and a division of a percolation path

of a time slot for transmission. ■

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

Proof: We refer to Figure 4 for the proof. Each path is constrained within a strip of width $\kappa \log(n_s p)\frac{L}{\sqrt{n_s p}}$. Thus D is upper bounded by $(\kappa \log(n_s p)\frac{L}{\sqrt{n_s p}})$. Recall that the radius of the disk is R, then we know the farthest cell that could be used is the black cell, a distance $\mathcal{R} \leq \sqrt{2}(R + \frac{(1+\kappa \log(n_s p))\tau L}{\sqrt{n_s p}}) \leq \sqrt{2}(R + \frac{2\kappa \log(n_s p))\tau L}{\sqrt{n_s p}})$ 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 5.6: Assume that the length of a SD pair is ℓ, then the number of hops required is $\frac{c_3}{\tau}\left(\frac{\sqrt{2n_s p}\ell}{L} + 4\kappa \log(n_s p)\right)$, where c_3 is a constant.

Proof: The proof is presented in Appendix C. ■

Lemma 5.7: Assume that cell s is within a distance $\mathcal{R}$ from $\mathcal{C}_j$'s head, then if $R \geq \frac{2\kappa \log(n_s p)L}{5\sqrt{n_s}}$, the probability $\mathcal{P}$ that s is utilized to transmit for cluster $\mathcal{C}_j$ is upper bounded as:

$$\mathcal{P} \leq \frac{5c_3\delta\tau}{\kappa}\frac{L\sqrt{|\mathcal{C}_j|}}{\sqrt{n_s p}R}. \tag{6}$$

Proof: The proof is presented in Appendix D. ■

We know that $|\mathcal{C}_j| = n^{1-\alpha_i}$ if $\mathcal{C}_j \in \mathcal{SC}_i$. Let $\mathcal{P}_i \leq \frac{5c_3\delta\tau}{\kappa}\frac{L\sqrt{n^{1-\alpha_i}}}{\sqrt{n_s p}R} \leq \frac{5c_3\delta\tau}{\kappa}\frac{L}{\sqrt{n^{\alpha_i}}R}$ denote the probability that a cell is used for $\mathcal{C}_j$ if $\mathcal{C}_j \in \mathcal{SC}_i$ and s is within the influential range $\mathcal{R}$ of $\mathcal{C}_j$'s head. 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 $\epsilon, \delta > 0$,

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

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

Theorem 5.2: For each $\mathcal{SC}_i$ $(1 \leq i \leq k)$, let $k(1) = k$ and $k(i) = 1/c_0$ for $2 \leq i \leq k$, then the following inequality should be satisfied based on *VC-Theorem*.

$$\mathbf{Pr}\left(\sup_{s \in \mathcal{O}}\left(\mathcal{FL}_i(s) \leq \frac{20\pi c_3 \delta\tau}{\kappa} \frac{Rn^{\alpha}}{k(i)L\sqrt{n^{\alpha_i}}}\right)\right) \geq 1 - \delta(n),$$

where $\mathcal{FL}_i(s)$ is the number of flows using s in clusters belonged to $\mathcal{SC}_i$ and $\delta(n)$ is a set of sequence approaching 0 when n goes to infinity. Eqn. (9) should be satisfied when $2 \leq i \leq k$ and Eqn. (10) should be satisfied when $i = 1$.

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

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

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

where $d = O(\log p)$ is the VC-dimension. Then substitute Eqn. (6) into it and note that $|\mathcal{C}_j| = \Theta(n^{1-\alpha_i})$, we obtain

$$\mathbf{Pr}\left(\sup_{s \in \mathcal{O}} \frac{\mathcal{FL}_i(s)}{N} \leq \frac{10c_3\delta\tau L}{\kappa\sqrt{n^{\alpha_i}}R} + \epsilon(n)\right) > 1 - \delta(n). \quad (7)$$

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

$$N \geq \max\left\{\frac{8d}{\epsilon(n)} \log \frac{13}{\epsilon(n)}, \frac{4}{\epsilon(n)} \log \frac{2}{\delta(n)}\right\} = \frac{4\kappa Rn^{\alpha_i/2}\log n}{5c_3\delta\tau L}, \quad (8)$$

