

On The Topology of Wireless Sensor Networks

Sen Yang, Xingbing Wang, Luoyi Fu

Dept. of Electrical Engineering

Shanghai Jiao Tong University, China

Email: {twood, xwang8}@sjtu.edu.cn, yiluofu@gmail.com

Abstract—In this paper, we explore methods to generate optimal network topologies for wireless sensor networks (WSNs) with and without obstacles. Specifically, we investigate a dense network with n sensor nodes and $m = n^b$ ($0 < b < 1$) helping nodes, and assess the impact of topology on its throughput capacity. For networks without obstacles, we find that uniformly distributed sensor nodes and regularly distributed helping nodes have some advantages in improving the throughput capacity. We also explore properties of networks composed of some isomorphic sub-networks. For networks with obstacles, we assume there are $M = \Theta(n^v)$ ($0 < v \leq 1$) arbitrarily or randomly distributed obstacles, which block cells they are located in, i.e., sensor nodes cannot be placed in these cells and nodes' communication cannot cross them directly. We find that the overall throughput capacity is bounded by the transmission burden in areas around these blocked cells and introduce a novel algorithm of complexity $O(M)$ to generate optimal sensor nodes' topologies for any given obstacles' distributions. We further analyze its performance for regularly distributed obstacles, which can be taken to estimate the lower bound of the algorithm's performance.

I. INTRODUCTION

Capacity is a fundamental issue in wireless sensor networks (WSNs). A typical WSN involves little or no infrastructure and sensor nodes may communicate in an ad hoc manner. In Gupta and Kumar's seminal work, they adopt Protocol and Physical Model to describe a successful transmission and show that the per-node throughput capacity scales as $\Theta(1/\sqrt{n \log n})$ in random networks, and the per-node transport capacity scales as $\Theta(1/\sqrt{n})$ in arbitrary networks, respectively [1]. These results provide us not only a theoretical bound but also a foundation in the network optimization and protocol design. Following their work, extensive research are conducted to understand the scaling laws in wireless sensor networks better. On the other hand, in some applications, helping nodes are introduced to improve the performance, which results in a heterogeneous network. In heterogeneous networks, access control is studied in [12], [13], routing protocol is studied in [7], [8], [10], N. Li et al. studied topology control in [15], P. Li et al. studied the throughput capacity of networks with rectangular areas in [16]. Many other schemes such as multicast, multiple-input multiple-output (MIMO), hierarchical regime and adding infrastructure are also explored in literatures to improve the network capacity.

However, most of works above are for networks with regularly or uniformly distributed sensor nodes. While in practice, sensor nodes may not be placed uniformly, which could have a huge impact on network properties, including the capacity. For example, if lots of nodes are confined in a small

region, it would lead to great interference and deteriorates the capacity. Also, if nodes are too sparse in a particular area, communication might get difficult, which also harms networks' performance. To the best of our knowledge, only a few works have dealt with the capacity of networks with inhomogeneous node density. In [17], [18], [19], [20], [21], capacity of inhomogeneous clustered networks are analyzed. Corresponding scheduling and routing schemes to approach the upper bounds are discussed in [22].

On the other hand, almost all the previous works dealt with flat network region. While in practice, sensor networks are often deployed in complex environments, such as battle fields or mountainous areas, and there are often many obstacles distributed in these regions. These obstacles may constrain the distribution of sensor nodes and the transmission of packets. For example, in a building monitoring WSN, electromagnetic wave signal can be attenuated significantly when passing through furniture, walls or floors, which could have a great impact on network performance. Another example is WSNs deployed in a mountainous area, in which both routing strategy design and deployment of sensor nodes should consider the constraint of the complex landform. Generally, obstacles have a negative impact on the network capacity. However, if we design the network topology appropriately, it could lead to a noneligible improvement. For example, in building monitoring WSNs, capacity can be improved if we place less nodes in areas shadowed by obstacles. Also, in a mountainous region, if we deploy more nodes in open areas, network capacity can be much larger than that we put most of them in valleys or laps.

These motivate us to explore better network topologies for given network regions, especially for networks with obstacles. In this paper, we investigate how nodes' spatial distributions influence the throughput capacity and explore the optimal nodes distribution on given conditions. We firstly derive some useful conclusions on generating the optimal topology for flat network areas. For networks with obstacles, it's difficult to derive a general solution for various obstacles distributions. However, a feasible algorithm with linear complexity can be proposed by dividing the whole network region into some small pieces and dealing with them respectively.

Our main contributions are as follows:

- For networks which consist of many isomorphic sub-networks, compared with the topology of sub-networks, the overall network's topology results in a larger throughput capacity.

- For networks without helping nodes, uniform sensor nodes' distribution is order optimal on maximizing throughput capacity.
- For networks with uniformly distributed sensor nodes, we find that regularly distributed helping nodes are optimal to maximize the network throughput capacity.
- For networks with non-uniformly distributed sensor nodes, though regularly distributed helping nodes are no longer optimal, any improvement on helping nodes' distribution cannot change the throughput capacity on scale.
- For networks with obstacles, we introduce a novel algorithm of complexity $O(M)$ to generate the optimal sensor nodes' topology for any given obstacles' distribution.

The rest of the paper is organized as follows. Section II gives the network model. In Section III, we study the connectivity of networks with different nodes' distributions. In Section IV, we derive the throughput capacity of networks with different topologies. In Section V and VI, we explore some general properties of network topologies. In Section VII, we investigate the throughput capacity of networks with obstacles and introduce an algorithm to generate the optimal user nodes' topology for any given obstacles' distribution. We finally conclude this paper in section VIII.

II. NETWORK MODEL

In this section, we introduce the heterogeneous wireless network model, definition of obstacles, routing strategy and scheduling scheme.

A. Network Components

A heterogeneous wireless network is a dense network with n user nodes and $m = n^b$ ($0 < b < 1$) helping nodes. Here, instead of symmetric traffic pattern, we assume that the network has asymmetric traffic as that defined in [16], i.e., all the n user nodes are sources while only n^d ($0 < d < 1$) user nodes are randomly chosen as destinations. Also, user nodes can serve as relays if needed. On the other hand, helping nodes do not have information to transmit or receive and they only help relay packets from user nodes. According to whether user nodes' packets are forwarded by helping nodes, we divide network traffic into two modes, namely, user mode and helping mode. In user mode, packets are forwarded only by user nodes. While in helping mode, packets are firstly transmitted to the nearest helping nodes, and then forwarded to intended destinations in the helping network. Meanwhile, we assume that all the user nodes have a total bandwidth of 1 and split it into three parts as follow

$$W_1 + W_2 + W_3 = 1$$

where W_1 , W_2 and W_3 are for ad hoc transmissions in user mode, uplink transmissions in helping mode and downlink transmissions in helping mode, respectively. Besides, we assume that ad hoc transmissions in helping network have an independent bandwidth of $W_4 = \Omega(1)$.

B. Definition of Obstacles

To describe networks with obstacles, we assume the network area is partitioned into $\mathcal{K} = \Theta(n^w)$ ($0 < w \leq 1$) cells. When there is no user node distributed in a cell, we assume at the cell's center there is a relay working in the same bandwidth as user nodes, which keeps the network's connectivity. Assume there are $M = \Theta(n^v)$ ($0 < v \leq w$) number of obstacle nodes in the network area, which can be arbitrarily or randomly distributed. Cells are blocked when there are obstacle nodes in them. Here, "blocked" has two implications: no user node can be distributed in blocked cells and communications between user nodes cannot cross them.

C. Interference Model

To bound the interference between different nodes, we suppose the system is based on a cellular network model. Assume that the network is an unit square and we divide it into non-overlapping cells with equal size. Nodes can communicate with each other only when they are in the adjacent cells. Furthermore, we assume that communications between different cells has taken time division multiplexing (TDMA) scheme. Therefore, to avoid interference between adjacent cells, we adopt a rotating scheduling scheme as that described in [16]. Thus, at the same time, in all of the adjacent cells there is at most one that can transmit or receive packets and each cell has the same opportunity to be active.

Following the power propagation model introduced in [23], the reception power at node X_j of the signal from node X_i is

$$P_{ij} = C \frac{P_i}{d_{ij}^\alpha} \quad (1)$$

where d_{ij} is the distance between node X_i and node X_j and P_i is the power emitted by node X_i . According to Shannon Theorem, the achievable transmission rate R_{ij} from node X_i to node X_j is:

$$R_{ij} = W \log(1 + SINR_{ij}) \quad (2)$$

where W is the channel bandwidth, and $SINR_{ij} = \frac{C \frac{P_i}{d_{ij}^\alpha}}{N + \sum_{k \neq i} C \frac{P_k}{d_{kj}^\alpha}}$ is the Signal-to-Interference and Noise Ratio of the transmission from node X_i to node X_j . In this paper we assume that all of the user nodes and all of the helping nodes adopt the same transmission power, respectively. As it derived in [16], we have the following lemma.

