Multicast Capacity with Max-Min Fairness for Heterogeneous Networks

Yixuan Li, Qiuyu Peng, Xinbing Wang
Department of Electronic Engineering
Shanghai Jiao Tong University, China
Email: {lyx1990116, pqy, xwang8}@sjtu.edu.cn

Abstract—In this paper, we investigate the multicast capacity for static ad hoc networks with heterogeneous clusters. We study the effect of heterogeneous cluster traffic (HCT) on the achievable capacity. HCT means cluster clients are more likely to appear near the cluster head, instead of being uniformly distributed across the network. Such property is commonly found in real networks. By adopting max-min fairness, the minimum among all individual multicast capacities of clusters can be maximized. Since this minimal individual multicast capacity will not be maximized unlimitedly, our work focuses on deriving the upper bound of the minimum individual multicast capacity (we refer it as minimum capacity for simplicity) in HCT, which provides the best performance for the minimum multicast capacity to attain in the whole network. We find that HCT increases minimum capacity for ad hoc networks. Further, the multicast capacity achieving scheme is provided to justify the derived asymptotic upper bound for the minimum capacity. Our work can generalize various results obtained under non-heterogeneous networks in previous literature.

Index Terms—Static Clustered Network, Multicast Capacity, Heterogeneous Cluster.

I. INTRODUCTION

A 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 a static network. Later in [2], the capacity result is analyzed under more general fading channel condition 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.

The multicast network, which generalizes the above unicast network, received more attention recently and the estimation of the achievable multicast capacity is required in many applications like sensor networks and TV streaming [25]. Li, et al. [4] studied the achievable capacity in multicast networks. In their work, there are $n$ multicast sessions, each comprised of 1 source and $k$ destinations and they find that the capacity scales as $\Theta(1/\sqrt{k n \log n})$ based on the Manhattan

Routing scheme. Their results generalize both unicast and broadcast [5] capacity. In [6], Shakkottai, et al. studied a different multicast framework where there are $n^\epsilon$ multicast sources and $n^{1-\epsilon}$ destinations per flow. Their network can support a rate of $\Theta(1/\sqrt{n^\epsilon \log n})$ 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 optimal 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 networks can also be viewed as clustered networks. However, there are few works concerning the cluster behavior of multicast networks. Uniformly distributed cluster (multicast session) traffic is assumed and cluster heterogeneities are rarely involved in previous works. Actually, most real networks are characterized by various clustered heterogeneities and some aspects have already been investigated in unicast networks, 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. Since spontaneous grouping of the nodes around a few attraction points occurs commonly in wireless network, G. Alfano, et al. [15], [16] extended the capacity scaling to networks with inhomogeneous node density. In their work, nodes are generated according to a specified point process and they show that the bottleneck is in the node sparse region because the network capacity is related to the minimum node intensity.

Pattern heterogeneity: It is likely that there exists more than one type of traffic pattern in the network and nodes of the same traffic pattern constitute a cluster. In [17], Wang, et al. studied a unified modeling framework composed of unicast, multicast, broadcast traffic. Later in [18], Ji, et al. explored networks composed of both unicast and converge-cast traffic and show that MIMO cooperation can be applied to increase capacity for both types of traffic. Li, et al. [19] dealt with networks containing some helping nodes for packet delivery. Using this method, the normal nodes and helping nodes can be viewed as two clusters.

The primary incentive motivates us to investigate clustered networks is because of its practical use in real life. For instance, in military battlefield, commanders from different

---

\(^1\)Given two functions $f(n) > 0$ and $g(n) > 0$: $f(n) = o(g(n))$ means $\lim_{n \to \infty} f(n)/g(n) = 0$; $f(n) = \Omega(g(n))$ 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 $g(n) = O(f(n))$. 
places must send requirements through a common wireless channel to their respective soldiers around them. In sensor networks, local schedulers also need to send packets to their adjacent client sensors. In addition, the study of capacity performance in clustered networks under multicast traffic is still not sufficient yet and that is another reason why we set the model. Since clients of the same data flow are non-uniformly distributed, the open question is:

- What are the impacts of heterogeneous traffic on multicast capacity in clustered networks?

In this paper, nodes in each multicast session are comprised of a cluster and heterogeneous cluster traffic (HCT) is defined by: clients of the same cluster (data flow) are likely to be deployed around a cluster head specified by an inhomogeneous Poisson process (IPP). We describe this clustering behavior with a variable $\sigma$, which describes the extent of HCT.

Our asymptotic analysis for multicast capacity is based on the max-min fairness, through which the minimum among all individual multicast capacities of clusters can be maximized. The formal definition of max-min fairness will be provided in the next section. Typically, there are more than one criterion in evaluating the efficiency of a certain scheduling policy. Most previous literature such as [4] [15] [16] aims at deriving the bounds for total throughput. However, maximizing the total throughput is sometimes not desirable because it may cause imbalance among individual multicast capacities of clusters and cannot avoid the abnormal minimum value of individual multicast capacity. For example, consider a network of 100 clusters, among which 99 clusters have equal rates of 100Mbps and one cluster has a rate of 1Mbps. Such abnormal individual multicast rate of 1Mbps can be regarded as service outage comparing with other large rates, and should be improved if possible. By adopting max-min fairness, such rate gap among different clusters can be narrowed and potential service outages can be reduced. The relative unfair individual multicast capacity of 1Mbps caused by traffic congestion or unreasonable rate allocation can be enhanced to a higher value.

As we know, this minimal individual multicast capacity will not be maximized unlimitedly, so there must exist an upper limit that the minimum capacity cannot exceed. Our paper, being set upon the max-min metric, focuses on deriving the upper bound of the minimum multicast capacity. Investigating the minimum capacity is meaningful because it provides the best performance for the minimum multicast capacity to attain in the whole network. Note that for the purpose of brevity, we will use the notion “minimum capacity” to represent the “minimum individual multicast capacity” throughout the paper.

Previously, researchers had established a framework for multicast rate allocation with max-min fairness in wired networks [20] [21]. In [20], the authors not only justified the max-min fairness for multicast traffic, but also gave efficient algorithms to compute the rate allocation vector based on the metric. Nevertheless, providing algorithms is sometimes not a straightforward way of showing the capacity performance of networks, although it does help in practical congestion control computing. Our work, instead, directly shows a quantitative analysis and bounds the value of minimum capacity using explicit formulations rather than iterative computing procedures. In other words, the seminar work [20] can help validate the max-min fairness for multicast traffic, our work then goes a step further to provide a scalable analysis for the best capacity performance that the minimum capacity can attain. We find that the minimum capacity increases in HCT because the total transmission length is shortened and we offer a quantitative relationship between the minimum capacity and $\sigma$.

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, close forms of the upper bound for minimum capacity with distribution function and distribution variance are derived respectively. In section V, we provide a routing scheme for achieving the upper bound of minimum capacity in 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 extended networks composed of $n_s = n^\alpha$ ($0 \leq \alpha < 1$) clusters distributed over a 2-dimensional torus region $\mathcal{O}$ of edge $2L = n^\beta$ ($0 \leq \beta \leq \alpha/2$). 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 $\xi$ is given by $|C_j|\phi(k_j, \xi)$, where $|C_j| \leq p = n^{1-\alpha}$ is the expected size of the cluster and $\phi(\cdot)$ is the dispersion density function. In order to investigate the impact of HCT on minimum capacity, we will discuss in detail here what further assumptions should be made for both $\phi(\cdot)$ and $|C_j|$. Because the dispersion density function $\phi(\cdot)$ determines the client distribution of each multicast session, intuitively, HCT is mainly described by the characteristics of $\phi(\cdot)$. In addition, HCS deals with the cardinality of each cluster, so it is strongly related to $|C_j|$ ($1 \leq j \leq n_s$).

A cluster dense regime is assumed, which means the distance between adjacent clusters $\frac{1}{\sqrt{n_s}}$ tends to 0 when $n$ approaches infinity. We will assume $\frac{1}{\sqrt{n_s}} = O(1)$ throughout the paper.

![Fig. 1: Demonstration of Network topology. Nodes in the same cluster are labeled with the same number. H and C represent head (source) and clients (destination), respectively.](image-url)
First, we outline the properties of the dispersion density function $\phi(\cdot)$ as follows:

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, non-negative, bounded, and continuous function with respect to the Euclidean distance $|k_j - \xi|$.

2. Integration $\phi(k_j, \xi)$ of $\xi$ over the whole torus $\mathcal{O}$ is equal to 1, $\int_{\mathcal{O}} \phi(k_j, \xi) d\xi = 1$.