Eqn. (7) is satisfied. Applying the same technique as Lemma 3.3, the number of clusters belonged to $\mathcal{SC}_i$ $(2 \leq i \leq k)$ within the disk $N \in (\frac{\pi n_s R^2}{2kL^2}, \frac{2\pi n_s R^2}{kL^2})$ if $R = \omega(\sqrt{\frac{k}{n_s}}L)$. Therefore when $N \geq \frac{\pi n_s R^2}{2kL^2} \geq \frac{4\kappa Rn^{\alpha_i/2}\log n}{5c_3\delta\tau L}$, which means if

$$R \geq \frac{8\kappa kLn^{\alpha_i/2}\log n}{5\pi c_3\delta\tau\sqrt{n_s}} = \frac{8\kappa}{5\pi c_3\delta\tau} kLn^{\alpha_i/2-\alpha}\log n, \quad (9)$$

Eqn. (8) is satisfied. Substitute $N \leq \frac{2\pi n_s R^2}{kL^2}$ into Eqn. (7), we complete the proof when $2 \leq i \leq k$. A similar technique is employed for $\mathcal{SC}_1$ if the following condition is satisfied.

$$R \geq \frac{8\kappa Ln^{\alpha_i/2}\log n}{5\pi c_0 c_3\delta\tau\sqrt{n_s}} = \frac{8\kappa}{5\pi c_0 c_3\delta\tau} \frac{L\log n}{\sqrt{n^{\alpha}}}. \quad (10)$$

■

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. In addition, we find that if the number of flows over s in $\mathcal{SC}_{i'}$ is upper bounded by $\triangle$, $\triangle$ is also a upper bound of number of flows in $\mathcal{SC}_i$ $(i' \leq i \leq k)$ due to α_i is an increasing sequence over i.

Lemma 5.8: Given a cluster radius $R = \omega(L/\sqrt{n_s})$, the data rate that the highway can sustain for $\mathcal{C}_j$ is denoted by λ_j^h. When $R \leq O(\sqrt{\frac{k}{n_s}}L)$,

$$\lambda_j^h = \Omega\left(\frac{L}{R\sqrt{n^{\alpha}}}\right) \quad 1 \leq j \leq k,$$

when $R \geq \omega(\sqrt{\frac{k}{n_s}}L)$,

$$\lambda_j^h = \begin{cases} \Omega(\frac{Ln^{\alpha_i/2}}{Rn^{\alpha}}) & \mathcal{C}_j \in \mathcal{SC}_i, \quad 1 < i \leq k', \\ \Omega(\frac{Ln^{\alpha_{k'}/2}}{Rn^{\alpha}}) & \mathcal{C}_j \in \mathcal{SC}_i, \quad k' < i < k. \end{cases} \quad (11)$$

Proof: In case of $R \leq O(\sqrt{\frac{k}{n_s}}L)$, there are at most $O(\frac{n_s R^2}{L^2}) \leq O(k)$ clusters within a disk of radius R therefore only a small portion of super clusters may have members inside the disk. The number of flows through each node is upper bounded by $\Theta(\frac{R\sqrt{n_s}}{L})$ w.h.p according to Theorem 5.2. Therefore $\hbar = O(\frac{R\sqrt{n_s}}{L})$ in our FDM network and the highway can sustain a data rate of $W/\hbar = \Omega(\frac{L}{R\sqrt{n_s}})$.

When $R \geq \omega(\sqrt{\frac{k}{n_s}}L)$, the number of clusters belonged to $\mathcal{SC}_i$ $(2 \leq i \leq k)$ is $\Theta(\frac{n_s R^2}{kL^2})$ and $\hbar = k$ which means each $\mathcal{SC}_i$ $(2 \leq i \leq k)$ is allocated $(1-c_0)W/k$ bandwidth and $\mathcal{SC}_1$ is allocated $c_0 W/k$ bandwidth for transmission in the highway. We specify a k' as follows:

$$k' = \arg\sup_{1 \leq i \leq k}\left\{\frac{8\kappa}{5\pi c_0 c_3\delta\tau} kLn^{\alpha_i/2-\alpha}\log n \leq R\right\}.$$