Lemma 1. *Each cell in the network can transmit at a transmission rate $c_1 W_1$, where c_1 is a deterministic positive constant.*

D. Routing Strategy

As user nodes can only communicate with nodes in neighboring cells, packets from source nodes may need to be forwarded through multi-hop transmissions to reach destination nodes. Thus, for networks with and without obstacles, we adopt following routing strategies, respectively.

Routing Strategy I - for networks without obstacles:
Suppose a source node is located in cell S_i and its destination

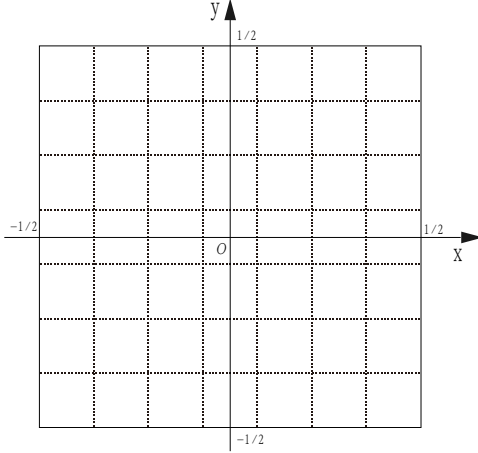


Fig. 1. Coordinate system for the network region

is located in cell S_j . Packet sent from the source node firstly be forwarded along the cells in the same vertical line of cell S_i until it get the cell in the same horizontal line of cell S_j , then the packet is forwarded along the cells in the same horizontal line of cell S_j until it reaches the destination node.

Routing Strategy II - for networks with obstacles:

- 1) If packet sent from the source node can be relayed to its destination by Routing Strategy I, do it.
- 2) Otherwise, if there are only convex obstacles polygons, firstly forward the packet along the routing path generated by Routing Strategy I. When it can no longer be forwarded in current direction (vertical or horizontal), change the forwarding direction to another one (horizontal or vertical). Repeat this until it arrives at the destination node.
- 3) If there exist concave polygons obstacles and neither of the source node and the destination node are in the groove of a concave obstacles polygon, replace the concave obstacles polygons by their convex hulls, respectively, and then following Step 1) and 2) to forward the packet.
- 4) If source and destination nodes (or either of them) are in the grooves of concave obstacles polygons, we can find cells outside the corresponding convex hulls and nearest to source and destination nodes, respectively. Denote them by S_A and S_B , respectively. We firstly transmit the packet from the source node to cell S_A , then following Step 1), 2) and 3) to forward this packet from cell S_A to cell S_B , and finally we forward the packet from cell S_B to the destination node.

E. Network Topology

We first established coordinate system for the network region. As shown in Fig. 1, the coordinate origin is located at the center of the network. As the edge length of the network region is 1, the maximal scales of x-axis and y-axis are both 1/2.

1) *Uniform Distribution:* For networks with uniformly distributed nodes, the probabilities that nodes located at any place

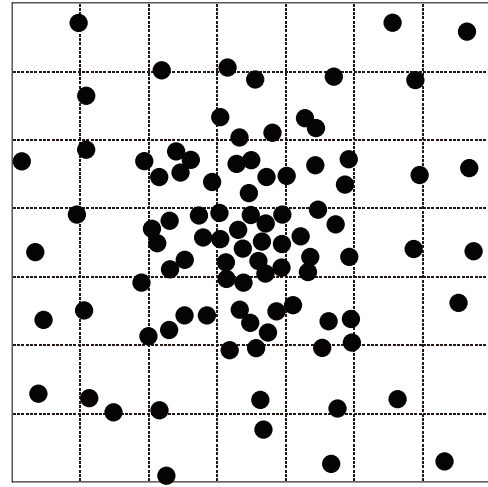


Fig. 2. Centralized distribution networks

of the network is the same with each other, i.e., it has following probability density function:

$$\begin{cases} f(x) = 1 & (-\frac{1}{2} < x < \frac{1}{2}) \\ f(y) = 1 & (-\frac{1}{2} < y < \frac{1}{2}) \end{cases} \quad (3)$$

2) *Centralized Distribution:* We first consider a simple case of the non-uniform distribution. As shown in Fig. 2, nodes density is large in the center of the network and small at the edge. We call it “centralized distribution”. One of its possible probability density functions can be shown as follows:

$$\begin{cases} f(x) = (4a - 4)x + 2 - a & (-\frac{1}{2} < x < \frac{1}{2}) \\ f(y) = (4a - 4)y + 2 - a & (-\frac{1}{2} < y < \frac{1}{2}) \end{cases} \quad (4)$$

where a ($0 \leq a \leq 1$) is centralization coefficient and its value determines the extension that network nodes aggregate to the center. The larger a is, the more uniform the nodes are distributed, or vice versa. In particular, when $a = 1$ are uniformly distributed; when $a = 0$ probability of nodes distributed at the edge of the network would become 0.

3) *Multi-centralized Distribution:* In practice, network nodes often aggregate in several locations of the network, not just the center of the network. We call it “multi-centralized distribution”. As shown in Fig. 3, We can divide the whole network into many small sub-network according to the aggregation centers and each sub-network has similar network topology. In this paper, we assume that all the sub-network is a small centralized distribution network as defined before, i.e., it satisfied the probability density function shows in (4).

III. NETWORK CONNECTIVITY

In this section, we analyze the smallest cell size that can keep the network connectivity.

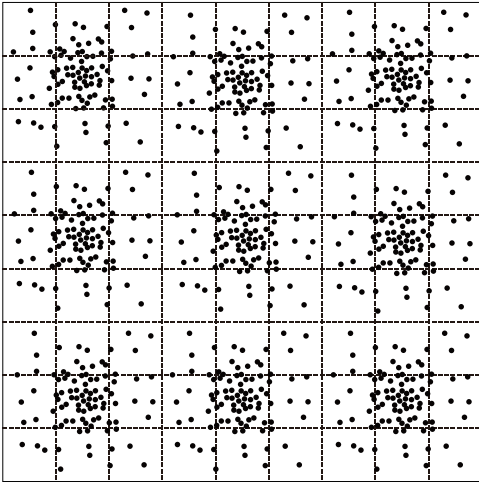


Fig. 3. Multi-centralized distribution networks

A. Uniform Network

For uniformly distributed nodes, network connectivity has been adequately researched in various literatures, here we only give some brief results.

Lemma 2. For uniformly distributed nodes, if we divide the network area into cells of length $l = \sqrt{\frac{c_2 \log n}{n}}$, where $c_2 > 1$ is a constant, each cell has at least one node w.h.p. [24].

B. Centralized and Multi-centralized Network

For multi-centralized network, because each of its sub-network is a small centralized network, we only need to discuss the connectivity of centralized distribution network.

For centralized network, we have the following lemma:

Lemma 3. For centralized, if cell's length satisfies the following condition, the network is connected w.p.h..

$$l \geq \begin{cases} \frac{-a + \sqrt{a^2 + 8(1-a)\sqrt{\frac{c_3 \log n}{n}}}}{4(1-a)} & \text{when } 0 \leq a < 1 \\ \sqrt{\frac{c_3 \log n}{n}} & \text{when } a = 1 \end{cases} \quad (5)$$

where c_3 is a deterministic constant.

Proof: Different from uniform network, in non-uniform network, probabilities that nodes distributed in different cells are not the same. For network satisfying probability density function (4), if there is at least one node in cells located in the corners, there must be at least one node in every cell w.h.p..

For centralized network, the joint probability density function is

$$f(x, y) = [(4a - 4)x + 2 - a] \cdot [(4a - 4)x + 2 - a] \quad (6)$$

where $-1/2 < x < 1/2$ and $-1/2 < y < 1/2$. Without loss of generality, we consider the cell S in the the upper right

corner, for a user node X_i , the probability that it is in cell S , is

$$\begin{aligned} P_i &= \int_{\frac{1}{2}-1}^{\frac{1}{2}} \int_{\frac{1}{2}-1}^{\frac{1}{2}} [(4a - 4)x + 2 - a] \\ &\quad \cdot [(4a - 4)x + 2 - a] dx dy \\ &= ((2a - 2)x^2 + (2 - a)x) \Big|_{\frac{1}{2}-1}^{\frac{1}{2}} \\ &\quad \cdot ((2a - 2)y^2 + (2 - a)y) \Big|_{\frac{1}{2}-1}^{\frac{1}{2}} \\ &= (al - 2(a - l)l^2)^2 \end{aligned}$$

Denote the probability that there is at least one node in cell S by P_S , when $n \rightarrow \infty$, we have

$$P_S = 1 - (1 - P_i)^n \rightarrow 1 - e^{-nP_i} \quad (7)$$

when P_i satisfy

$$P_i \leq \frac{c_3 \log n}{n} \quad (8)$$

we can get

$$P_S \leq 1 - \frac{1}{n^2} \quad (9)$$

i.e., $P_S \rightarrow 1$ as $n \rightarrow \infty$. Put (8) into (7), we can get

$$(al - 2(a - l)l^2)^2 \leq \frac{c_3 \log n}{n} \quad (10)$$

Solving this function and considering that $0 < l < 1$, we can obtain

$$l \geq \begin{cases} \frac{-a + \sqrt{a^2 + 8(1-a)\sqrt{\frac{c_3 \log n}{n}}}}{4(1-a)} & \text{when } 0 \leq a < 1 \\ \sqrt{\frac{c_3 \log n}{n}} & \text{when } a = 1 \end{cases}$$

Conclusions for multi-centralized distribution network is similar to Lemma 3. From Lemma 3, it's not difficult to get the following corollary.