The first property restricts the dispersion density function to a regime that clients are more likely to distribute around the cluster head, that is to say there are more clients around the cluster head and less clients in remote areas with respect to the cluster head. The second property can be derived by defining a non-negative, non-increasing continuous function $s(\rho)$ such that $\int_0^\infty \rho s(\rho)d\rho < \infty$ and then normalizing it over the whole area

$$\phi(k_j, \xi) = \frac{s(|\xi - k_j|)}{\int_{\mathcal{O}} s(|\xi - k_j|) d\xi}.$$ 

Notice that $\phi(k_j, \xi) = \Theta(s(|\xi - k_j|))$ in asymptotic analysis if we neglect the factor $\int_{\mathcal{O}} s(|\xi - k_j|) d\xi = \Theta(1)$.

Now we need to offer a quantitative value for a given dispersion density function $\phi(\cdot)$ to depict its degree of heterogeneity, then we can analyze the relationship between the minimum capacity and the degree of HCT. As we know, the expectation can describe the average value of a function and in this case, it corresponds to the average density distribution,

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

And variance can describe the fluctuation level of a function around its expectation. In this case, we define the variance of $\phi(\cdot)$ as distribution variance $\sigma_{\mathcal{O}}$, which can be utilized to depict HCT.

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

We omit the term $k_j$ due to the wrap around property of a torus and it can liberate us from the border effect. In case of uniform traffic, $\phi(\xi) \equiv \frac{1}{\mathcal{O}}$, 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 can achieve the largest value of the minimum capacity given a fixed $\sigma_{\mathcal{O}}$ in section IV.

We classify these $n_s$ clusters into $k$ super clusters ($\mathcal{S}\mathcal{C}$) based on their cluster size, in order to explain the effect of cluster size. For each cluster $C_j \in \mathcal{S}\mathcal{C}_i$ ($1 \leq i \leq k$), its size $|C_j| = \Theta(n^{1-\alpha_i})$, where $\alpha_i$ is an increasing sequence over $i$ and the clusters in $\mathcal{S}\mathcal{C}_1$ possess the largest size. In addition, some further assumptions are shown below:

1. $\alpha = \min\{\alpha_1, \alpha_2, \ldots, \alpha_{k-1}, \alpha_k\} = \alpha_j$.
2. There are $c_0 n_s$ ($0 < c_0 < 1$) clusters in $\mathcal{S}\mathcal{C}_1$, and the other $(1 - c_0) n_s$ clusters are randomly allocated to $\mathcal{S}\mathcal{C}_i$ ($2 \leq i \leq k$).

The second assumption indicates that the number of clusters with size $\Theta(p)$ is the same order as the total number of clusters. Although it is somewhat restrictive here, it is constructive when we design our capacity achieving scheme and allows us to gain important insight into the structure of the problem.

Under the above assumptions, cluster clients are likely to be distributed near the cluster head and the cluster size conforms to a poisson distribution with rate $|C_j|$. In addition, there are $\Theta(n_s/k)$ clusters in $\mathcal{S}\mathcal{C}_i$ for $2 \leq i \leq k$ applying the Chernoff bound. In order to simplify our analysis, we take both $|C_j|$ and $n^{1-\alpha_i}$ as the cluster size and $n_s/k$ as the number of clusters in $\mathcal{S}\mathcal{C}_i$ for $2 \leq i \leq k$. Such simplification does not influence our results in order sense. Figure 1 is an example of the network topology.

---

**B. Transmission Protocol**

All wireless transceivers can communicate over a common channel with 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 $j$; $\Delta$ 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 and the source. In this paper, we assume that the minimal transmission range is denoted by $r$ and we will prove that nearest neighbor transmission is also dominant. Figure 2 illustrates two concurrent transmissions.

---

**C. Traffic Model**

A multicast traffic pattern is assumed where each cluster head generates data flows to their respective clients, e.g. in Figure 1. The one to many data flow in $C_j$ can be modeled as a multicast tree $T_j$ spanning 1 head and $|C_j|$ clients. In [4], Euclidean minimal spanning tree (EMST) is employed to
bound the length of transmission for each multicast session in non-clustered networks with uniform node distribution. We employ new techniques to study heterogenous clustered networks, which generalizes uniform cases. Let $EMST(C_j)$ denote the EMST for $C_j$ and the definition is given below.

Definition of EMST: Assume that cluster head $v_j$ generates its cluster members $\{v_j^{(1)}, v_j^{(2)}, \ldots, v_j^{(|C_j|)}\}$ according to an IPP. The EMST for $C_j$ in 2-dimensional space $\mathbb{R}^2$ connects points $v_j \cup \{v_j^{(1)}, v_j^{(2)}, \ldots, v_j^{(|C_j|)}\}$ using lines such that the total Euclidean length of all the lines is minimized and any point can be reached from any other by following the lines.

Note that the communication between any SD pairs can also go through multiple relays from other clusters.

Since the network topology defined above is fundamentally related to both the number of clusters $n_s$ and the network physical extension $L$, we hope to find out how the minimum capacity asymptotically scales. Now we give the definition of capacity.

Definition of Asymptotic Capacity: Let $\lambda_j (1 \leq j \leq n_s)$ denote the individual multicast capacity for cluster $C_j$. The multicast rate vector $\lambda_j = (\lambda_1, \lambda_2, \ldots, \lambda_{n_s-1}, \lambda_n)$ consists of the individual multicast capacity of all clusters, and it is also identified as network capacity. The minimum multicast capacity is defined as $\lambda = \text{min}\{\lambda_1, \lambda_2, \ldots, \lambda_{n_s-1}, \lambda_n\}$, which is the minimal achievable data rate in these clusters. Then $\lambda = \Theta(f(n))$ is defined as the asymptotic minimum capacity if there exist constants $c > c' > 0$, such that

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

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

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

D. Max-min Fairness

A feasible multicast rate vector $\lambda_s$ is max-min fair if and only if an increase of any individual multicast capacity within the domain of feasible allocations must be at the cost of a decrease of some already smaller individual multicast capacity. Formally, we have the following.

Definition of Max-min Fairness: [22] A multicast rate vector $\lambda_s$ is max-min fair if and only if, for any other feasible multicast rate vector $\hat{\lambda}_s$, if the individual multicast capacity of cluster $i$ satisfies $\lambda_i > \hat{\lambda}_i$ then there must exists some cluster $j$ such that $\lambda_j < \hat{\lambda}_j$ and $\hat{\lambda}_j < \lambda_j$.

In [20], the authors justified the properties of max-min fairness for multicast service in a wired network. In their model, the rate allocation for different multicast sessions was also denoted by a rate vector, in which an element represents the individual multicast capacity for one cluster. The properties can be adapted to the wireless networks since our mathematical model is almost the same as the single-rate multicast service model shown in [20]. That is to say, max-min fairness can be applied in multicast traffic to enhance the minimum capacity. Based on this, our major concern is to find out how large can the minimum capacity attain at most. The rest of the paper will focus on this $\lambda$, i.e., the minimum capacity. Since the main purpose of our paper lies in deriving the upper limit for minimum capacity, we omit the detailed proof of the max-min fairness properties to avoid redundancy.

E. Mathematical Notations

Throughout our paper, we denote $h^b_{C_j}$ as the number of hops required for transmitting bit $b$ to all clients in $C_j$, $\ell^b_{h}$ is the length of transmission of bit $b$ in its $h_{th}$hop. $\delta_{n,b}$ 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 $\xi$ with radius $R$. $R$ is the radius of the influential region centered at the head. Nodes outside the influential region cannot act as relays for that cluster. $|T_j|$ is the total Euclidean length of a multicast tree $T_j$, which can be calculated by summing up the Euclidean length of all edges belonging to the tree.

F. Useful Known Results

Throughout the paper, the following two known results would be used for proving theorems. In particular, the Chernoff bound is applied to bound the probability of sums of independent variables and the Hölder’s inequality will be consulted for deriving the bound for minimum capacity.

Lemma 2.1: Chernoff bound: Given a Poisson random variable $X$ with expectation $\eta$, for any $\delta > 0$, the Chernoff bound is given by

$$Pr(X > (1 + \delta)\eta) < \left(\frac{e^\delta}{(1 + \delta)^{1+\delta}}\right)^\eta.$$

Lemma 2.2: Hölder’s inequality: If $S$ is a measurable subset of $R^n$ with the Lebesgue measure, and $f$ and $g$ are measurable real- or complex-valued functions on $S$, then Holder inequality is

$$\int_S |f(\xi)g(\xi)|d\xi \leq \left(\int_S |f(\xi)|^p d\xi\right)^{\frac{1}{p}} \left(\int_S |g(\xi)|^q d\xi\right)^{\frac{1}{q}},$$

where $1/p + 1/q = 1$.

G. Main Results

Our main results are summarized as follows:

- Given the dispersion density function $\phi(\cdot)$, the upper bound of minimum capacity is given as follows:

$$\lambda \leq \min \left\{W, \frac{cWL}{\sqrt{n_s} \int_\mathbb{C} \sqrt{\phi(\xi)} d\xi}\right\},$$

where $c$ is some constant.