Then we know the number of flows across each percolation cell s can be upper bounded by $\frac{20\pi c_3\delta\tau}{\kappa}\frac{Rn^{\alpha}}{kL\sqrt{n^{\alpha_i}}}$ for clusters belonged to an arbitrary super cluster $\mathcal{SC}_i$ $(1 \leq i \leq k')$. Therefore

$$\lambda_j^h \geq \frac{W/k}{\frac{20\pi c_3\delta\tau}{\kappa}\frac{Rn^{\alpha}}{kL\sqrt{n^{\alpha_i}}}} = \frac{\kappa WL\sqrt{n^{\alpha_i}}}{20\pi c_3\delta\tau Rn^{\alpha}} = \Omega(\frac{L\sqrt{n^{\alpha_i}}}{Rn^{\alpha}}).$$

For $\mathcal{C}_j \in \mathcal{SC}_i$ $(k'+1 \leq i \leq k)$, the number of flows cross a cell s is also upper bounded by $\frac{20\pi c_3\delta\tau}{\kappa}\frac{Rn^{\alpha}}{kL\sqrt{n^{\alpha}k'}}$, the same with clusters of $\mathcal{SC}_{k'}$ and in this case $\lambda_j^h = \Omega(\frac{Ln^{\alpha_{k'}/2}}{Rn^{\alpha}})$. ■

Recall Lemma 5.4, we know the bottleneck is due to data delivery on the highway when $R \geq \Omega(\frac{L\log n}{\sqrt{n^{\alpha}}})$ and let λ_j^u, λ_j^d denote the data rate of the uplink and downlink, respectively. And $\lambda_j = \min\{\lambda_j^u, \lambda_j^d, \lambda_j^h\} = \lambda_j^h$. The achievable capacity λ for the whole network is therefore

$$\lambda = \min\{\lambda_j | 1 \leq j \leq n_s\} = \Omega\left(\frac{L}{R\sqrt{n^{\alpha}}}\right) = \Omega\left(\frac{L}{R\sqrt{n_s}}\right).$$

Therefore our scheme approaches the maximized capacity as Eqn. (5) in order sense if we substitute $R = \frac{L}{\sqrt{\pi(1+L^2\sigma_{\mathcal{O}})}}$ into λ and $\lambda = \Omega(\max\{\frac{1}{\sqrt{n_s}}, \frac{L\sigma_{\mathcal{O}}}{\sqrt{n_s}}\})$.

Our results can generalize various results obtained in non-heterogeneous network like [3], [4], [6]–[8]. Here we highlight

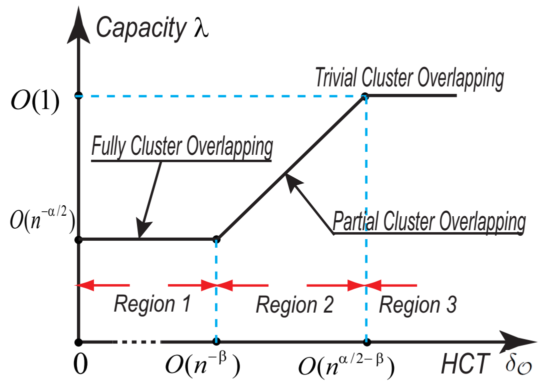

Fig. 5: Relationship between capacity λ and $\sigma_{\mathcal{O}}$ (log scale). In region 2, the network capacity increases with $\sigma_{\mathcal{O}}$. $\sigma_{\mathcal{O}} = 0$ means a non-heterogeneous network with the same achievable capacity as prior works.

two features in our clustered network, *heterogeneous cluster traffic* and *heterogeneous cluster size*, both of which are characterized in realistic multicast network.

C. The Impact of Heterogeneous Cluster Traffic

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. (3) 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 5 illustrates the relationship between λ and $\sigma_{\mathcal{O}}$ derived from Eqn. (5). 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 $O(W)$.

D. The Impact of Heterogeneous Cluster Size

HCS characterizes a network property that each multicast session does not comprises of clients with identical number. We show that HCS is not a deterministic factor for the achievable network capacity λ. However, we find that smaller cluster can achieve a larger capacity by relying on the *information highway* constructed mainly by larger clusters. Based on Eqn. (11), we provide a roughly estimation of λ_j: $\lambda_j = \Theta(\sqrt{\frac{n^{\alpha_i}}{n^\alpha}}\lambda)$ if $\mathcal{C}_j \in \mathcal{SC}_i$. During our derivation, we propose a new concept of super clusters, which contains clusters of the same size in order sense. The reason why we propose such a concept is that the scaling property of the achievable capacity is sustained by a certain group. It is more convenient to handle a discrete variable than a continuous one and we can gain a general insight of the impact of the cluster size.