Corollary 1. For centralized network defined by (4), if $l = \sqrt[4]{\frac{c_3 \log n}{n}}$, for any $0 \leq a \leq 1$, there is at least one node in each cell w.h.p.

IV. THE THROUGHPUT CAPACITY OF HETEROGENEOUS WIRELESS NETWORKS WITHOUT OBSTACLES

In this section, we explore a lower bound on the throughput capacity of heterogeneous wireless networks without obstacles by deriving the achievable per-node throughput. We further divide the communication process under helping mode into three phases as that in [16]: First, packets are sent from the source node to the nearest helping node, then forwarded in the helping-network until it reaches the helping node nearest to the destination, and in the final phase packet is transmitted from that helping node to the destination node. We analyze the throughput capacity user mode and the three phases of helping mode. Denote achievable per-node throughput in user mode and helping mode by T_u and T_h , respectively. Thus, the

achievable per-node throughput of the heterogeneous wireless networks, denoted by T , can be calculated as follows:

$$T = \max\{T_u, T_h\} \quad (11)$$

where

$$T_h = \min\{T_{h1}, T_{h2}, T_{h3}\} \quad (12)$$

Here, T_{h1}, T_{h2}, T_{h3} are achievable per-node throughput in the three phases of helping mode. In this section, we assume that all the helping nodes are placed regularly and only investigate the impact of the user nodes' topology. The impact of helping nodes' topology will be studied in the following sections.

A. Uniform Network

As the uniform network we proposed above is a special case of that in [16], which assumed the network placed in a rectangular area, here we only give the derivative results briefly.

Theorem 1. *An achievable throughput in uniform networks, denoted by $T^{uniform}$, is*

$$T^{uniform} = \Omega \left(\max \left\{ \min \left\{ \frac{1}{\sqrt{n \log n}}, n^{d-1} \right\}, \min \left\{ n^{\frac{b}{2}-1}, n^{d-1} \right\} \right\} \right) \quad (13)$$

B. Centralized Network

To facilitate the calculation, here we will only consider the case that centralization coefficient is 0, i.e., the density of nodes in the center goes to the maximum. Moreover, results of such extreme case are also easier for us to compare with that of the uniform network. When $a = 0$, probability density function given in (4) becomes

$$\begin{cases} f(x) = -4x + 2 & (-\frac{1}{2} < x < \frac{1}{2}) \\ f(y) = -4y + 2 & (-\frac{1}{2} < y < \frac{1}{2}) \end{cases} \quad (14)$$

Additionally, due to the non-uniformity, properties of the network are related to cells' locations. As shown in Fig. 4, we number the rows and columns by integrates and let $S(i, j) : i, j = \pm 1, \pm 2, \dots, \pm \frac{1}{2}$ denote the cell located in the i th column and the j th row. If the number of rows or columns is odd, the numbering of cells should start from 0, i.e., $i, j = 0, \pm 1, \pm 2, \dots, \pm \frac{1}{2}$. From the following derivation, we can see that the starting value of the numbering is irrelative to the result.

1) *Achievable Throughput in User Mode:* Let N_x^i and N_y^j denote the number of source nodes located in the same column of $S(i, j)$ and the number of destination nodes located in the same row of $S(i, j)$, respectively. Thus, we have

$$\begin{aligned} E[N_x^i] &= n \cdot \int_{(i-1)l}^{il} f(y) dy \\ &= n \cdot (-2y^2 + 2y) \Big|_{(i-1)l}^{il} \\ &= n \cdot (-4il^2 + 2l^2 + 2l) \end{aligned} \quad (15)$$

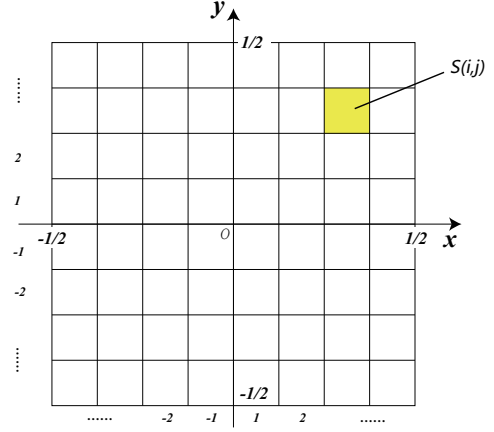


Fig. 4. A numbering of cells

Similarly,

$$\begin{aligned} E[N_y^j] &= n \cdot \int_{(j-1)l}^{jl} f(x) dx \\ &= n \cdot (-4jl^2 + 2l^2 + 2l) \end{aligned} \quad (16)$$

According to Corollary 1, we have

$$\begin{aligned} E[N_x^i] &= 2n(-2i + 1) \left(\frac{c_3 \log n}{n} \right)^{\frac{1}{2}} + 2n \left(\frac{c_3 \log n}{n} \right)^{\frac{1}{4}} \\ E[N_y^j] &= 2n^d(-2j + 1) \left(\frac{c_3 \log n}{n} \right)^{\frac{1}{2}} + 2n^d \left(\frac{c_3 \log n}{n} \right)^{\frac{1}{4}} \end{aligned}$$

Recall the Chernoff bounds, we can obtain the following lemma.

Lemma 4. *For each cell, w.h.p.,*

- 1) *The number of source nodes which are located in cells with the same x-coordinate is at most $4n(-2i + 1) \left(\frac{c_3 \log n}{n} \right)^{\frac{1}{2}} + 4n \left(\frac{c_3 \log n}{n} \right)^{\frac{1}{4}}$.*
- 2) *The number of destination nodes located in cells with the same y-coordinate is at most $4n^d(-2j + 1) \left(\frac{c_3 \log n}{n} \right)^{\frac{1}{2}} + 4n^d \left(\frac{c_3 \log n}{n} \right)^{\frac{1}{4}}$ when $1/4 < d < 1$, and at most c_4 when $0 < d < 1/4$, where c_4 is a constant and $c_4 > \frac{2}{1-4d}$.*

Techniques used to prove Lemma 4 is similar to that of Lemma 3 in [16] and we would not discuss them in details here. Interest reader can refer to [16].

On the other hand, according to [16], for the network with n source nodes and n^d destination nodes, we also have

Lemma 5. *For each destination node, w.h.p., there are at most $2n^{1-d}$ source nodes destined to it [16].*

Let F_k^{ij} denote the number of data flows crossing cell

$S(i, j)$. For each cell, we have

$$F_k^{ij} \leq N_x^i + 2n^{1-d}N_y^j$$

$$\leq \begin{cases} 4n(-2i+1)\left(\frac{c_3 \log n}{n}\right)^{\frac{1}{2}} + 4n\left(\frac{c_3 \log n}{n}\right)^{\frac{1}{4}} \\ + 2n^{1-d}\left[4n^d(-2i+1)\left(\frac{c_3 \log n}{n}\right)^{\frac{1}{2}}\right. \\ \left.+ 4n^d\left(\frac{c_3 \log n}{n}\right)^{\frac{1}{4}}\right] & \text{when } \frac{1}{4} < d < 1 \\ 4n(-2i+1)\left(\frac{c_3 \log n}{n}\right)^{\frac{1}{2}} + 4n\left(\frac{c_3 \log n}{n}\right)^{\frac{1}{4}} \\ + 2c_4n^{1-d} & \text{when } 0 < d < \frac{1}{4} \end{cases} \quad (17)$$

Notice that the left part of (17) are monotonically decreasing functions, we have

$$F_{k,max}^{ij} = F_k^{11}$$

$$= \begin{cases} 9n\left[\left(\frac{c_3 \log n}{n}\right)^{\frac{1}{4}} - \left(\frac{c_3 \log n}{n}\right)^{\frac{1}{2}}\right] \\ \text{when } \frac{1}{4} < d < 1 \\ 4n\left[\left(\frac{c_3 \log n}{n}\right)^{\frac{1}{4}} - \left(\frac{c_3 \log n}{n}\right)^{\frac{1}{2}}\right] + 2c_4n^{1-d} \\ \text{when } 0 < d < \frac{1}{4} \end{cases}$$

i.e.,

$$F_{k,max}^{ij} = O\left(\max\left\{n \cdot \left(\frac{\log n}{n}\right)^{\frac{1}{4}}, n^{1-d}\right\}\right) \quad (18)$$

Form Lemma 1, each cell can achieve a constant transmission rate. Denote the achievable throughput in user mode by $T_n^{central}$, from (18), we can obtain that

$$T_n^{central} = \Omega\left(\min\left\{n^{-1} \cdot \left(\frac{n}{\log n}\right)^{\frac{1}{4}}, n^{d-1}\right\}\right) \quad (19)$$

2) *Achievable Throughput in Helping Mode:* Recall that we have divided communication in helping mode into three phases, in this section we derive the throughput capacity in these three phases, respectively, and synthesize them to get the achievable throughput in helping mode.