- HCT increases the upper bound of minimum capacity $\lambda$ and a universal relationship between $\lambda$ and $\sigma_{\mathbb{C}}$ is obtained as:

$$\lambda \leq \min \left\{O(W), O\left(\frac{\max\{1, L\sigma_{\mathbb{C}}\} W}{\sqrt{n_s}}\right)\right\}.$$
III. UPPER BOUND TO MINIMUM CAPACITY

In this section, we will provide an upper bound of the minimum capacity, which is a crucial step for deriving the relationship between the capacity and the distribution variance. First, we will prove several key results inherent to the static network with multicast traffic. There are some tradeoffs that must be tolerated among which are the number of hops, transmission range, limited radio resources. Therefore, a thorough comprehension of the implicit relationships among them is constructive for deriving the upper bound of the achievable capacity.

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 and 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 sufficiently large.

\[
\sum_{j=1}^{n_x} \sum_{b=1}^{h_c} \sum_{h=1}^{b_c} \frac{\pi}{16} \Delta^2(\ell_c^b)^2 \leq WT^2. \tag{3}
\]

**Proof:** When \( T \) is sufficiently large, the total number of bits communicated from the head to its clients in cluster \( C_j \) is \( \lambda_j T \). Assume two SD pairs \( X_i \rightarrow X_h \) and \( X_j \rightarrow X_l \) are active in the given time slot, then according to the transmission protocol model,

\[
|X_k - X_h| \geq \frac{\Delta}{2} (|X_i - X_h| + |X_j - X_l|),
\]

which is derived in [1]. Thus disks of radius \( \frac{\Delta}{2} \) times the transmission radius centered at the receiver can be viewed as an “exclusion region” that rejects transmitters from other data flows. Such a property also holds for broadcast that the transmission range is defined as the furthest node that can receive the packets. Let \( \ell_c^b \) be the transmission radius centered at the receiver for the \( h \)-th hop of bit \( b \), and \( S^b_h \) be the overlap area between the “exclusion region” of bit \( b \)'s \( h \)-th hop and the deployed region \( O \). Note that at least a quarter of the disk is within the given region (the extreme situation happens when the node is near the periphery of the region). Then

\[
S^b_h \geq \frac{\pi}{4} \left( \frac{\Delta \ell_c^b}{2} \right)^2 = \frac{\pi \Delta^2(\ell_c^b)^2}{16}.
\]

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:

\[
\sum_{j=1}^{n_x} \sum_{b=1}^{h_c} \sum_{h=1}^{b_c} \frac{\pi}{16} \Delta^2(\ell_c^b)^2 \leq \sum_{j=1}^{n_x} \sum_{b=1}^{h_c} \sum_{h=1}^{b_c} \frac{S^b_h}{W} \leq WT^2. \]

The above inequality is a basic requirement for multihop transmission type and the cooperative MIMO as in [14] is not considered here. For simplicity, we define the area of the exclusion region consumed by one bit in a certain hop as a single unit of transmission resources. The left side sums up all the transmission resources consumed by each hop of each bit \( b \) in each cluster \( C_j \). Since at most \( WT \) bits can be carried in time \( T \) with bandwidth \( W \), for area \( L^2 \), the upper bound of total transmission resource is \( WT \cdot L^2 = WT^2 \). This inequality cannot only be used in our heterogeneous case, but also in homogeneous case as well since it is derived regardless of node distribution. In static network, there is a minimal transmission range \( r \) to guarantee full network connectivity. In networks 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 \( D(\xi, r) \) is nonzero all the time.

**Lemma 3.2:** Let \( \mathcal{N}(r) \) represent the number of nodes within the transmission range \( r \) when a node wants to transmit its packets, then \( \mathcal{N}(r) = \Theta\left( \frac{n_r + n_p^2}{W^2} \right) \) when \( r = \Omega(\sqrt{n/(\mu p)}) \).

**Proof:** In our cluster dense regime, when \( r = \omega\left( \frac{1}{\sqrt{n/(\mu p)}} \right) \), \( \mathcal{N}(r) \leq 2\pi \sum_{j=1}^{n_c} |C_j|^2 \leq 2\pi n_c \mu p^2 \) 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 \( \mathcal{N}(r) \geq c_n \frac{\pi n_c \mu p^2}{2} \).

Then we come to \( r = \Theta\left( \frac{1}{\sqrt{n/(\mu p)}} \right) \) case. Note that the node density is upper bounded by a HPP with rate \( \mu = \frac{\pi r^2 n}{L^2} \), the probability that the number of nodes inside the disk exceeding \( n_0 \) is

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

During the above derivation, we used the Lagrange form of the remainder term and \( 0 \leq \theta \leq 1 \). Therefore when \( n_0 = \omega(\frac{n_r \mu p^2}{L^2}) \), \( Pr(\mathcal{N}(r) > n_0) = 0 \) 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 \( r = \ell_c^b \geq r = \Omega(\sqrt{\frac{L}{\pi^2 p}}) \). Now we need to characterize the total length for transmitting a bit \( \sum_{h=1}^{b_c} \ell_c^b \). In [4], they prove that nearest neighbor transmission can achieve optimal multicast capacity in uniform traffic distribution. They obtain that the number of hops required is \( \sqrt{k n} \) for \( k \) destinations. The situation is complicated here since traffic is not uniformly distributed across the network. Clients in the same multicast session are 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 the many more destinations involved in a larger transmission range can compensate for the sacrifice of 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 upper bound of minimum capacity in non-heterogeneous networks. To obtain the results under heterogeneous cluster traffic, we first introduce the following lemma, Theorem 1 in [24] as follows.

**Lemma 3.3:** If is the density of the probability function for picking points, then for large n and d ≥ 1, the size of the EMST is approximately \( c(d)n^{d-2} \int_{(0, \ldots , 0)} f(x) \frac{dx}{x^d} \), where \( c(d) \) is a constant depending only on the dimension d.

Our case corresponds to \( d = 2 \) and such that

\[
|EMST(C_j)| = \Theta \left( \sqrt{|C_j|} \int_{(0, \ldots , 0)} \sqrt{\phi(\xi)} d\xi \right).
\]

**Algorithm 1** Generation of \( \rho \)-simplified EMST from EMST.

Input: \( C_j, \phi(\cdot) \)  
Output: \( \rho - C_j \)

1. Specify two point sets \( S \leftarrow \emptyset, S' \leftarrow C_j \).
2. Randomly label each node in \( S' \) with numbers \( 1, 2, \ldots , |C_j| - 1, |C_j| \).
3. Choose nodes with the smallest label number in \( S' \).
4. Add the chosen nodes to \( S \) and discard all the nodes within \( D(\xi', \rho) \) in \( S' \), where \( \xi' \) is the position of the chosen node.
5. Back to step 3 until no node is left in \( S' \), \( (\rho - C_j) \leftarrow S' \).

It reveals that the length of the Euclidean minimum multicast spanning tree constitutes two terms, the square root of the number of nodes \( \sqrt{|C_j|} \) and \( \int_{(0, \ldots , 0)} \sqrt{\phi(\xi)} d\xi \). Note that we eliminate the constant \( c(d) \) for simplicity. Intuitively, the minimum length for connecting all the nodes is the minimum total transmission length (MTTL). For a given graph, MTTL can be counted by summing all the irredundant transmission lengths with extra intermediate nodes and edges. Although MTTL is superficially similar to the EMST problem, the length of MTTL is far less than EMST for the following reasons.

- Larger transmission range can cover more than one node, 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 relays on the MTTL.

Figure 3 illustrates two examples of the above two questions, respectively. In order to answer the first question (derive the relationship between the transmission range \( r \) and MTTL), we construct a new concept \( \rho \)-simplified cluster \( \rho - C_j \). Algorithm 1 illustrates how to generate \( \rho - C_j \) given \( C_j \) and the dispersion density function \( \phi(\cdot) \).