The capacity achieving scheme is only for UCRM in this work. Such property can be generalized to other network topology as well because large clusters can be viewed as helping clusters in all the networks. And small clusters usually can take advantage of this benefits. In addition, a strong assumption is made in our paper that the number of cluster with size $\Theta(n^{1-\alpha}) = \Theta(p)$ must be larger than $c_0 n_s$. Such an assumption is crucial for highway construction and we will discuss some more general HCS in our future work.

VI. 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.

References

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

[2] L. L. Xie and P. R. Kumar, "*A network information theory for wireless communication: Scaling laws and optimal operation*", in Proc. IEEE Trans. Info. Theory, vol. 50, pp. 748-767, 2004.

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

[4] X.-Y. Li, S.-J. Tang, and O. Frieder. "*Multicast capacity for large scale wireless ad hoc networks*", in Proc. ACM Mobicom'08, San Francisco, CA, USA, Sept. 2008.

[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] S. Li, Y. Liu, X.-Y. Li, "*Capacity of Large Scale Wireless Networks Under Gaussian Channel Model*", in Proc. ACM MobiCom'08, San Francisco, CA, USA, Sept. 2008.
[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 Proc. 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 Proc. IEEE Journal on Selected Areas of Communications, 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] Fima C Klebaner "*Introduction to Stochastic Calculus with Applications, 2nd Edition*" published by Imperial College Press.

Appendix A
Proof of Theorem 3.1

According to Lemma 3.3, $\delta_{h,b} \geq \frac{c_0 n_s p r^2}{2L^2}$ when $r = \Omega(L/\sqrt{n_s p})$. Then according to Lemma 3.2, we know

$$\sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T}\sum_{h=1}^{h^b_{\mathcal{C}_j}} 1 \leq \frac{\sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T}\sum_{h=1}^{h^b_{\mathcal{C}_j}} \delta_{h,b}}{\frac{c_0 n_s p r^2}{2L^2}} \leq \frac{2L^2 T W}{c_0 r^2}$$

Utilizing Lemma 3.1 and Cauchy inequality, we have

$$\begin{aligned}\sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T}\sum_{h=1}^{h^b_{\mathcal{C}_j}} \ell_b^h &\leq \sqrt{\sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T}\sum_{h=1}^{h^b_{\mathcal{C}_j}} (\ell_b^h)^2 \sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T}\sum_{h=1}^{h^b_{\mathcal{C}_j}} 1} \\ &\leq \sqrt{\frac{32W^2T^2L^4}{c_0\Delta^2 r^2}}. \qquad (12)\end{aligned}$$

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

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

Applying Lemma 3.5, 3.7, we obtain that if $\mathbb{S}_1 \neq \emptyset$, there exists a constant $c_4 > 0$, such that

$$\begin{aligned}\sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T}\sum_{h=1}^{h^b_{\mathcal{C}_j}} \ell_b^h &\geq c_4 \sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T} |EMST(r-\mathcal{C}_j)| \\ &\geq c_4 \left(\sum_{\mathcal{C}_j \in \mathbb{S}_1}\sum_{b=1}^{\lambda_j T} |EMST(r-\mathcal{C}_j)| \right. \\ &\quad \left. + \sum_{\mathcal{C}_j \in \mathbb{S}_2}\sum_{b=1}^{\lambda_j T} \sqrt{|\mathcal{C}_j|}\int_{\mathcal{O}} \sqrt{\phi(\xi)}d\xi \right) \\ &\geq c_4 T \left(\sum_{j=1}^{n_s} \lambda_j \sqrt{|\mathcal{C}_j|} \right) \psi\left(\phi(\cdot), \mathbb{S}_1\right), \quad (13)\end{aligned}$$

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. (12) into Eqn. (13), we can obtain

$$\begin{aligned}\sum_{j=1}^{n_s} \lambda_j \sqrt{|\mathcal{C}_j|} &\leq \frac{1}{c_4 T \psi\left(\phi(\cdot), \mathbb{S}_1\right)} \sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T}\sum_{h=1}^{h^b_{\mathcal{C}_j}} \ell_b^h \\ &\leq \frac{4\sqrt{2}WL^2}{\Delta\sqrt{c_0} c_4 r \psi\left(\phi(\cdot), \mathbb{S}_1\right)}. \qquad (14)\end{aligned}$$