Phase I: transmit from sources to the helping nodes

Since helping nodes are regularly placed, we can re-divide the network into m cells of length $l = \sqrt{1/m} = n^{-\frac{b}{2}}$. Then we can obtain the following lemma.

Lemma 6. *There are at most $8n\left(-n^{-b} + n^{-\frac{b}{2}}\right)^2$ user nodes in each cell.*

Proof: Denote the number of user nodes located in cell $S(i, j)$ by random variable N_{ij} . Thus, the expectation of N_{ij} is

$$E[N_{ij}] = n \cdot \int_{(i-1)l'}^{il'} \int_{(j-1)l'}^{jl'} f(x)f(y)dydx$$

$$= n\left[(-4i+2)n^{-b} + 2n^{\frac{b}{2}}\right]$$

$$\cdot \left[(-4j+2)n^{-b} + 2n^{\frac{b}{2}}\right] \quad (20)$$

Notice that the left part of (20) is monotonically decreasing function, thus we have

$$E[N_{ij}]_{max} = E[N_{11}] = 4n\left(-n^{-b} + n^{-\frac{b}{2}}\right)^2 \quad (21)$$

Similar to the proof of Lemma 4, we can easily show that $P(N_{ij} \leq 2E[N_{ij}]_{max} \forall S_{ij}) \rightarrow 1$ as $n \rightarrow \infty$. ■

Let $T_{h1}^{central}$ denote the achievable throughput in phase I. We can obtain that

$$T_{h1}^{central} = \Omega\left(\frac{W_2}{8n\left(-n^{-b} + n^{-\frac{b}{2}}\right)^2}\right) = \Omega(n^{b-1}) \quad (22)$$

Phase II: forwarding in helping network

Notice that packets forwarding in helping network is also ad hoc transmission. Similarly, we can obtain the following lemmas.

Lemma 7. *Each cell in the network can transmit at a transmission rate c_5W_4 , where c_5 is a deterministic positive constant.*

Lemma 8. *For every cell, w.h.p.,*

- 1) *There are at most $4n\left((-2i+1)n^{-b} + n^{-\frac{b}{2}}\right)$ source nodes located in the same column.*
- 2) *The number of destination nodes located in the same row is at most $4n^d \cdot \left((-2j+1)n^{-b} + n^{-\frac{b}{2}}\right)$ when $d > \frac{b}{2}$, and at most c_6 when $d < \frac{b}{2}$, where c_6 is a constant*

Thus, Similar to the derivation of achievable throughput capacity in user mode, we can obtain the throughput capacity in phase II, denoted by $T_{h2}^{central}$, as follows

$$T_{h2}^{central} = \Omega\left(\max\left\{n^{\frac{b}{2}-1}, n^{d-1}\right\}\right) \quad (23)$$

Phase III: transmitting from the helping nodes to destinations

For the re-divided network in helping mode, we have the following lemma.

Lemma 9. *For centralized distribution network, in each cell, w.h.p., there are at most $8n^d\left[-n^{-b} + n^{-\frac{b}{2}}\right]^2$ destination nodes when $0 < b < d < 1$, and at most c_7 when $0 < d < b < 1$, where $c_7 > \frac{b}{b-d}$*

Proof: Denote the number of destination nodes located in cell $S(i, j)$ by random variable D_{ij} . Thus, the expectation of D_{ij} is

$$E[D_{ij}] = n^d \cdot \int_{(i-1)l'}^{il'} \int_{(j-1)l'}^{jl'} f(x)f(y)dydx$$

$$= n^d\left[(-4i+2)n^{-b} + 2n^{\frac{b}{2}}\right]$$

$$\cdot \left[(-4j+2)n^{-b} + 2n^{\frac{b}{2}}\right]$$

$$\leq 4n\left(-n^{-b} + n^{-\frac{b}{2}}\right)^2 \quad (24)$$

Similar to the proof of Lemma 4, we can easily show that $P(D_{ij} \leq 8n^d\left[-n^{-b} + n^{-\frac{b}{2}}\right]^2 \forall S_{ij}) \rightarrow 1$ as $n \rightarrow \infty$. ■

Similar to Lemma 1, we can obtain that transmissions in Phase III can also achieve a constant transmission rate $c_8 W_3$, where $0 < c_8 < \infty$ is a deterministic constant. Furthermore, we find that w.h.p. the number of data flows from each helping node to its destination nodes is at most $8n^d \left[-n^{-b} + n^{-\frac{b}{2}} \right]^2 \times 2n^{1-d} = \Omega(n^{1-b})$ when $0 < b < d < 1$, and at most $c_8 \times 2n^{1-d} = 2c_8 n^{1-d}$ when $0 < d < b < 1$. Then, we can obtain the throughput capacity in phase III, denoted by $T_{h3}^{central}$, as follows

$$T_{h3}^{central} = \begin{cases} \Omega(n^{b-1}) & \text{when } 0 < b < d < 1 \\ \Omega(n^{d-1}) & \text{when } 0 < d < b < 1 \end{cases} \quad (25)$$

Combining (22), (22) and (22), we can get

$$T_h^{central} = \Omega\left(\min\left\{n^{\frac{b}{2}-1}, n^{d-1}\right\}\right) \quad (26)$$

Substituting (19) and (26) into (12), we can get the following theorem

Theorem 2. *An achievable throughput in centralized distribution networks, denoted by $T^{central}$, is*

$$T^{central} = \Omega\left(\max\left\{\min\left\{n^{-1} \cdot \left(\frac{n}{\log n}\right)^{\frac{1}{4}}, n^{d-1}\right\}, \min\left\{n^{\frac{b}{2}-1}, n^{d-1}\right\}\right\}\right) \quad (27)$$

C. Multi-centralized Network

Following similar trace of derivation, we can obtain the following theorem.

Theorem 3. *An achievable throughput in multi-centralized distribution networks, denoted by T^{multi} , is*

$$T^{multi} = \Omega\left(\max\left\{\min\left\{\frac{k}{n} \cdot \left(\frac{\frac{n}{k^2}}{\log \frac{n}{k^2}}\right)^{\frac{1}{4}}, n^{d-1}\right\}, \min\left\{kn^{\frac{b}{2}-1}, n^{d-1}\right\}\right\}\right) \quad (28)$$

Proof: We also need to find an achievable throughput in user mode and helping mode, respectively. An achievable throughput in multi-centralized distribution networks can be obtained by choosing the maximum.

1) user mode

For multi-centralized distribution network, in every sub-network, there are n/k^2 nodes. Thus, in user mode, cells' length is $\frac{1}{k} \cdot \sqrt[4]{\frac{c_3 \frac{n}{k^2}}{\log \frac{n}{k^2}}}$. Similar to that in section IV-B (1), we can obtain that

$$T_n^{multi} = \min\left\{\frac{k}{n} \cdot \left(\frac{\frac{n}{k^2}}{\log \frac{n}{k^2}}\right)^{\frac{1}{4}}, n^{d-1}\right\}$$

2) helping mode

The cell length in helping mode is still $\sqrt{\frac{1}{m}}$, however, the area of each centralized sub-network is only $\frac{1}{k^2}$. Thus,

the relative coverage of each cell is increased. Similar to that in section IV-B (2), we can obtain that

$$\begin{aligned} T_{h1}^{multi} &= \Omega\left(\frac{W_2}{8\frac{n}{k^2} \left(-k^2 \cdot n^{-b} + k \cdot n^{-\frac{b}{2}}\right)^2}\right) \\ &= \Omega(n^{b-1}) \\ T_{h2}^{multi} &= \Omega\left(\min\left\{\frac{c_5 W_4}{9k \cdot \frac{n}{k^2} \left(-n^{-b} + n^{-\frac{b}{2}}\right)}, n^{d-1}\right\}\right) \\ &= \Omega\left(\min\left\{kn^{\frac{b}{2}-1}, n^{d-1}\right\}\right) \\ T_{h3}^{multi} &= \begin{cases} \Omega(n^{b-1}) & \text{when } 0 < b < d < 1 \\ \Omega(n^{d-1}) & \text{when } 0 < d < b < 1 \end{cases} \end{aligned}$$

Thus, an achievable throughput in multi-centralized distribution networks is

$$T^{multi} = \Omega\left(\max\left\{\min\left\{\frac{k}{n} \cdot \left(\frac{\frac{n}{k^2}}{\log \frac{n}{k^2}}\right)^{\frac{1}{4}}, n^{d-1}\right\}, \min\left\{kn^{\frac{b}{2}-1}, n^{d-1}\right\}\right\}\right) \quad \blacksquare$$

V. GENERAL PROPERTIES OF "COMBINED NETWORKS"

From the results in section IV, we can see that compared to sub-networks of the same network scales, the overall networks has a larger achievable throughput. Here, "scale" means the size of network area, number of nodes and size of cells.