Let \( \rho - T_j \) denote the multicast tree spanning \( \rho - C_j \). The length of each branch in \( \rho - T_j \) is larger than \( \rho \) and all the abandoned nodes are within a distance \( \rho \) from the nodes in \( \rho - C_j \), \( |\rho - T_j| \geq |EMST(\rho - C_j)| \) according to the definition of EMST. After the thinning process following 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
  \[ \tilde{\rho}_j = \sup\{\rho_j, \phi(\rho_j) \geq \frac{1}{\pi \rho^2 |C_j|}\} \]
  for this circular region \( S_1,j \) because \( \phi(\cdot) \) is invariant under rotations. After the thinning process, \( \phi'(\xi) \geq \Theta(\frac{1}{\pi \rho^2 |C_j|}) \) on the basis of Chernoff bound.
- **Sparse Region** \((S_2,j = O/S_1,j)\) Nodes in this region are relatively 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 3.4:** There exists a constant \( c_2 > 0 \) such that

\[
|EMST(\rho - C_j)| \geq c_2 \sqrt{|C_j|} \left( \frac{\sqrt{\pi \rho_j^2}}{\rho} + \int_{S_{2,j}} \sqrt{\phi'(\xi)} d\xi \right).
\]

**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

\[
|EMST(\rho - C_j)| \geq c' \int_{(0, \ldots , 0)} |C_j| |\phi'(\xi)| d\xi
\]

\[
= c' \sqrt{|C_j|} \left( \int_{S_{1,j}} \sqrt{\phi'(\xi)} d\xi + \int_{S_{2,j}} \sqrt{\phi'(\xi)} d\xi \right)
\]

\[
\geq c_2 \sqrt{|C_j|} \left( \frac{\sqrt{\pi \rho_j^2}}{\rho} + \int_{S_{2,j}} \sqrt{\phi'(\xi)} d\xi \right),
\]

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 second question. The following lemma tells us that \( |EMST(\rho - C_j)| \) is at most \( \frac{4 \sqrt{2}}{3} \) times larger than MTTL when relays are utilized, which is proved in [23].

**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 of at least \( q |EMST(U)| \), where \( q = \sqrt{3}/2 \) and \( |EMST(U)| \) is the EMST spanning \( U \).

Then we obtain the lower bound of MTTL for \( C_j \) as

\[
MTTL(C_j) \geq \Theta(|EMST(r - C_j)|).
\]

Based on the above inequality, we obtain a lower bound of \( h_{C_j}^b \) as:

\[
h_{C_j}^b \geq \frac{MTTL(C_j)}{r} \geq \Theta(|EMST(r - C_j)|) \frac{1}{r}.
\]

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 minimum capacity \( \lambda \) 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 networks, the following tradeoffs must be satisfied by any scheduling policies.

\[
\sum_{j=1}^{n_x} \lambda_j \sqrt{|C_j|} \leq c \sqrt{\frac{n_p W L}{\int_0^\infty \sqrt{\phi(\xi)} d\xi}}.
\]

(5)

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

Proof: Using the Cauchy-Schwarz inequality and Eqn.(3), we have

\[
\left( \sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \right)^{1/2} \leq \left( \sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \right)^{1/2} \leq \frac{16WTL^2}{\pi \Delta^2} \left( \sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \right)^{1/2}.
\]

(6)

Because \(r\) is the minimal transmission range in the network, the following inequality is satisfied.

\[
h_j^b \leq \frac{\sqrt{\rho_h^2}}{r}.
\]

(7)

and substitute it into inequality (6), we can obtain that

\[
\sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \leq \frac{16WTL^2}{\pi \Delta^2} \cdot \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2}.
\]

On the other hand, we will get a lower bound on the left side of the above inequality. First we divide \(n_x\) clusters into two sets \(S_1, S_2\) as:

\[
S_1 = \{C_j | \text{Dense Region } S_{1,j} \neq \emptyset\},
\]

\[
S_2 = \{C_j | \text{Dense Region } S_{1,j} = \emptyset\}.
\]

Based on Eqn. (4) and the fact that \(\frac{|EMST(C_j)|}{\sqrt{|C_j|}} = \Theta \left( \int_0^\infty \sqrt{\phi(\xi)} d\xi \right)\)

given in Lemma 3.3, we obtain that if \(S_1 \neq \emptyset\), there exists a constant \(c_4, c_4, c_5 > 0\), such that

\[
\sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \geq \frac{c_4}{r} \sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} |EMST(r - C_j)|
\]

\[
\geq \frac{c_4}{r} \left( \sum_{j \in S_1} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} |EMST(r - C_j)| + \sum_{j \in S_2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} |EMST(r - C_j)| \right)
\]

\[
\geq \frac{c_4}{r} \left( \sum_{j \in S_1} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \right) \psi(\phi(\cdot), S_1)
\]

\[
+ \frac{c_4}{r} \left( \sum_{j \in S_2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \right) \psi(\phi(\cdot), S_1)
\]

\[
\geq \frac{c_5}{r} \left( \sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \right) \psi(\phi(\cdot), S_1),
\]

(8)

where

\[
\psi(\phi(\cdot), S_1) = \min_{C_j \in S_1} \left\{ \frac{|EMST(r - C_j)|}{\sqrt{|C_j|}} \right\}.
\]

Then by substituting Eqn. (7) into Eqn. (8), we can obtain

\[
\sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \leq \frac{4\sqrt{2WTL^2}}{\Delta \sqrt{c_6 c_5 \rho}} \psi(\phi(\cdot), S_1).
\]

(9)

According to the definition of minimum capacity,

\[
\sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \geq \lambda \sum_{j=1}^{n_x} \sqrt{|C_j|} + \sum_{j \in S \notin S_1} \sqrt{|C_j|}
\]

\[
= \lambda \left( \sum_{j \in S_1} \sqrt{|C_j|} + \sum_{j \in S \notin S_1} \sqrt{|C_j|} \right)
\]

(10)

Recalling the previous assumption that the number of clusters in \(S \in S_1\) with size \(\Theta(p)\) is the same order with total cluster numbers \(n_x\), we have \(\sum_{j \in S \notin S_1} \sqrt{|C_j|} = \Theta(n_x \sqrt{p})\). Hence, there exists a constant \(c_6\) such that

\[
\lambda \left( \sum_{j \in S_1} \sqrt{|C_j|} + \sum_{j \in S \notin S_1} \sqrt{|C_j|} \right) \geq c_6 \lambda n_x \sqrt{p}
\]

\[
+ \lambda \sqrt{n_x - 1} (1 - c_5) n_x \geq c_6 \lambda n_x \sqrt{p}.
\]

(11)

Combine Eqn. (9) and Eqn. (10), we obtain

\[
\lambda \leq \frac{4\sqrt{2WTL^2}}{\Delta \sqrt{c_6 c_5 \rho \sqrt{n_x \sqrt{p}}}} \psi(\phi(\cdot), S_1).
\]

(12)

Else if \(S_1 = \emptyset\):

\[
\sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \leq \frac{c_4}{r} \sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} |EMST(r - C_j)|
\]

\[
\geq \frac{c_4}{r} \sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \psi(\phi(\cdot), S_1)
\]

\[
\sum_{j \in S_1} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \psi(\phi(\cdot), S_1)
\]

\[
\geq \frac{c_5}{r} \left( \sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \right) \psi(\phi(\cdot), S_1),
\]

(8)

where

\[
\psi(\phi(\cdot), S_1) = \min_{C_j \in S_1} \left\{ \frac{|EMST(r - C_j)|}{\sqrt{|C_j|}} \right\}.
\]

Then by substituting Eqn. (7) into Eqn. (8), we can obtain

\[
\sum_{j=1}^{n_x} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \lambda_j \frac{T_{h_j}^2}{\rho_h^2} \leq \frac{4\sqrt{2WTL^2}}{\Delta \sqrt{c_6 c_5 \rho \sqrt{n_x \sqrt{p}}}} \psi(\phi(\cdot), S_1).
\]

(9)

During the above derivation, \(r \geq c' \frac{\sqrt{n_x \sqrt{p}}}{\sqrt{c_6 \rho}}\) is a necessary condition to guarantee network connectivity, where \(c' > 0\) is a constant. After comparing the results in the two cases, the
only thing required to do is to prove that there exists a constant $c_7 > 0$ such that
\[
\frac{L \sqrt{n_s p}}{c'} \int_0^1 \sqrt{\phi(\xi)} d\xi \geq c_7 L^2
\]
Recalling Lemma 3.4, it is equivalent to prove
\[
c' c' \int_0^1 \sqrt{\phi(\xi)} d\xi \leq c_7 \sqrt{p} + r \int_0^1 \sqrt{\phi(\xi)} d\xi.
\]
Note that $r \geq c' L / \sqrt{n_s p}$ and $\phi(\xi) \leq \frac{2n}{p L^2} \Theta(1/L^2)$ (See Section III.B), and there exists a constant $c_7$ that can meet the above inequality and $c$ satisfies Eqn. (5).

**Theorem 3.2**: The achievable minimum capacity $\lambda$ in our network is upper bounded as follows:
\[
\lambda \leq \min \left\{ W, \frac{c W L}{\sqrt{n_s} \int_0^1 \sqrt{\phi(\xi)} d\xi} \right\}. \tag{14}
\]

**Proof**: Since we assume that the number of clusters with size $\Theta(p)$ is in the same order with total number of clusters $n_s$, we have
\[
\sum_{j=1}^{n_s} \lambda_j \sqrt{|C_j|} \geq \lambda \sum_{j=1}^{n_s} \sqrt{|C_j|} = \Theta(\lambda n_s \sqrt{p}).
\]
Combining with Theorem 3.1, the term $\sqrt{p}$ can be eliminated.