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

$$\begin{aligned}\sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T}\sum_{h=1}^{h^b_{\mathcal{C}_j}} \ell_b^h &\geq c_4 \sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T} |EMST(r-\mathcal{C}_j)| \\ &\geq c_4 \sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T} \sqrt{|\mathcal{C}_j|}\int_{\mathcal{O}} \sqrt{\phi(\xi)}d\xi \\ &= c_4 T \int_{\mathcal{O}} \sqrt{\phi(\xi)}d\xi \left(\sum_{j=1}^{n_s} \lambda_j \sqrt{|\mathcal{C}_j|} \right). \quad (15)\end{aligned}$$

Substitute (15) into (12), we obtain

$$\begin{aligned}\sum_{j=1}^{n_s} \lambda_j \sqrt{|\mathcal{C}_j|} &\leq \frac{1}{c_4 T \int_{\mathcal{O}} \sqrt{\phi(\xi)}d\xi} \sum_{j=1}^{n_s}\sum_{b=1}^{\lambda_j T}\sum_{h=1}^{h^b_{\mathcal{C}_j}} \ell_b^h \\ &\leq \frac{4\sqrt{2}WL^2}{\Delta\sqrt{c_0} c_4 r \int_{\mathcal{O}} \sqrt{\phi(\xi)}d\xi} \\ &\leq \frac{4WL\sqrt{2n_s p}}{\Delta c' \sqrt{c_0} c_4 \int_{\mathcal{O}} \sqrt{\phi(\xi)}d\xi}. \qquad (16)\end{aligned}$$

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 3.5, 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. (3).

Appendix B
Proof of Lemma 5.1

Let $\mathcal{A}(d, r_1, r_2)$ denote the overlapping area of two circles of radius r_1, r_2 with centers of distance d away, and $\mathcal{C}(r)$ denotes number of nodes within radius r. Note that the distribution of cluster clients is HPP within a circle of radius R, thus we could obtain

$$\begin{aligned}
&\mathbf{Pr}(\mathcal{N}(r) = k) \\
&= \sum_{m=0}^{n_s} \mathbf{Pr}(\mathcal{N}(r) = k | \mathcal{C}(r+R) = m) \mathbf{Pr}(\mathcal{C}(r+R) = m) \\
&= \sum_{m=0}^{n_s} (\int_0^{R+r} \sum_{\mathcal{S}} (\prod_{i=1}^{m} (e^{-\mu_1} \frac{\mu_1^{\upsilon_i}}{\upsilon_i!})) \frac{2x}{(R+r)^2} dx) e^{-\mu_2} \frac{\mu_2^m}{m!} \\
&= \int_0^{R+r} (\sum_{m=0}^{n_s} e^{-m\mu_1} \frac{(m\mu_1)^k}{k!} e^{-\mu_2} \frac{(\mu_2)^m}{m!}) \frac{2x}{(R+r)^2} dx \\
&\leq \int_0^{R+r} (e^{-\mu_1 \mu_2} \frac{(\mu_1 \mu_2)^k}{k!}) \frac{2x}{(R+r)^2} dx \\
&= \int_0^{R+r} e^{-\frac{\mathcal{A}(x,R,r) n_s p}{L^2}} \frac{(\frac{\mathcal{A}(x,R,r) n_s p}{L^2})^k}{k!} \frac{2x}{(R+r)^2} dx.
\end{aligned}$$

During the above derivation, we utilize the following notations to simplify our calculations:

$$\mu_1 = \frac{\mathcal{A}(x,R,r)p}{\pi(R+r)^2}, \mu_2 = \frac{\pi(R+r)^2 n_s}{L^2} \text{ and } \sum_{i=1}^{m} \upsilon_i = k$$

$$\mathcal{S} = \{\upsilon_0, \upsilon_1 \dots \upsilon_m | \sum_{i=1}^{m} \upsilon_i = k\}$$

The overlapping area of two circles are smaller than the area of the small circle therefore

$$\mathcal{A}(x,R,r) \leq \begin{cases} \pi r^2 & x \in [0, R+r) \\ 0 & x \in [R+r, \infty) \end{cases}$$