To explain this phenomenon, we can divide the impacts of combination into two categories:

- 1) The interference of different sub-networks
- 2) Flows passing across different sub-networks

For impact 1), from the proof of Lemma 1 in [16], we can see that since in one time slot in each cell there is only one node that can transmit packet, the interference between cells is only relative to the size of cells and irrelative to the number of nodes in it. Thus, the interference in combined network is the same to that in sub-networks of the same scales.

Impact 2) is not so easy to explore, however, from the derivation in section IV, we can see that the achievable throughput of the network is bounded by the busiest cells, and the transmission burden of the busiest cells is further determined by the cells with the largest nodes density. Thus, we can obtain the following Theorem.

Theorem 4. *For network composed of some isomorphic sub-networks, the throughput capacity of the overall network, denoted by \tilde{T} , and the throughput capacity of sub-network of same network scales, denoted by $\tilde{\tilde{T}}$, have the following relationship.*

$$\tilde{T} \geq \tilde{\tilde{T}} \quad (29)$$

Proof: Firstly, we give the abstract expression of the derivation of the network throughput capacity.

- 1) In user mode, let $N_{x,max}$ and $N_{y,max}$ denote the maximal number of source nodes located in cells with the same x-coordinate, and the maximal number of destination nodes located in cells with the same y-coordinate in the sub-network, respectively. Thus, the maximal number of flows cross a cell, denoted by $F_{ij,max}$, is

$$F_{ij,max} = N_{x,max} + 2n^{1-d}N_{y,max} \quad (30)$$

Thus, the achievable throughput in user mode is

$$T_u = \frac{c_1 W_1}{F_{ij,max}} \quad (31)$$

- 2) In helping mode, let C_{max} , D_{max} , $N'_{x,max}$ and $N'_{y,max}$ denote the maximal number of source nodes in one cell, the maximal number of destination nodes in one cell, the maximal number of source nodes located in cells with the same x-coordinate, and the maximal number of destination nodes located in cells with the same y-coordinate in the sub-network, respectively. Following the trace of derivation in section IV, we have

- In the first phase

$$T_{h1} = \frac{c_5 W_2}{C_{max}} \quad (32)$$

- In the second phase

$$F'_{ij,max} = N'_{x,max} + 2n^{1-d}N'_{y,max} \quad (33)$$

$$T_{h2} = \frac{W_4}{F'_{ij,max}} \quad (34)$$

- In the third phase

$$T_{h3} = \frac{c_6 W_3}{D_{max}} \quad (35)$$

From above, we can see that the achievable throughput of a network without obstacles is determined by the cells and rows of the largest nodes density, i.e., by variables $N_{x,max}$, $N_{y,max}$, C_{max} , D_{max} , $N'_{x,max}$ and $N'_{y,max}$.

For the combined network consisting of $k \times k$ sub-networks, denote the corresponding variables by $\tilde{(\cdot)}$. We have

$$\begin{cases} \tilde{C}_{max} = C_{max} \\ \tilde{D}_{max} = D_{max} \\ \tilde{N}_{x,max} = kN_{x,max} \\ \tilde{N}_{y,max} = kN_{y,max} \\ \tilde{N}'_{x,max} = kN'_{x,max} \\ \tilde{N}'_{y,max} = kN'_{y,max} \end{cases} \quad (36)$$

Furthermore, we have

$$\begin{aligned} \tilde{F}_{ij,max} &= \tilde{N}_{x,max} + 2n^{1-d}\tilde{N}_{y,max} \\ &= kN_{x,max} + 2n^{1-d}kN_{y,max} \\ &= kF_{ij,max} \end{aligned} \quad (37)$$

Similarly, we have

$$\tilde{F}'_{ij,max} = kF'_{ij,max} \quad (38)$$

Substituting (36), (37) and (38) into (31), (32), (34) and (35), we can obtain that

$$\tilde{T}_n = \frac{\hat{c}_1 W_1}{\tilde{F}_{ij,max}} = \frac{\hat{c}_1 W_1}{kF_{ij,max}}$$

$$\tilde{T}_{h1} = \frac{\hat{c}_5 W_2}{\tilde{C}_{max}} = \frac{\hat{c}_5 W_2}{C_{max}}$$

$$\tilde{T}_{h2} = \frac{W_4}{\tilde{F}'_{ij,max}} = \frac{W_4}{kF'_{ij,max}}$$

$$\tilde{T}_{h3} = \frac{\hat{c}_6 W_3}{\tilde{D}_{max}} = \frac{\hat{c}_6 W_3}{D_{max}}$$

Similarly, for sub-network of the same scales as the overall network, denote the corresponding variables by $\tilde{\tilde{(\cdot)}}$. We have

$$\tilde{\tilde{T}}_n = \frac{\hat{c}_1 W_1}{\tilde{\tilde{F}}_{ij,max}} \leq \frac{\hat{c}_1 W_1}{k^2 \cdot \frac{1}{k} \cdot F_{ij,max}} = \tilde{T}_n$$

$$\tilde{\tilde{T}}_{h1} = \frac{\hat{c}_5 W_2}{\tilde{\tilde{C}}_{max}} \leq \frac{\hat{c}_5 W_2}{k^2 \cdot \frac{1}{k^2} \cdot C_{max}} = \tilde{T}_{h1}$$

$$\tilde{\tilde{T}}_{h2} = \frac{W_4}{\tilde{\tilde{F}}'_{ij,max}} \leq \frac{W_4}{k^2 \cdot \frac{1}{k} \cdot F'_{ij,max}} = \tilde{T}_{h2}$$

$$\tilde{\tilde{T}}_{h3} = \frac{\hat{c}_6 W_3}{\tilde{\tilde{D}}_{max}} \leq \frac{\hat{c}_6 W_3}{k^2 \cdot \frac{1}{k^2} \cdot D_{max}} = \tilde{T}_{h3}$$

Thus, we can obtain that

$$\begin{aligned} \tilde{T} &= \max\{\tilde{T}_n, \tilde{T}_h\} \\ &= \max\{\tilde{T}_n, \min\{\tilde{T}_{h1}, \tilde{T}_{h2}, \tilde{T}_{h3}\}\} \\ &\geq \max\{\tilde{T}_n, \min\{\tilde{\tilde{T}}_{h1}, \tilde{\tilde{T}}_{h2}, \tilde{\tilde{T}}_{h3}\}\} \\ &= \tilde{\tilde{T}} \end{aligned} \quad (39)$$

■

VI. IMPACT OF NETWORK TOPOLOGY ON THROUGHPUT CAPACITY

A. Impact of User Nodes' Topology

Comparing the results in section IV with each other, we can find that achievable throughput of networks with the three different topologies have similar scales (take k as a constant). In general, we have the following theorem.

Theorem 5. *For the topology of user nodes, if the value range of nodes distribution's probability density function (PDF) is limited, the gap in achievable throughput of non-uniform networks and uniformly networks is at most a constant time.*

Proof: Firstly, from the analysis in section V, we can see that if size of cells stay unchanged, the interference between cells can not be changed by network topologies. Secondly, in the proof of Theorem 4, we have conclude that for networks without obstacles the achievable throughput is determined by cells and rows of the largest nodes density, i.e., by variables $N_{x,max}$ $N_{y,max}$, C_{max} , D_{max} , $N'_{x,max}$ and $N'_{y,max}$. Let $\bar{N}_{x,max}$ and $\hat{N}_{x,max}$ denote the the maximal number of source nodes located in cells with the same x-coordinate in uniform and non-uniform, respectively. Since the value range of the nodes distribution's PDF is limited, i.e., $\exists M \in R^+$, for $\forall x, y$ we have $|f(x)| \leq M, |f(y)| \leq M$. Thus, we have

$$E[\bar{N}_x] = n \cdot \frac{l}{1} \quad (40)$$

$$\begin{aligned} E[\hat{N}_x^i] &= n \cdot \int_{(i-1)l}^{il} f(y)dy \\ &\leq n \cdot \int_{(i-1)l}^{il} Mdy \\ &= Mnl \\ &= ME[\bar{N}_x] \end{aligned}$$

Using Chernoff Bounds, we can prove that w.h.p. $\bar{N}_x \leq 2E[\bar{N}_x]$ and $\hat{N}_x^i \leq 2E[\hat{N}_x^i]$. Thus, we have

$$\hat{N}_{x,max} \leq M\bar{N}_{x,max} \quad (41)$$

Similar results can be proved for $N_{y,max}$, C_{max} , D_{max} , $N'_{x,max}$ and $N'_{y,max}$. Denoted the achievable throughput in uniform and non-uniform network by \bar{T} and \hat{T} , respectively. Following the trace of derivation in Theorem 4' proof, we can conclude that

$$\bar{T} \leq C(M)\hat{T} \quad (42)$$

where $C(M)$ is a constant relative to M . ■

B. Impact of Helping Nodes' Topology

In the above analysis, we have only considered regularly distributed helping nodes. In this subsection, we explore the impact of different helping nodes' distributions. In general, we have the following theorem.