IV. UPPER BOUND OF MINIMUM CAPACITY WITH DISTRIBUTION VARIANCE CONSTRAINED

In the previous section, we have completed our analysis of the relationship between the upper bound of minimum capacity $\lambda$ and the dispersion density function $\phi(\cdot)$. However, we cannot tell what the impact is on the minimum capacity. Therefore a quantitative value $\sigma_\Omega$ is needed to describe the degree of heterogeneity of each dispersion density function $\phi(\cdot)$. There are two reasons for studying their relationship. First, we cannot see whether heterogeneity increases or decreases minimum capacity $\lambda$ based on Eqn. (14). $\sigma_\Omega$ is a good indicator of the degree of heterogeneity. Second, we can predict the achievable minimum capacity just by knowing the distribution variance. The exact point process is not needed using this model. This is useful because it is often difficult to obtain the dispersion density function.

Recalling the definition of $\sigma_\Omega$, there are various dispersion density functions $\phi(\cdot)$ satisfied given a fixed $\sigma_\Omega$. Below we will show how to find the point process that achieve the largest value of minimum capacity.

**Theorem 4.1**: Given the distribution variance $\sigma_\Omega$, the minimum capacity $\lambda$ is bounded as follows:
\[
\lambda \leq \min \left\{ O(W), O\left(\frac{\max\{1, L\sigma_\Omega\} W}{\sqrt{n_s}}\right) \right\}. \tag{15}
\]
The respect dispersion function is identical to $\phi^{\nu}(\xi)$.

According to Theorem 3.1, a smaller $|EMST(C_j)|$ results in a larger capacity. Therefore we will derive the minimum $|EMST(C_j)|$ given a fixed $\sigma_\Omega$.

**Theorem 4.2**: Define a real variable function $\Upsilon(\phi(\cdot)) = \int_0^1 \sqrt{\phi(\xi)} d\xi$. Then we can prove that $\phi^{\nu}(\cdot)$ can minimize $\Upsilon(\phi(\cdot))$ among all $\phi(\cdot)$ with distribution variance $\sigma_\Omega$.

**Proof**: We will refer to Hölder’s inequality for the proof of this theorem as follows.

Let $\sigma_\Omega^2 = \frac{1}{\pi R^2} - \frac{1}{L}$ and we can obtain that
\[
\int_0^1 \phi^2(\xi) d\xi \leq \frac{1}{\pi R^2}.
\]
Therefore our minimization problem is transformed as follows:
\[
\begin{align*}
& \min \ U(\phi(\cdot)) = \int_0^1 \sqrt{\phi(\xi)} d\xi \\
& \text{s.t.} \quad \int_0^1 \phi(\xi) d\xi = 1 \\
& \quad \int_0^1 \phi^2(\xi) d\xi \leq \frac{1}{\pi R^2} \\
& \quad |\phi(\xi)| \leq \phi(|\xi_1|) \text{ for all } |\xi_1| \geq |\xi_2|
\end{align*}
\]
We substitute $f(\xi) = q^{2/3}(\xi), g(\xi) = \phi^{1/3}(\xi), p = 3$ and $q = 3/2$ into Eqn. (2). Then we can obtain the following inequality in space $\Omega$.
\[
\int_0^1 \sqrt{\phi(\xi)} d\xi \geq \left( \frac{\int_0^1 \phi(\xi) d\xi}{\left(\int_0^1 \phi^2(\xi) d\xi \right)^{1/3}} \right)^{3/2} \geq \sqrt{\pi} R.
\]
Note that $\phi(\cdot)$ is a monotonously decreasing function and the inequality will become an equation when
\[
\phi(\xi) = \begin{cases} 
\frac{1}{\pi R^2} & |\xi| \leq R \\
0 & \text{otherwise}
\end{cases}
\]
which means that $\phi^{\nu}(\cdot)$ can achieve the minimal value of $\Upsilon(\phi(\cdot))$.

Recalling Theorem 3.2, we complete the proof of Theorem 4.1. In this case, the node distribution conforms exactly 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 the next section, we will provide the capacity achieving scheme to approach this bound and verify the upper bound of minimum capacity is achievable in order sense.

V. CAPACITY ACHIEVING SCHEME FOR UNIFORM CLUSTER RANDOM MODEL

We have already derived upper bound of minimum capacity when two types of heterogeneities are involved. In this section, we will provide a capacity achieving scheme for UCRM based on percolation theory. We prove that the minimum capacity achieved by our scheme matches the asymptotic upper bounds when the number of sessions $n_s$ is large enough.

We find that only when the length of the multicast spanning tree $|T_j|$ is on the same order with $|EMST(C_j)|$, the individual multicast capacity can approach the theoretical upper bound in order sense. Intuitively, when the degree of HCT is high enough, the nodes in each cluster are only overlapped.
with a constant number of clusters, which corresponds to 
\[ \sigma_{\mathcal{O}} = \Omega(\frac{\sqrt{n}}{L}) \] . We denote such a case as **trivial cluster overlapping**. On the contrary, when the degree of HCT is relatively low when \( \sigma_{\mathcal{O}} = O(\frac{1}{L}) \). We denote such a case as **fully cluster overlapping**. The network topology between these two extreme cases is denoted by **partial cluster overlapping**. The capacity achieving scheme for uniform cases has been developed in [4] and we discuss the other two cases in this section.

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

In this case, 
\[ R = \frac{L}{\sqrt{\pi(1+L^2(\sigma_{\mathcal{O}})^2)}} = O(\frac{\sqrt{L}}{\sqrt{n}}) \]. 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 \( \sigma_{\mathcal{O}} = O(\frac{1}{L}) \)**

In this case, 
\[ R = \frac{L}{\sqrt{\pi(1+L^2(\sigma_{\mathcal{O}})^2)}} = \Omega(L) \]. All the nodes are fully overlapped and we find that the achievable minimum capacity is identical to the uniform case.

**C. When \( \sigma_{\mathcal{O}} = o(\frac{\sqrt{n}}{L}) \)**

In this case, 
\[ R = \Theta(\frac{1}{\sigma_{\mathcal{O}}}) = \omega(\frac{\sqrt{L}}{\sqrt{n}}) \] and the traffic in each cluster is not so aggregated because \( \sigma_{\mathcal{O}} \) is relatively smaller. In [3], an **information highway** is proposed to approach the capacity upper bound for unicast non-clustered networks based on percolation theory and we apply this concept to our routing scheme. Before we outline the definition of an **information highway**, some useful lemmas are provided.

**Lemma 5.1:** Let \( \mathcal{N}(r) \) denote the number of nodes within a disk of radius \( r \) as \( \Theta(L/\sqrt{n_\ast p}) \). Then if the cluster radius \( R = \Omega(L/\sqrt{n_\ast}) \), the following inequality is satisfied.
\[ \Pr(\mathcal{N}(r) = k) \leq 2 \exp(-\frac{\pi r^2 n_\ast p}{L^2}) \left(\frac{\pi r^2 n_\ast p}{L^2}\right)^k k! \].

**Proof:** Let \( A(d, r_1, r_2) \) denote the overlapping area of two circles of radius \( r_1, r_2 \) with centers of distance \( d \) away, and \( C(r) \) denotes number of nodes within radius \( r \). Note that the distribution of cluster clients is HPP within a circle of radius \( R \), thus we can obtain
\[ \Pr(\mathcal{N}(r) = k) = \sum_{m=0}^{\infty} \frac{\Pr(\mathcal{N}(r) = k|C(r + R) = m)\Pr(C(r + R) = m)}{n_\ast} \]
\[ = \sum_{m=0}^{\infty} \left( \int_{0}^{R+r} \sum_{S} \left( \prod_{i=1}^{m} \left( e^{-\mu_1 v_i} \right) \right) \frac{2x}{(R+r)^2} dx \right) \left( e^{-\mu_2 \frac{m^2}{m!}} \right) \]
\[ \leq \int_{0}^{R+r} \left( e^{-\mu_1 \frac{m_1 k}{k!}} \left( e^{-\mu_2 \frac{m_2}{m!}} \right) \right) \frac{2x}{(R+r)^2} dx \]
\[ \leq \int_{0}^{R+r} \frac{\pi r^2 n_\ast p}{L^2} \left( \frac{\pi r^2 n_\ast p}{L^2} \right)^k k! \]
\[ = \sum_{m=0}^{\infty} \frac{\Pr(\mathcal{N}(r) = k|C(r + R) = m)\Pr(C(r + R) = m)}{n_\ast} \]
\[ \leq 2 \exp(-\frac{\pi r^2 n_\ast p}{L^2}) \left(\frac{\pi r^2 n_\ast p}{L^2}\right)^k k! \].