Then we substitute it into $\mathbf{Pr}(\mathcal{N}(r) = k)$ and obtain

$$\mathbf{Pr}(\mathcal{N}(r) = k) \leq 2e^{-\frac{\pi r^2 n_s p}{L^2}} \frac{(\frac{\pi r^2 n_s p}{L^2})^k}{k!}.$$

Appendix C
Proof of Lemma 5.6

Denote the Abscissa and Ordinate of a point $\xi \in \mathcal{O}$ as ξ_a and ξ_o. Assume ξ^1 and ξ^2 are two representing points in two percolated cells on the same horizontal path. Let $h(\xi^1, \xi^2)$ denote the number of hops required for transmission from ξ^1 to ξ^2. We can prove that there exists a constant c_3, such that

$$\mathbf{Pr}\left(h(\xi^1, \xi^2) \leq \frac{c_3 \sqrt{n_s p} |\xi_a^1 - \xi_a^2|}{\tau L}\right) \geq 1 - \delta(n_s p), \quad (17)$$

where $\lim_{n_s p \to \infty} \delta(n_s p) = 0$. Now we divide the path into $K = \frac{L}{|\xi_a^1 - \xi_a^2|}$ sub-paths as in the second figure in Figure. 3 and $\ell_i = \frac{L\wp(n_s p, i)}{K} (1 \leq i \leq K)$ denotes the length of the i_{th} sub-path. Then the proof of Eqn. (17) is identical to the following inequality.

$$\lim_{n_s p \to \infty} \mathbf{Pr}(\ell_i = \omega(\frac{L}{K})) = 0,$$

which is also identical to prove

$$\lim_{n_s p \to \infty} \mathbf{Pr}(\wp(n_s p, i) = \infty) = 0.$$

For each path on the highway system, the length of it is on the order of $\Theta(L)$ according to its property. Thus there exists a constant c', such that the length of every path is upper bounded by $c' L$, which means $\sum_{i=1}^{K} \ell_i \leq c' L$. Taking the expectation on both sides, we can obtain

$$\begin{aligned}
&\sum_{i=1}^{K} \inf_{\wp(n_s p, i) = \infty} \{\wp(n_s p, i)\} \mathbf{Pr}(\ell = \omega(L/K)) \frac{L}{K} \\
&\leq \sum_{i=1}^{K} \mathbf{E}[\ell_i] = \mathbf{E}[\sum_{i=1}^{K} \ell_i] \leq c' L.
\end{aligned}$$

Then we can obtain

$$\mathbf{Pr}(\ell = \omega(L/K)) \leq \frac{c'}{\inf_{\wp(n_s p, i) = \infty} \{\wp(n_s p, i)\}}$$

Taking $\delta(n_s p) = \frac{c'}{\inf_{\wp(n_s p, i) = \infty} \{\wp(n_s p, i)\}}$ and note that each hop can forward the packet of distance $\Theta(L/\sqrt{n_s p})$ on the highway, we complete the proof of Eqn. (17). The condition of vertical path is the dual of the horizontal case.

Now assume that a sender at ξ^s generates a traffic to a receiver ξ^r of distance ℓ away. Let ξ^i be the intersection of the horizontal and vertical path in which $\mathcal{F}_h(\xi^s)$ and $\mathcal{F}_v(\xi^r)$ reside. According to the properties of percolation path, every percolation path is constrained within a strip of width $\kappa \log(n_s p) \frac{L}{\sqrt{n_s p}}$ and, such that

$$\begin{cases}
\max\{|[\mathcal{F}_h(\xi^s)]_a - \xi_a^i|\} \leq |\xi_a^s - \xi_a^i| + \frac{\kappa L \log(n_s p)}{\sqrt{n_s p}} \\
\max\{|[\mathcal{F}_v(\xi^r)]_o - \xi_o^i|\} \leq |\xi_o^r - \xi_o^i| + \frac{\kappa L \log(n_s p)}{\sqrt{n_s p}} \\
\max\{\sqrt{(\xi_a^s - \xi_a^i)^2 + (\xi_o^r - \xi_o^i)^2}\} \leq \ell + \frac{2\kappa L \log(n_s p)}{\sqrt{n_s p}}
\end{cases}$$