Theorem 6. *For networks with uniformly distributed user nodes, regularly distributed helping nodes are optimal to maximize the network throughput capacity.*

Proof: Firstly, we analyze the impact of helping nodes' distribution on interference. In network with non-uniformly distributed helping nodes, cells' sizes in helping mode are also different. Let l' denote length of cells in network with regularly distributed helping nodes and l''_{max} denote the maximal length of cells of network with non-uniformly distributed helping nodes, respectively. We can easily obtain that $l''_{max} \geq l'$. Denote the achievable rate of cells in networks with regularly and non-uniformly distributed network by W' and W'' , respectively. From the derivation of Lemma 1, we can obtain that $W' \geq W''$.

Secondly, according to the proof of Theorem 4, achievable throughput in helping mode is determined by variables C_{max} , D_{max} , $N'_{x,max}$ and $N'_{y,max}$. Let C'_{max} , D'_{max} , $N'_{x,max}$ and $N'_{y,max}$ denote the corresponding variables in network with regularly distributed helping nodes, and C''_{max} , D''_{max} , $N''_{x,max}$ and $N''_{y,max}$ denote corresponding variables in network with non-uniformly distributed helping nodes, respectively. In networks with non-uniformly distributed helping nodes, there must be some cells of length larger than the average value. Thus, we can obtain that

$$\begin{cases} C''_{max} > C'_{max} \\ D''_{max} > D'_{max} \\ N''_{x,max} > N'_{x,max} \\ N''_{y,max} > N'_{y,max} \end{cases} \quad (43)$$

Denote the achievable throughput capacity of network with regularly and non-uniformly distributed network by T' and T'' , respectively. Since that throughput capacity in helping mode is inversely proportional to these variables, from above we can obtain the following conclusion

$$T' \geq T'' \quad (44)$$

■

Theorem 7. *For networks with non-uniformly distributed user nodes, though regularly distributed helping nodes topology is no longer optimal, any improvement on the helping nodes' topology cannot change the network throughput capacity on scale.*

Proof: Firstly, considering interference, according to the proof of Theorem 6, non-uniformly distributed helping nodes can only increase the interference between cells and decrease the achievable rate of cells. Thus, it has a negative impact on the network throughput capacity.

Secondly, if we do not consider change of interference, network throughput capacity in helping mode is determined by variables C_{max} , D_{max} , $N'_{x,max}$ and $N'_{y,max}$. Thus, if we change the helping nodes' topology, network throughput capacity will achieve the maximal value when each cell has the same number of nodes. However, similar to that in the proof of Theorem 5, we can easily show that this improvement is not larger than a constant time.

Combining conclusions above together, we can conclude that improvement on the helping node topology cannot change the network throughput capacity on scale. ■

VII. OPTIMAL TOPOLOGY FOR NETWORKS WITH OBSTACLES

In this section, we introduce a novel algorithm to generate the optimal network topology for any given obstacles distributions and analyze its performance.

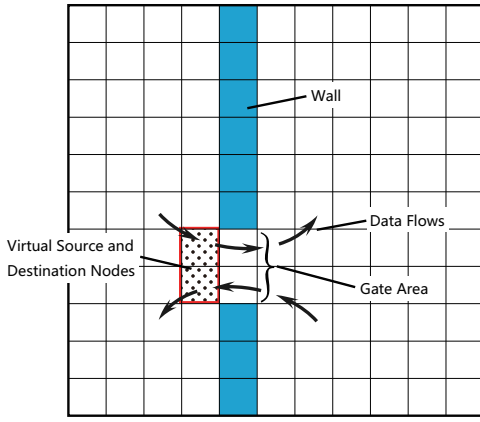


Fig. 5. A wall with gate in the network area

A. Algorithm to Obtain the Optimal Network Topology

To design the optimization algorithm, we first consider a simple scenario. As shown in Fig. 5, assume that there is a “wall” with a “gate” in the network, which divides the network area into two parts. In this case, the area around the gate is the “hot spot” and the bottleneck of the network achievable throughput since any data flow passing from one side of the wall to another side must pass through the gate. To maximize the network throughput capacity, we can minimize the transmission burden of this area by the following algorithm.

Algorithm - “Wall with Gate”:

- 1) Assume that there are \hat{n} , n_1 and n_2 number of nodes in the gate area, the left and the right part of the network, respectively, where $\hat{n} + n_1 + n_2 = n$. The expect number of data flows passing through the gate is $u = f(\hat{n}, n_1, n_2)$, where function $f(\cdot)$ can be decided using the methods given in Section IV. Thus the transmission burden of the gate area is $B_0 = u/k_0$, where k_0 is the number of cells in the gate area (the nodes’ distribution in the gate area is assumed to be uniform since this area is relatively small).
- 2) Ignore the wall and the right part of the network. Put $\varphi_1 = g_s(\hat{n}, n_1, n_2)$ number of virtual source nodes and $\phi_1 = g_d(\hat{n}, n_1, n_2)$ number of virtual destination nodes uniformly in front of the gate (i.e., the area illustrated in Fig. 5) to replace the ignored user nodes. Virtual source and destination nodes work as sources and destinations, respectively, generating virtual data flows. Functions $g_s(\cdot)$ and $g_d(\cdot)$ are determined by the routing strategy so that this number of virtual nodes have the same influence on the left part of the network as the ignored parts.¹ Then we obtain a degraded sub-network without any obstacles.
- 3) For the degraded sub-network, use methods and conclusions given in Sections IV - VI to generate an optimal topology $\mathfrak{T}_1 = T_1(\hat{n}, n_1, n_2)$ and calculate the corresponding transmission burden $B_1 = h_1(\hat{n}, n_1, n_2)$.
- 4) Repeat Step 2 and 3 to the right part of the network, respectively. Generate the optimal topology $\mathfrak{T}_2 =$

¹For the routing strategy given in Section II, we can let $g_s(\hat{n}, n_1, n_2) = n_1 n_2 / n$ and $g_d(\hat{n}, n_1, n_2) = n_1 n_2 / n$, respectively.

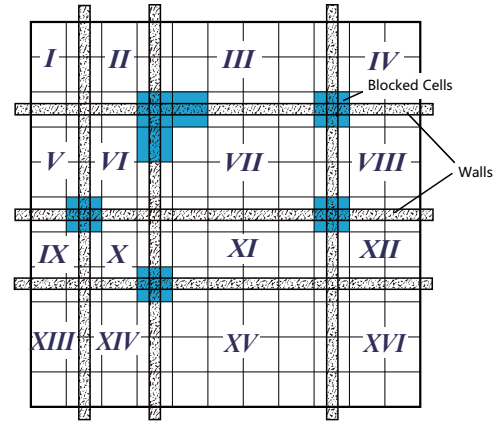


Fig. 6. Divide the network by walls - Method I

$T_2(\hat{n}, n_1, n_2)$ and calculate the transmission burden $B_2 = h_2(\hat{n}, n_1, n_2)$.

- 5) Use appropriate optimization methods to minimize the cost function $B = \max(B_0, B_1, B_2)$. Calculate corresponding \hat{n} , n_1 and n_2 . Since the topologies obtained in Step 3 and 4 are functions of \hat{n} , n_1 and n_2 , the optimal topology for the whole network can thus be determined by combining \mathfrak{T}_1 and \mathfrak{T}_2 .

This “Wall with Gate” algorithm can be generalized to obtain optimal topology for any given networks with obstacles. Firstly, divide the network area into pieces by the following method.

Divide the network by walls - Method I: As shown in Fig. 6, take blocked cells in a row (vertical or horizontal) as a wall and cells without obstacles in this row as gates. Then the network is divided into some sub-networks by these walls.

The optimal topology for this network area can be obtained by applying “Wall with Gate” algorithm to each of these sub-networks and gate areas. Note that since the gate areas here might be relatively large, nodes distribution in these areas can no longer be assumed to be uniform and Step 2 - 3 must be applied to these gate areas. Denote the optimal network topology by \mathfrak{T} , assume that there are S sub-networks and R gates, we have

$$\begin{aligned} \mathfrak{T} &= T(n_1, n_2, \dots, n_S, \hat{n}_1, \hat{n}_2, \dots, \hat{n}_R,) \\ &= \{\mathfrak{T}_1, \dots, \mathfrak{T}_S, \hat{\mathfrak{T}}_1, \dots, \hat{\mathfrak{T}}_2, \hat{\mathfrak{T}}_R\} \end{aligned}$$

where $n_i, i = 1, 2, \dots, S$ is the number of user nodes in the i th sub-network, $\hat{n}_j, j = 1, 2, \dots, R$ is the number of user nodes in the j th gate area, $\mathfrak{T}_k, k = 1, 2, \dots, S$ is the optimal topology of the k th sub-network and $\hat{\mathfrak{T}}_l, l = 1, 2, \dots, R$ is the optimal topology of the l th gate area.