**Lemma 5.2:** There exists a constant \( \tau \) such that if we partition the square \( \mathcal{O} \) into cells with equal side length \( \tau L/\sqrt{n_\ast p} \), the probability that at least one node resides in a cell is larger than \( \frac{1}{2} \). Now we outline the definition of an information highway:

Choose \( \tau \) large enough so that \( \frac{1}{2} > 5/6 \) and we equally partition the square \( \mathcal{O} \) into cells of edge \( \tau L/\sqrt{n_\ast p} \). This is because \( \frac{\sqrt{n_\ast p}}{\tau} \times \frac{\sqrt{n_\ast p}}{\tau} \) cells and each cell in \( i_{th} \) row and \( j_{th} \) column is denoted by \( s_{i,j} \), and \( s_{i,j} \) 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 4 is an example of percolation path, which represents a wireless backbone that is used to multihop...
data across the network[3]. 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. \( \Theta(n,p) \) disjoint paths crossing from left
to right (top to bottom). A set containing these \( \Theta(n,p) \) horizontal
and vertical paths are called information highways. Note that
the cells composed of the highway are called percolated cells,
and the nodes are called representing nodes. The following
properties are proved in [3] via percolation theory.

- In each horizontal (vertical) rectangle of size \( L \times (\kappa \log(n,p)\frac{L}{\sqrt{n^p}} - \epsilon_L) \),
  there are at least \( \delta \log(n,p) \) horizontal (vertical) highway paths w.h.p. It indicates
  there are \( \Theta(\sqrt{n^p}) \) 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)
  paths is at most \( O(L \log(n,p) / \sqrt{n^p}) \).
- There exists a spatial and temporal scheme that can
  achieve \( O(W) \) throughput on the highway. It means
  that each cell composing the highway can be considered
  linked by optical wires to its neighbors with bandwidth
  \( \Theta(W) \). (Theorem 3)

The information highway is an infinitely large component
such that each node in the deployed region can connect to
it within a hop of length \( O(\frac{L \log(n,p)}{\sqrt{n^p}}) \). Now we need to
construct a multicast spanning tree \( T_j \) for \( C_j \) to route its
packets.

Lemma 5.3: The Prim’s algorithm is utilized to construct
an \( T_j \) for cluster \( C_j \) and we prove that \( |T_j| \leq 6 \sqrt{2}|C_j| R \).

Proof: Prim’s algorithm: Initially, each node is a separate
part, then we iteratively find the shortest edge to compose a
lager part until one part is left. Each member in \( C_j \) is confined
in a disk of radius \( R \). We utilize a square of edge \( 2R \) to cover
the whole circle. At each \( i \)th \( (1 \leq j \leq |C_j|) \) step, there are
\( |C_j| + 1 - i \) parts remaining. We equally partition the square into
\( \frac{\sqrt{|C_j|} + 1 - i}{2} \) cells with edge length \( \frac{2R}{\sqrt{|C_j|} + 1 - i} \), 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 \( \frac{2\sqrt{2}R}{\sqrt{|C_j|} + 1 - i} \). Therefore, the upper bound of \( |T_j| \) is:

\[
|T_j| \leq \sum_{i=1}^{\sqrt{|C_j|}} \frac{2\sqrt{2}R}{\sqrt{|C_j|} + 1 - i}
= 2\sqrt{2}R \left( \sqrt{|C_j|} + \sum_{i=1}^{\sqrt{|C_j|}} \frac{1}{i} + \frac{|C_j|}{\sqrt{|C_j|}} - 2 \right)
\leq 6 \sqrt{2}|C_j| R.
\]

(16)

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 is \( \frac{5}{6} \). The next lemma illustrates
the minimum achievable data rate in the uplink and downlink
phases.

\[^3\kappa \text{ and } \delta \text{ is some constant, } \epsilon_L = o(\log(n,p)\frac{L}{\sqrt{n^p}}) \text{ and is to make} \]
\( \kappa \log(n,p)\frac{L}{\sqrt{n^p}} - \epsilon_L \text{ an integer.} \)

**Algorithm 2** Capacity Achieving Scheme for UCRM

1. **Access Point Mapping**: Establish mappings \( F_h(X) \) and
   \( F_v(X) \) for each node \( X \). Horizontally divide the \( L \times L \)
square into slices of size \( L \times (\kappa \log(n,p)\frac{L}{\sqrt{n^p}} - \epsilon_L) \). Then
   there are at least \( \delta \log(n,p) \) paths in each slice. Denote
   each path in the slice as \( path_i \) (\( 1 \leq i \leq \delta \log(n,p) \)).
   We further divide each slice into \( \delta \log(n,p) \) sub-slices of
   size \( L \times (\frac{\kappa L}{\sqrt{n^p}}) \) each. If node \( X \) is in the \( i \)th sub-slice,
   \( F_h(X) \) denotes the percolated cell on \( path_i \) with the same
   ordinate. Mapping of \( F_v(X) \) is the dual of \( F_h(X) \), and is
   mapped by applying the above algorithm to vertical paths.

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. **Routing Protocol**: A multicast spanning tree \( T_j \) is con-
   structed in Lemma 5.3. Each time slot is divided into 3
   mini-slot and corresponds to 3 phases as in Figure 4. For
   each branch in \( T_j \) linking nodes \( X_1 \) and \( X_2 \), the 3 phases
   are as follows:
   - **Uplink**: \( X_1 \) drains its data to \( F_h(X_1) \).
   - **Highway**: This phase corresponds to step 2 and uti-
     lizes multihop transmission along the horizontal path
     from \( F_h(X_1) \) to the intersection with the vertical path
     in which \( F_v(X_2) \) resides then forwards to \( F_v(X_2) \).
   - **Downlink**: \( X_2 \) downloads the data from \( F_v(X_2) \).

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

Proof: According to Algorithm 2, for each node \( X \), both
\( |F_h(X) - X| \) and \( |F_v(X) - X| \) is upper bounded by
\( \kappa \log(n,p)\frac{L}{\sqrt{n^p}} \times 2\sqrt{2}(R + \frac{2\kappa \log(n,p)}{\sqrt{n^p}} L) \). \( X \) and \( F_h(X) \) must reside in the same cell.
The same is true for \( X \) and \( F_v(X) \). 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 3.2, there
are at most \( 8\kappa^2 \log^2(n,p) \) nodes in each cell, which means
there are at most \( 8\kappa^2 \log^2(n,p) \) \( SD \) pair. Thus, each \( SD \) pairs
can be allocated \( \frac{1}{O(\log^{-2}(n,p))} = \Omega(\log^{-2}(n,p)) \) fraction of a
one time slot for transmission.

Then we begin to analyze the second phase. Although
the highway can be virtually considered as a wired network with
bandwidth \( \Theta(W) \), each path must help relay data for many
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 can serve.

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

Proof: We refer to Figure 5 for the proof. Each path
is constrained within a strip of width \( \kappa \log(n,p)\frac{L}{\sqrt{n^p}} \). Thus
\( D \) is upper bounded by \( (\kappa \log(n,p)\frac{L}{\sqrt{n^p}}) \). Recalling that
the radius of the disk is $R$, we know the farthest cell that can be used is the black cell, which is at a distance $R \leq \sqrt{2(R + \frac{1 + \kappa \log(n,p)}{\sqrt{n,p}})\ell} \leq \sqrt{2(R + \frac{2 \kappa \log(n,p)}{\sqrt{n,p}})}$ away from the kernel.

Fig. 5: Demonstration of influential range and a division of a percolation path

Each branch in the spanning tree $T_j$ 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 $\ell$, then there exists a constant $c_3$, such that $\Pr\left(h(\xi_1^1, \xi_2^1) \leq \frac{c_3 \sqrt{n,p} \log(n,p)}{\tau L} + 4\kappa \log(n,p)\right) \geq 1 - \delta(n,p)$, where $\delta(n,p) \to 0$ as $n,p \to \infty$. We can prove that there exists a constant $c_3$, such that

$$\Pr\left(h(\xi_1^1, \xi_2^1) \leq \frac{c_3 \sqrt{n,p} \log(n,p)}{\tau L} + 4\kappa \log(n,p)\right) \geq 1 - \delta(n,p), \quad (17)$$

where $\lim_{n,p \to \infty} \delta(n,p) = 0$. Now we divide the path into $K = \frac{\ell}{\ell_1}$ sub-paths as in the second figure in Figure. 5 and $\ell_i = \frac{L}{\ell} \log(n,p)$ (1 \leq i \leq K) denotes the length of the $i^{th}$ sub-path. Then the proof of Eqn. (17) is as follows.