Recall (12), then the number of hops required for transmission with nodes ℓ away is upper bounded by

$$
\begin{aligned}
& h(\mathcal{F}_h(\xi^s), \xi^i) + h(\xi^i, \mathcal{F}_v(\xi^r)) \\
\le & \frac{c_3\sqrt{n_s p}}{\tau L}\left(|[\mathcal{F}_h(\xi^s)]_a - \xi_a^i| + |[\mathcal{F}_v(\xi^r)]_o - \xi_o^i|\right) \\
\le & \frac{c_3}{\tau}\left(\frac{\sqrt{n_s p}}{L}(|\xi_a^s - \xi_a^i| + |\xi_o^r - \xi_o^i|) + 2\kappa\log(n_s p)\right) \\
\le & \frac{c_3}{\tau}\left(\frac{\sqrt{n_s p}}{L}\sqrt{2((\xi_a^s - \xi_a^i)^2 + (\xi_o^r - \xi_o^i)^2)} + 2\kappa\log(n_s p)\right) \\
\le & \frac{c_3}{\tau}\left(\frac{\sqrt{2n_s p}\ell}{L} + 4\kappa\log(n_s p)\right).
\end{aligned}
$$

APPENDIX D
PROOF OF LEMMA 5.7

There are at least $\frac{\kappa}{\delta}\lfloor\frac{\sqrt{2n_s p}\mathcal{R}}{\tau L}\rfloor^2 \ge \frac{\kappa}{\delta}(\frac{\sqrt{2n_s p}\mathcal{R}}{\tau L} - 1)^2$ percolated cells within the disk of radius $\mathcal{R}$. Follow the prim algorithm used in Lemma 5.3 to construct a spanning tree. Let $\mathcal{I}(s,i)$ be the indicator whether cell s is used in the i_{th} step and $\mathcal{P}$ be the probability that s is used in the whole process. Then based on Lemma 5.5, 5.6:

$$
\begin{aligned}
\mathcal{P} \le & \sum_{j=1}^{|\mathcal{C}_j|}\mathbf{Pr}(\mathcal{I}(s,i)=1) \\
\le & \sum_{j=1}^{|\mathcal{C}_j|}\frac{c_3}{\tau\frac{\kappa}{\delta}(\frac{\sqrt{2n_s p}\mathcal{R}}{\tau L}-1)^2}\left(\frac{4\sqrt{n_s p}R}{\lfloor\sqrt{|\mathcal{C}_j|+1-i}\rfloor L} + 4\kappa\log(n_s p)\right) \\
\le & \frac{c_3\delta\tau L^2}{2\kappa n_s p R^2}\sum_{j=1}^{|\mathcal{C}_j|}\left(\frac{4\sqrt{n_s p}R}{\lfloor\sqrt{|\mathcal{C}_j|+1-i}\rfloor L} + 4\kappa\log(n_s p)\right) \\
\le & \frac{c_3\delta\tau L^2}{2\kappa n_s p R^2}\left(\frac{10\sqrt{n_s p|\mathcal{C}_j|}R}{L} + 4\kappa|\mathcal{C}_j|\log(n_s p)\right) \\
\le & \frac{10 c_3\delta\tau L\sqrt{|\mathcal{C}_j|}}{\kappa\sqrt{n_s p}R}.
\end{aligned}
$$

APPENDIX E
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. Here we do not explicitly explain how to conduct the transformation but only list the results. For the thorough proof, one can refer to [22].

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

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 $\phi_i^m \le \phi_j^m$ iff $i \ge j$.

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

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

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 $k \le i$, $\phi_k^m > \frac{1}{\pi R^2}$ and for $k > i$, $\phi_k^m \le \frac{1}{\pi R^2}$
- For $k < j$, $\phi_k^m \ge \frac{1}{\pi R^2}$ and for $k \ge j$, $\phi_k^m < \frac{1}{\pi R^2}$
- For $k \le l$, $\phi_k^m > 0$ 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:

$$\phi_i^m \ge \frac{1}{\pi R^2} + \Delta_i \quad \phi_j^n \le \frac{1}{\pi R^2} - \Delta_i - \Delta_l \quad \phi_l^m \ge \Delta_l \tag{18}$$

$$
\begin{aligned}
(\phi_i^m - \Delta_i)^2 + (\phi_j^m + \Delta_i + \Delta_l)^2 + (\phi_l^m - \Delta_l)^2 \\
\le (\phi_i^m)^2 + (\phi_j^m)^2 + (\phi_l^m)^2
\end{aligned} \tag{19}
$$

$$\Delta_i \le \frac{\phi_j^m(\sqrt{\phi_i^m} - \sqrt{\phi_j^m})(\sqrt{\phi_i^m} - \sqrt{\phi_k^m})}{\sqrt{\phi_i^m}(1 + (\frac{\phi_i^m - \phi_j^m}{\phi_j^m - \phi_l^m})^2)(\sqrt{\phi_j^m} + \sqrt{\phi_l^m})} \tag{20}$$

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.