B. Complexity of the Algorithm

Noticing that the algorithm complexity is proportional to the number of sub-networks S and number of gates R , to decrease the algorithm complexity, we can simplify the division of the network area. Note that in Fig. 6, sub-networks I - IV can be combined into a larger sub-network, so do sub-networks V

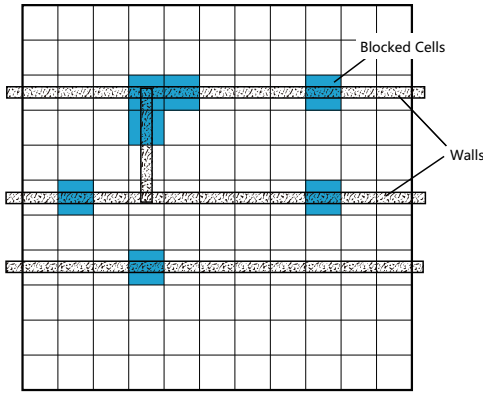


Fig. 7. Divide the network by walls - Method II

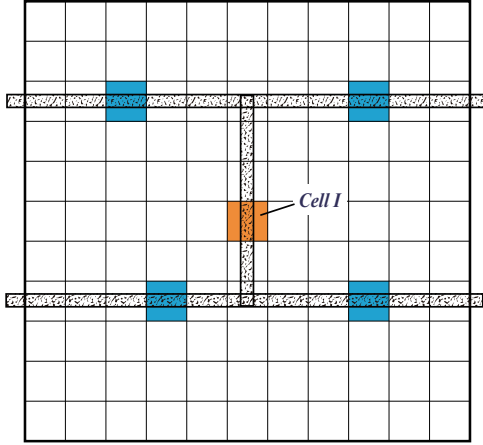


Fig. 8. Cell I add four gates to the divided network

- VI, VII - VIII, IX - XII and XIII - XVI. We can modify the network dividing method as follows.

Divide the network by walls - Method II: As shown in Fig. 7, first construct a wall in the row (either vertical or horizontal) with the most number of blocked cells, dividing the network area into two parts. For each part, repeat this step iteratively until all the blocked cells are crossed by at least one wall.

The complexity of the algorithm is given by the following lemma.

Lemma 10. *The algorithm complexity is $O(M^2)$ when using network dividing method I and is $O(M)$ when using method II.*

Proof: Here we consider the worst cases, i.e., the cases that all cells with obstacles are not collinear. When using method I to divide the network area, there are $2M$ walls. Denote the number of sub-networks and number of gates by S_I and R_I , respectively. In the worst case, these $2M$ walls divide the network into M^2 areas, i.e., $S_I = M^2$. Since there is a gate between any neighboring sub-network areas, there are at least $4 * M^2 / 2$ gates, i.e., $R_I = 2M^2$. So the algorithm complexity is

$$\eta_I = O(S_I + R_I) = O(M^2) \quad (45)$$

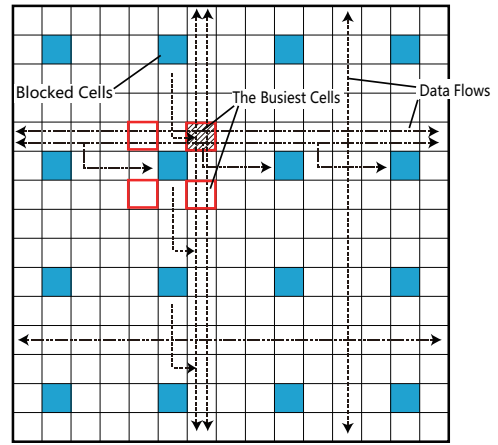


Fig. 9. Data flows in network with regularly distributed obstacles

When using method II, each cell with obstacles generate at most one wall, so there are at most M walls. Denote the number of sub-networks and number of gates by S_{II} and R_{II} , respectively. These walls divide the network into $M + 1$ areas, thus, $S_{II} = M + 1$. Furthermore, as shown in Fig. 8, each cell with obstacles can generate at most four additional gates to the network, i.e., $R_{II} \leq 4M$. So the algorithm complexity is

$$\eta_{II} = O(S_{II} + R_{II}) = O(M) \quad (46)$$

■

Although method I generate more sub-networks, each sub-network is relatively simple and easy to analyze. In some particular cases, for example, in the case that the obstacles' distribution has some symmetry properties, using method I might result in a smaller algorithm complexity.

C. Performance of the Algorithm for Network with Regularly Distributed Obstacles

Generally, the more uniform the obstacles' distribution is, the larger the achievable throughput is, and thus the less room for improvement there is. Therefore we take regular obstacles distribution, one of the most uniform distributions, as an example to estimate the lower bound of the algorithm performance.

1) *Achievable throughput of network with uniformly distributed user nodes:* Since obstacle nodes are regularly distributed, cells blocked by obstacles are also regularly distributed. Denote the number of blocked cells by \mathfrak{M} . To keep the network connectivity, we have $M \leq \frac{\mathcal{K}}{4}$ and $\mathfrak{M} = M$. As shown in Fig. 9, the marked cells are the busiest ones in the network since they have to relay extra data flows from or to the nodes located in the adjacent columns or rows, respectively. Denote the number of data flows that cross the i th ($1 \leq i \leq \mathcal{K}$) cell by F_i . Following similar trace of derivation in Section IV,

we can obtain that for all i

$$F_i \leq \begin{cases} 6 \left(1 + C \frac{\sqrt{\mathcal{K}} - \sqrt{M}}{\sqrt{\mathcal{K}}}\right) \cdot \frac{n}{\mathcal{K} - M} \cdot \sqrt{\mathcal{K}} & \text{when } \frac{w}{2} < d < 1 \\ \left(1 + C \frac{\sqrt{\mathcal{K}} - \sqrt{M}}{\sqrt{\mathcal{K}}}\right) \left(\frac{2n}{\mathcal{K} - M} \cdot \sqrt{\mathcal{K}} + 4c_9 n^{1-d}\right) & \text{when } 0 < d < \frac{w}{2} \end{cases} \quad (47)$$

where c_9 is a deterministic constant, $C = 1/2$ when $M < \mathcal{K}/4$ and $C = 1$ when $M = \mathcal{K}/4$. The achievable throughput is $T_r = c_1 W_1 / F_i$. For comparison, in network without obstacles, the number of data flows that cross the i th cell, denoted by \hat{F}_i , is

$$\hat{F}_i \leq \begin{cases} 6 \frac{n}{\sqrt{\mathcal{K}}} & \text{when } \frac{w}{2} < d < 1 \\ 2 \frac{n}{\sqrt{\mathcal{K}}} + 4c_9 n^{1-d} & \text{when } 0 < d < \frac{w}{2} \end{cases} \quad (48)$$

The corresponding achievable throughput is $\hat{T} = c_1 W_1 / \hat{F}_i$. Considering that $\frac{n}{\mathcal{K} - M} \cdot \sqrt{\mathcal{K}} = o(n^{1-d})$ and $\frac{n}{\sqrt{\mathcal{K}}} = o(n^{1-d})$ when $0 < d < w/2$, as n goes to infinity, the gap in achievable throughput between these two networks is

$$\frac{\hat{T}}{T_r} = \begin{cases} \frac{\mathcal{K}}{\mathcal{K} - M} \cdot \left(1 + C \frac{\sqrt{\mathcal{K}} - \sqrt{M}}{\sqrt{\mathcal{K}}}\right) & \text{when } \frac{w}{2} < d < 1 \\ \left(1 + C \frac{\sqrt{\mathcal{K}} - \sqrt{M}}{\sqrt{\mathcal{K}}}\right) & \text{when } 0 < d < \frac{w}{2} \end{cases} \quad (49)$$

From equation (49), we can obtain the following theorem.

Theorem 8. *Uniform user nodes' distribution is order optimal for throughput capacity of networks with regularly distributed obstacles.*

Proof: Since \hat{T} is the upper bound of achievable throughput for networks with obstacles, no matter how user nodes and obstacles are distributed, we only need to prove that $T_r = \Theta(\hat{T})$. According to equation (49), when $M = o(\mathcal{K})$, we have $\hat{T}/T_r = 1 + C$. When $M = A \cdot \mathcal{K}$ ($0 < A < 1/4$), we have

$$\frac{\hat{T}}{T_r} = \begin{cases} \frac{1 + C - C\sqrt{A}}{1 - A} & \text{when } \frac{w}{2} < d < 1 \\ 1 + C - C\sqrt{A} & \text{when } 0 < d < \frac{w}{2} \end{cases} \quad (50)$$

Since $1/2 \leq C \leq 1$ and $0 < A < 1/4$, for any $0 < M \leq \mathcal{K}/4$, we can obtain that

$$1 \leq \frac{\hat{T}}{T_r} \leq 2$$

i.e., $T_r = \Theta(\hat{T})$. ■

2) *Performance of the topology optimization algorithm:* A better network topology is shown in Fig. 10, whose achievable throughput can be taken as the lower bound of the optimal network topology. User nodes are distributed uniformly in the marked areas. Denote the number of data flows crossing the i th cell by F'_i , we have that for all i

$$F'_i \leq \begin{cases} 6 \frac{n}{\sqrt{\mathcal{K}} - \sqrt{M}} & \text{when } \frac{w}{2} < d < 1 \\ 2 \frac{n}{\sqrt{\mathcal{K}} - \sqrt{M}} + 4c_9 n^{1-d} & \text{when } 0 < d < \frac{w}{2} \end{cases}$$