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

which is also identical to prove

$$\lim_{n,p \to \infty} \Pr(\varphi(n,p, i) = \infty) = 0.$$

For each path on the highway system, the length of the path 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 expected value on both sides, we can obtain

$$\sum_{i=1}^{K} \Pr(\ell = \omega(\frac{L}{K})) \leq K \Pr(\ell = \omega(\frac{L}{K})) \leq c'L.$$

Then we can obtain

$$\Pr(\ell = \omega(\frac{L}{K})) \leq \frac{c'}{\inf_{\varphi(n,p, i) = \infty} \{\varphi(n,p, i)\}}$$

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

Now assume that a sender at $\xi^s$ generates traffic to a receiver $\xi^t$ at a distance $\ell$ away. Let $\xi^1$ be the intersection of the horizontal and vertical path in which $\mathcal{F}_h(\xi^s)$ and $\mathcal{F}_v(\xi^t)$ reside. According to the properties of percolation paths, every percolation path is constrained within a strip of width $\kappa \log(n,p)$ that

$$\left\{ \begin{array}{l l} 
\max\{|[\mathcal{F}_h(\xi^s)]_a - [\mathcal{F}_v(\xi^t)]_a| \leq |\xi^a - \xi^a_1| + \frac{\kappa L \log(n,p)}{\sqrt{n,p}}, \\
\max\{|[\mathcal{F}_v(\xi^t)]_o - [\mathcal{F}_h(\xi^s)]_o| \leq |\xi^o - \xi^o_1| + \frac{\kappa L \log(n,p)}{\sqrt{n,p}}, \\
\max\{|\xi^o - \xi^o_1|^2 + (\xi^a - \xi^a_1)^2| \leq \ell + \frac{2\kappa L \log(n,p)}{\sqrt{n,p}}. 
\end{array} \right. \quad (18)$$

Then the number of hops required for transmission with nodes $\ell$ away is upper bounded by

$$h(\mathcal{F}_h(\xi^s), \xi^t) + h(\xi^t, \mathcal{F}_v(\xi^t)) \leq c_3 \frac{\sqrt{n,p}}{\tau L}( |[\mathcal{F}_h(\xi^s)]_a - [\mathcal{F}_v(\xi^t)]_a| + |[\mathcal{F}_w(\xi^t)]_o - [\mathcal{F}_h(\xi^s)]_o| ) \leq c_3 \left( \sqrt{n,p} \frac{\ell}{L} + 4\kappa \log(n,p) \right).$$

**Lemma 5.7:** Assume that cell $s$ is within a distance $R$ from $C_j$'s head. Then if $R \geq \frac{\kappa \log(n,p)}{\sqrt{n,p}}$, the probability $\mathcal{P}$ that $s$ is utilized to transmit for cluster $C_j$ is upper bounded as:

$$\mathcal{P} \leq \frac{12c_3 \delta \ell L \sqrt{|C_j|}}{\kappa \sqrt{n,p} R}. \quad (18)$$

**Proof:** There are at least $\frac{\kappa}{4} \left( \frac{\sqrt{n,p} R}{\tau L} \right)^2 \geq \frac{\kappa}{4} \left( \frac{\sqrt{n,p} R}{\tau L} - 1 \right)^2$ percolated cells within the disk of radius $R$. 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:

$$\mathcal{P} \leq \sum_{j=1}^{|C|} \mathcal{P}(\mathcal{I}(s,i) = 1) \leq \sum_{j=1}^{\frac{\kappa}{4}} \frac{c_3 \ell L^2}{2 \kappa n_p R^2} \left( \frac{4 \sqrt{n,p} R}{\ell} + 4\kappa \log(n,p) \right) \leq \frac{c_3 \delta \ell L^2}{2 \kappa n_p R^2} \left( \frac{12 \sqrt{n,p} |C_j|}{R} + 4\kappa |C_j| \log(n,p) \right) \leq \frac{12c_3 \delta \ell L \sqrt{|C_j|}}{\kappa \sqrt{n,p} R}. \quad (18)$$
The above lemma provides an upper bound on the probability that a multicast flow from a cluster $C_j$ is routed via the cell when it is within the influential range of that cluster. Let us now study the total number of times that the cell is used by our capacity achieving scheme. We know that $|C_j| = n^{1-\alpha_i}$ if $C_j \in SC_i$. Let $P_i \leq \frac{2d_{\delta/4} L \sqrt{\log n}}{8} \leq \frac{d_{\delta/4} L}{8}$ denote the probability that a cell is used for $C_j$ if $C_j \in SC_i$ and $s$ is within the influential range $R$ of $C_j$’s head. For each cluster $C_j$, $|C_j|$ 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 multicasting spanning tree is $O(\log p)$ according to [4].

**Theorem 5.1: (VC-Theorem):** If $S$ is a set of finite VC-dimension $VC-d(S)$, 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 $\epsilon, \delta > 0$,

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

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

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

$$
\Pr\left(\sup_{x \in O} \left| \frac{\mathcal{F}_i(s)}{N} - P_i \right| \leq \epsilon(n)\right) > 1 - \delta(n),
$$

where $\mathcal{F}_i(s)$ is the number of flows using $s$ by clusters belonging to $SC_i$ and $\delta(n)$ is a set of sequences approaching 0 when $n$ goes to infinity. Eqn. (21) should be satisfied when $2 \leq i \leq k$ and Eqn. (22) should be satisfied when $i = 1$.

**Proof:** The method to construct multicasts are i.i.d. variables within a disk of radius $R$. Then the multicast trees can be viewed as i.i.d variables. Therefore, we can use the VC-Theorem. Recalling that, given a cell $s$, the probability that a multicast tree will cross at most $\frac{12c_{\delta/4} L \sqrt{|C_j|}}{k \sqrt{n} \sqrt{R}}$ when the cell is within its influential range $R$. Hence, the whole set of the cell $O$:

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

when $N \geq \max\left\{\frac{8d_{\delta/4} L \sqrt{\log n}}{\epsilon} \log \frac{13}{\epsilon} + \frac{4}{\epsilon} \log \frac{2}{\delta}\right\}$, where $d = O(\log p)$ is the VC-dimension. Then substitute Eqn. (18) into it and note that $|C_j| = \Theta(n^{1-\alpha_i})$ and we obtain

$$
\Pr\left(\sup_{x \in O} \left| \frac{\mathcal{F}_i(s)}{N} \right| \leq \frac{12c_{\delta/4} L \sqrt{|C_j|}}{k \sqrt{n} \sqrt{R}} + \epsilon(n)\right) > 1 - \delta(n).\tag{19}
$$

Now let $\epsilon(n) = \frac{12c_{\delta/4} L \sqrt{|C_j|}}{k \sqrt{n} \sqrt{R}}$ and $\delta(n) = \frac{2}{n}$ and for inequality (19) to be true,

$$
N \geq \max\left\{\frac{8d_{\delta/4} L \sqrt{\log n}}{\epsilon(n)} \log \frac{13}{\epsilon(n)} + \frac{4}{\epsilon(n)} \log \frac{2}{\delta(n)}\right\}
= \frac{k R n^{\alpha_i/2} \log n}{3c_3 \delta \tau L},\tag{20}
$$

is required. In this case, $N$ denotes the number of clusters belonging to $SC_i (2 \leq i \leq k)$ within the disk of radius $R$. Applying the same technique as Lemma 3.2, we can obtain that $N \in \left(\frac{2n R^2}{h^2}, \frac{2n R^2}{h^2}\right)$ if $R = \omega(\sqrt{\frac{1}{n^2}})$. Therefore when

$$
R \geq \frac{2\kappa k Ln^{\alpha_i/2} \log n}{3\pi c_3 \delta \tau n}, = \frac{2\kappa k Ln^{\alpha_i/2-\alpha} \log n}{3\pi c_3 \delta \tau},\tag{21}
$$

Eqn. (20) is satisfied. Substitute $N \leq \frac{2n R^2}{h^2}$ into Eqn. (19) and we complete the proof when $2 \leq i \leq k$. A similar technique is employed for $SC_1$ if the following condition is satisfied.

$$
R \geq \frac{2\kappa \ln L \alpha_i/2 \log n}{3\pi c_3 \delta \tau n} = \frac{2\kappa L \alpha_i/2 \log n}{3\pi c_3 \delta \tau \sqrt{n^2},}\tag{22}
$$

The above constraint of $R$ means that only when $R$ is sufficiently 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 $SC_i$ is upper bounded by $\Delta$, $\Delta$ is also an upper bound of the number of flows in $SC_i (i' \leq i \leq k)$ because of $\alpha_i$ is an increasing sequence over $i$.