Actually, Eqn. (20) is not a necessary condition for the correctness of the derivation. However, such a constraint can guarantee the correctness of the following derivation based on Taylor's expansion. 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

$$
\begin{aligned}
& \sqrt{\phi_i^m - \Delta_i} + \sqrt{\phi_j^m + \Delta_i + \Delta_l} + \sqrt{\phi_l^m - \Delta_l} \\
\le & \sqrt{\phi_i^m} + \sqrt{\phi_j^m} + \sqrt{\phi_l^m}.
\end{aligned} \tag{21}
$$

According to step 3, we can obtain

$$\Delta_l = \rho\Delta_i - \frac{\Delta_i^2 + \Delta_l^2 + \Delta_i\Delta_l}{\phi_j^m - \phi_l^m}. \tag{22}$$

Rewrite Eqn. (21) and substitute Eqn. (22) into it, we can obtain

$$
\begin{aligned}
&\sqrt{\phi_i^m - \Delta_i} + \sqrt{\phi_j^m + \Delta_i + \Delta_l} + \sqrt{\phi_l^m - \Delta_l} \\
&-(\sqrt{\phi_i^m} + \sqrt{\phi_j^m} + \sqrt{\phi_l^m}) \\
\leq & -\frac{\Delta_i}{2\sqrt{\phi_i^m}} + \frac{\Delta_i + \Delta_l}{2\sqrt{\phi_j^m}} - \frac{\Delta_l}{2\sqrt{\phi_l^m}} + \delta(\Delta_i, \Delta_l) \\
\leq & \frac{\Delta_i}{2\sqrt{\phi_j^m}} \left(\frac{\sqrt{\phi_i^m} - \sqrt{\phi_j^m}}{\sqrt{\phi_i^m}} - \frac{\sqrt{\phi_j^m} - \sqrt{\phi_l^m}}{\sqrt{\phi_l^m}} \frac{\Delta_l}{\Delta_i} \right) + \delta(\Delta_i, \Delta_l) \\
\leq & \frac{\Delta_i}{2\sqrt{\phi_j^m}} \left(\frac{\sqrt{\phi_i^m} - \sqrt{\phi_j^m}}{\sqrt{\phi_i^n}} - \frac{\sqrt{\phi_j^m} - \sqrt{\phi_l^m}}{\sqrt{\phi_l^m}} \times \right. \\
& \left. (\rho + \frac{1 + \rho + \rho^2}{\phi_j^m - \phi_l^m} \Delta_i) \right) + \delta(\Delta_i, \Delta_l) \\
\leq & -\frac{\Delta_i(\sqrt{\phi_i^m} - \sqrt{\phi_j^m})(\sqrt{\phi_i^m} - \sqrt{\phi_l^m})}{2\sqrt{\phi_i^m \phi_j^m \phi_l^m}} + \\
& \left(\frac{1 + \rho + \rho^2}{2\sqrt{\phi_j^m \phi_l^m}(\sqrt{\phi_j^m} + \sqrt{\phi_l^m})} + \frac{1 + \rho^2}{2(\phi_j^m)^{1.5}} \right) \Delta_i^2 \\
\leq & -\frac{\Delta_i(\sqrt{\phi_i^m} - \sqrt{\phi_j^m})(\sqrt{\phi_i^m} - \sqrt{\phi_l^m})}{2\sqrt{\phi_i^m \phi_j^m \phi_l^m}} + \\
& \frac{\Delta_i^2(1 + \rho^2)(\phi_j^m + \phi_l^m)}{2\sqrt{\phi_j^m} \phi_l^m \phi_j^m} \\
\leq & \, 0.
\end{aligned}
$$

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.

Now we know Algorithm 3 can reduce $\Upsilon(\phi(\cdot))$ 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 $\phi^u(\cdot)$.