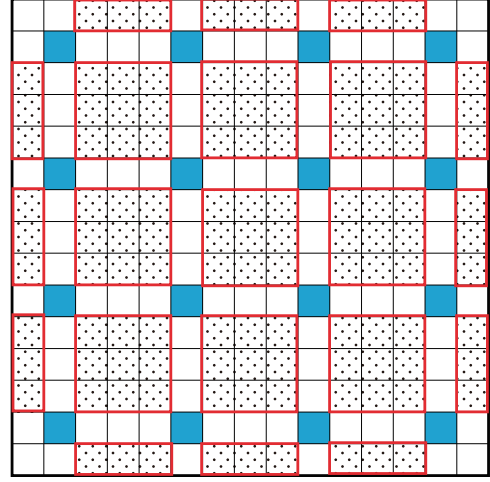


Fig. 10. A better network topology for regularly distributed obstacles

The corresponding achievable throughput is $T'_r = c_1 W_1 / F'_i$. When n goes to infinity, the improvement in achievable throughput is

$$\frac{T'_r}{T_r} = \begin{cases} \frac{(1+C)\sqrt{\mathcal{K}} - C\sqrt{M}}{\sqrt{\mathcal{K}} + \sqrt{M}} & \text{when } \frac{w}{2} < d < 1 \\ \frac{(1+C)\sqrt{\mathcal{K}} - C\sqrt{M}}{\sqrt{\mathcal{K}}} & \text{when } 0 < d < \frac{w}{2} \end{cases}$$

When $M = o(\mathcal{K})$, we have

$$\frac{T'_r}{T_r} = \frac{\hat{T}}{T_r} = 1 + C \quad (51)$$

When $M = A \cdot \mathcal{K}$ ($0 < A < 1/4$), we have

$$\begin{aligned} \frac{T'_r}{T_r} &= \begin{cases} \frac{1 + C - C\sqrt{A}}{1 + \sqrt{A}} & \text{when } \frac{w}{2} < d < 1 \\ 1 + C - C\sqrt{A} & \text{when } 0 < d < \frac{w}{2} \end{cases} \\ &= \begin{cases} \frac{\hat{T}/T_r}{1 - \sqrt{A}} & \text{when } \frac{w}{2} < d < 1 \\ \hat{T}/T_r & \text{when } 0 < d < \frac{w}{2} \end{cases} \end{aligned} \quad (52)$$

From equations (51) and (52) we can find that when the number of obstacles is small (i.e., $M = o(\mathcal{K})$), or when the number of destinations is small (i.e., $0 < d < w/2$), the asymptotic throughput capacity of networks with regularly distributed obstacles and optimal network topology can reach the that of networks without obstacles. When the number of obstacles and number of destinations are both large (i.e., $M = \Theta(\mathcal{K})$ and $w/2 < d < 1$), the asymptotic throughput capacity of networks with regularly distributed obstacles and optimal network topology may smaller than that of networks without obstacles by at most a constant time. The worst case happens when $M = \mathcal{K}/4$, in which $T'_r = \hat{T}/2$.

VIII. CONCLUSION

In this paper, we investigate the throughput capacity of heterogeneous wireless network with different network topologies and analyze the impact of topologies on the network

properties. We find that compared to the sub-networks with the same network scales, combined networks have a larger overall network achievable throughput. We also find that uniformly distributed user nodes and regularly distributed helping nodes have some advantages in improving the network capacity. Compared to regularly distributed helping nodes, any change of helping nodes' topology cannot improve the network achievable throughput on scale. We further investigate the impact of obstacles and introduce an algorithm to generate the optimal user nodes' distribution for any given network areas with obstacles.

REFERENCES

- [1] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388-404, Mar. 2000.
- [2] M. Grossglauser and M.D.Tse, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol.46,no.2, pp.388-404, March 2000.
- [3] A. Keshavarz-Haddad, V. Ribeiro, and R. Riedi, "Broadcast capacity in multihop wireless networks," in *Proc. ACM MobiCom*, Sept. 2006.
- [4] C. Hu, X. Wang and F. Wu, "MotionCast: on the capacity and delay tradeoffs," in *Proc. ACM MobiHoc*, May 2009.
- [5] A. Avudainayagam, Y. Fang, and W. Lou, "Dear: a device and energy aware routing protocol for heterogeneous ad hoc networks," *Journal of Parallel and Distributed Computing (Special Issue on Mobile Ad-hoc Networking and Computing)*, 63(2):228-236, February 2003.
- [6] W. Liu, Y. Zhang, W. Lou, and Y. Fang, "Delar: device/energy/load aware relaying in heterogeneous wireless ad hoc networks," In *Proceedings of IEEE MILCOM*, Monterey, CA, USA, November 2004.
- [7] W. Liu, Y. Zhang, and Y. Fang, "Conserving energy in heterogeneous mobile ad hoc networks," In *Proceedings of IEEE MILCOM*, Atlantic City, New Jersey, USA, October 2005.
- [8] W. Liu, Y. Zhang, K. Lu, and Y. Fang, "Energy conservation through resource-aware movement in heterogeneous mobile ad hoc networks," *Journal of Combinatorial Optimization (Special Issue on Wireless Network Applications)*, 24(1):7-20, February 2006.
- [9] E. Wu and Y. Huang. "Dynamic adaptive routing for a heterogeneous wireless network," *Mobile Networks and Applications*, 9(3):219-233, June 2004.
- [10] K. Yang, Y. Wu, and H.-H. Chen, "Qos-aware routing in emerging heterogeneous wireless networks," *IEEE Communications Magazine*, 45(2):74-80, 2007.
- [11] P. Li, X. Geng, and Y. Fang, "An adaptive power controlled mac protocol for wireless ad hoc networks," *IEEE Transactions on Wireless Communications*, 8(1):226-233, January 2009.
- [12] P. Li, Q. Shen, and Y. Fang, "Power controlled network protocols for multi-rate ad hoc networks," *IEEE Transactions on Wireless Communications*, 8(4):2142-2149, April 2009.
- [13] M. McGarry, M. Reisslein, and V. Syrotiuk, "Access control in heterogeneous multichannel wireless networks," In *Proceedings of the First International Conference on Integrated Internet Ad hoc and Sensor Networks (InterSense'06)*, Nice, France, May 2006.
- [14] H. Zhou, C. Yeh, and H. Moustah, "Access control in heterogeneous multichannel wireless networks," In *Proceedings of IEEE 60th Vehicular Technology Conference*, Los Angeles, CA, USA, September 2004.
- [15] N. Li and J. Hou, "Topology control in heterogeneous wireless networks: Problems and solutions," In *Proceeding of IEEE INFOCOM*, Hong Kong, China, March 2004.
- [16] P. Li and Y. Fang, "The Capacity of heterogeneous wireless networks," In *Proceeding of IEEE INFOCOM*, 2010.
- [17] S. R. Kulkarni, P.Viswanath, "Deterministic Approach to Throughput Scaling in Wireless Networks", *IEEE Trans. Inform. Theory*, vol. 50(6), pp. 10411049, June 2004
- [18] E. Perevalov, R. S. Blum, D. Safi, "Capacity of Clustered Ad Hoc Networks: How Large Is Large?," *IEEE Trans. Commun.*, Vol. 54, No. 9, pp. 1672C1681, Sept. 2006.
- [19] R. K. Ganti and M. Haenggi, "Interference and Outage in Clustered Wireless Ad Hoc Networks," *IEEE Trans. Inform. Theory*, to appear. Available at <http://arxiv.org/abs/0706.2434v1>
- [20] A.Keshavarz-Haddad and R.H. Riedi, "Bounds for the capacity of wireless multihop networks imposed by topology and demand," in *Proc. ACM MobiHoc '07*, pp. 256-265, Montreal, Canada, Sept. 2007.
- [21] G.Alfano, M.Garetto, E. Leonardi, "Capacity Scaling of Wireless Networks with Inhomogeneous Node Density: Upper Bounds", *IEEE Journal ON Selected Areas in Communications*, vol. 27, no. 7, Sep. 2009
- [22] G.Alfano, M.Garetto, E. Leonardi, "Capacity Scaling of Wireless Networks with Inhomogeneous Node Density: Lower Bounds", in *Proc. INFOCOM 09*, Rio de Janeiro, Brazil, April 2009.
- [23] T. Rappaport, "Wireless Communications: Principles and Practice (Second Edition)," *Prentice-Hall PTR*, 2002.
- [24] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Optimal Throughput-Delay Scaling in Wireless Networks-Part I The Fluid Model," Dept. Elect. Eng., Stanford Univ., Stanford, CA, 2005 [Online]. Available: <http://www.stanford.edu/~jmammen/>