**Lemma 5.8:** Given a cluster radius $R = \omega(L/\sqrt{n^2})$, the data rate that the highway can sustain for $C_j$ is denoted by $\lambda^h_j$. When $R = O(\sqrt{\frac{1}{n^2}})$,

$$
\lambda^h_j = \Omega\left(\frac{L}{R \sqrt{n^2}}\right) \quad 1 \leq j \leq k,
$$

when $R = \omega(\sqrt{\frac{1}{n^2}})$,

$$
\lambda^h_j = \begin{cases}
\Omega\left(\frac{L \alpha_i/2}{n^{\alpha_i}}\right) & C_j \in SC_i, \quad 1 \leq i < k', \\
\Omega\left(\frac{L \alpha_i/2}{n^{\alpha_i}}\right) & C_j \in SC_i, \quad k' \leq i < k.
\end{cases}\tag{23}
$$

**Proof:** In case of $R = O(\sqrt{\frac{1}{n^2}})$, there are at most $\Omega(\frac{R^2}{h^2}) \leq O(k)$ 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 cell is upper bounded by $\Theta(\frac{R^2}{h^2})$ w.h.p according to Theorem 5.2. Therefore $h = O(\sqrt{\frac{R^2}{h^2}})$ in our FDM network and the highway can sustain a data rate of $W/h = \Omega(\sqrt{\frac{1}{n^2}})$.

When $R = \omega(\sqrt{\frac{1}{n^2}})$, the number of clusters belonging to $SC_i (2 \leq i \leq k)$ is $\Theta(\frac{R^2}{h^2})$ and $h = k$ which means each $SC_i (2 \leq i \leq k)$ is allocated $(1-c_0)W/k$ bandwidth and $SC_1$ is allocated $c_0 W$ bandwidth for transmission in the highway. We specify a $k'$ as follows:

$$
k' = \arg\sup_{1 \leq i \leq k} \left\{\frac{2\kappa}{3\pi c_3 \delta \tau} \frac{k \ln n^{\alpha_i/2-\alpha} \log n \leq R}\right\}.\tag{24}
$$

Then we know the number of flows across each percolation cell $s$ can be upper bounded by $\frac{2\kappa}{3\pi c_3 \delta \tau} \frac{k \ln n^{\alpha_i/2-\alpha} \log n \leq R}$ for clusters belonging to an arbitrary super cluster $SC_i (1 \leq i \leq k')$. Therefore

$$
\lambda^h_j \geq \frac{W/k}{24\pi c_3 \delta \tau} \frac{L \alpha_i/2}{k L \ln n^2} = \frac{\kappa W \ln n^{\alpha_i/2}}{24\pi c_3 \delta \tau R n^2} = \Omega\left(\frac{L \ln n^{\alpha_i/2}}{R n^2}\right).
$$
For \( C_j \in SC_i \) \((k' + 1 \leq i \leq k)\), the number of flows crossing a cell \( s \) is also upper bounded by \( \frac{24m \sigma}{\sqrt{n} \frac{n^\alpha}{L}} \), which is the same as that of clusters belonging to \( SC_i \) and in this case \( \lambda_j^h = \Omega(\frac{\ln m}{\ln n}) \).

Recalling Lemma 5.4, we know the bottleneck is due to data delivery on the highway when \( R \geq \Omega(\frac{L \log n}{n}) \). Let \( \lambda_j^h, \lambda_j^d \) denote the data rate of the uplink and downlink, respectively, and \( \lambda_j = \min(\lambda_j^h, \lambda_j^d) \). The achievable upper bound of minimum capacity \( \lambda \) 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 same maximum capacity as Eqn. (15) in order sense if we substitute \( R = \frac{L}{\sqrt{\pi(1+L^2\sigma_C)}} \) into \( \lambda \) and \( \lambda = \Omega(\max\{\frac{1}{\sqrt{n_s}}, \frac{L\sigma_C}{\sqrt{n_s}}\}) \).

VI. DISCUSSION

Motivated by that max-min fairness framework is applicable for multicast traffic [20] and the minimum multicast capacity can be efficiently improved, we hope to find out the upper limit for the minimum capacity to attain. Investigating the minimum capacity can provide with insights into the worst situation of the whole network and solving the problem for minimum capacity can significantly benefit the whole network performance. Different from most previous literature, which mainly emphasized on maximizing the total throughput (or per-cluster throughput), our paper chooses another aspect in evaluating the network performance. By combining the max-min metric and heterogeneity property, we provide a novel perspective concerning the problem of network capacity.

Our results can generalize various results obtained in non-heterogeneous network like [3], [4], [6]–[8]. Here we highlight the feature in our clustered network, heterogeneous cluster traffic, which is characterized in real networks with multicast traffic patterns. Moreover, our results for networks with HCT heterogeneous traffic can be effectively applied to large scale networks such as wireless sensor networks [26].

![Graphical representation](image)

Fig. 6: Relationship between minimum capacity \( \lambda \) and \( \sigma_C \) (log scale). In region 2, the minimum capacity increases with \( \sigma_C \). \( \sigma_C = 0 \) means a non-heterogeneous network with the same achievable capacity as prior works.

HCT characterizes a network property that clients of each multicast session are not uniformly distributed across the network. Previous works are insufficient to estimate the achievable capacity. Indeed, Eqn. (5) is a precise formula for the achievable capacity. However, we cannot decide whether HCT increases the minimum capacity because there is no criteria to judge the level of heterogeneity for a dispersion density function. By introducing the variable distribution variance \( \sigma_C \), we offer a quantitative description of the extent of heterogeneous traffic. We find that if the traffic of a cluster is not uniformly disseminated across the region as is assumed in prior works, the minimum capacity \( \lambda \) increases.

Figure 6 illustrates the relationship between \( \lambda \) and \( \sigma_C \) derived from Eqn. (15). In region 1, HCT is relatively slight and each cluster is full overlap with other clusters. Fully overlapping indicates that in each small region \( \mathcal{O}' \subseteq \mathcal{O} \) with area \( \Theta(L^2/n) \), there are nearly \( \Theta(n^\alpha) \) clusters whose members have approximately equal probability to reside in \( \mathcal{O}' \). Therefore it is much like the uniform distributed cases and the minimum 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 \( \Theta(n^\alpha) \) clusters whose members have approximately equal probability to reside in \( \mathcal{O}' \). This case, \( \theta < \alpha \) and \( \theta \) is closely related to \( \sigma \). And the implicit reason for an increased minimum capacity is that each relay only needs to deliver packets for a smaller portion of the clusters compared to the 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 towards \( O(W) \).

VII. CONCLUSION AND FUTURE WORK

In this paper, we study the effect of heterogeneity on the asymptotic multicast capacity in clustered networks based on max-min fairness. Our contributions are mainly divided into two parts. First, we find that heterogeneous cluster traffic increases the minimum capacity among all the clusters. Through analyzing the fundamental constraints of wireless networks from global and local aspects, a quantitative relationship is provided between minimum capacity \( \lambda \) and distribution variance \( \sigma_C \). This is the first time this relationship is used to describe heterogeneity in literature. The effect of heterogeneity is also analyzed in uniform cluster random model, which is the optimal network layout given a fixed \( \sigma_C \).

There are several tasks for us to complete in the future. First, the cluster dense model should be studied. The cluster dense model indicates that the distance between cluster heads approaches 0 when \( n \to \infty \). Second, the dispersion density function \( \phi(\cdot) \) for each cluster is assumed identical. It would be interesting to study networks with clusters of different \( \phi(\cdot) \). Third, we mainly discuss the heterogeneous cluster traffic, while other aspects of heterogeneity such as heterogeneous cluster size (HCS) is not sufficiently investigated. If the clusters are formed with more diversity in size, it is meaningful to see whether HCS can increase the capacity or not.
VIII. ACKNOWLEDGMENT

This paper is supported by National Fundamental Research Grant (No. 2011CB302701); NSF China (No. 61271219, 61202373); Shanghai Basic Research Key Project (No. 11JC140510); China Ministry of Education Fok Ying Tung Fund (No. 122002); China Ministry of Education New Century Excellent Talent (No. NCET-10-0580).

REFERENCES


Xiyuan Li is currently pursuing her B.E degree in Electronic Engineering at Shanghai Jiao Tong University, China and working with Prof. Xingbing Wang in Institute of Wireless Communication Technology of Shanghai Jiao Tong University.
Her research of interests are in the area of scaling laws analysis and performance evaluation in wireless networks.

Qiuuy Peng received the BE degree in electronic engineering at Shanghai Jiao Tong University, China.
His research interests are in the area of asymptotic analysis of capacity and connectivity in wireless ad-hoc network.

Xinbing Wang received the B.S. degree (with hons.) from the Department of Automation, Shanghai Jiao-tong University, Shanghai, China, in 1998, and the M.S. degree from the Department of Computer Science and Technology, Tsinghua University, Beijing, China, in 2001. He received the Ph.D. degree, major in the Department of electrical and Computer Engineering, minor in the Department of Mathematics, North Carolina State University, Raleigh, in 2006. Currently, he is a faculty member in the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China. His research interests include scaling law of wireless networks and cognitive radio. Dr. Wang has been an associate editor for IEEE Transactions on Mobile Computing, and the member of the Technical Program Committees of several conferences including ACM MobiCom 2012, ACM MobiHoc 2012, IEEE INFOCOM 2009-2013.