Optimal Capacity-Delay Tradeoff in MANETs with Correlation of Node Mobility

Riheng Jia, Feng Yang, Shuochao Yao, Xiaohua Tian, Xinbing Wang, Wenjun Zhang, and Jun Xu

Abstract—In this paper, we analyze the capacity and delay in mobile ad hoc networks (MANETs) considering the correlation of node mobility (correlated mobility). Previous works on correlated mobility investigated the maximum capacity with the corresponding delay in several sub-case, the problem of optimal capacity under various delay constraints (the optimal capacity-delay tradeoff) still remains open. To this end, we deeply explore the characteristics of correlated mobility, and figure out the fundamental relationships between the network performance and scheduling parameters. Based on that we establish the overall upper bound of capacity-delay tradeoff in all the sub-case of correlated mobility. Then we try to obtain the achievable lower bound by identifying the optimal scheduling parameters on certain constrains. Results demonstrates the whole picture of how the correlation of node mobility impacts the capacity, delay and the corresponding tradeoff between them.

Index Terms—Mobile Ad Hoc Networks (MANETs), capacity and delay tradeoff, correlated mobility

I. INTRODUCTION

Since the ground-breaking work by Gupta and Kumar [1], the study of capacity in large scale wireless networks has gained great popularity within the research community. Gupta and Kumar showed that the per-node capacity is bounded by \( O\left(1/\sqrt{n \log n}\right) \) as the number of nodes \( n \) increases in a static network. Then Franceschetti and Dousse [2] utilized the percolation theory to improve the per-node capacity up to \( O\left(1/\sqrt{n}\right) \) and the capacity still decays rapidly as \( n \) increases.

In the seminal work [3], Grossglauser and Tse verified that the network can achieve a constant per-node capacity of \( O\left(1\right) \) when taking mobility into account, at the expense of unbounded delay. Therefore, researchers began to explore the relationship between the per-node capacity and the packet delay (capacity-delay tradeoff), for balancing the respective capacity and delay performance. It is well known that the i.i.d. mobility model plays an important role in the early investigations of how mobility influences the network performance for its mathematical tractability. Related works are [4], [5], and [6], which provide some insights of the relationship between node mobility and the capacity-delay tradeoff.

For better and comprehensively understanding the impact of mobility on the capacity-delay tradeoff in wireless networks, literatures of the scaling performance under various mobility patterns were consecutively created by researchers. The related works are Brownian motion mobility [7], random waypoint mobility [8], linear mobility [9], restricted mobility [10], [11], [12] and so on.

The aforementioned mobility models are either uniform or non-uniform over the network, basically covering the majority of the existing characteristics of node mobility except the correlation of node mobility (correlated mobility). The mobility in real world exhibits certain degree of correlation [14-17], which stimulates us to investigate the impact of correlated mobility on the capacity-delay tradeoff in wireless networks.

Correlated mobility can be divided into three sub-case based on different degrees of correlation of node mobility: 1) the cluster sparse regime (node mobility show strong correlation); 2) the cluster dense regime (node mobility show weak correlation); and 3) the cluster critical regime (node mobility show medium correlation). Ciullo et al. [13] first introduced correlated mobility into the scaling analysis of wireless networks. They obtained the maximum capacity with the corresponding packet delay in cluster sparse regime and the lower bound of capacity with the corresponding packet delay in cluster dense regime, respectively. Yet the problem of optimal capacity performance under various delay constraints remains to be solved, which can provide significant insights for the better design of wireless networks requiring operating in various delay conditions. In this paper, We study the following open question:

- What is the optimal capacity-delay tradeoff with correlated mobility (in all sub-cases) in mobile ad hoc networks (MANETs)?

We first study the correlated mobility model and establish an upper bound on the optimal capacity-delay tradeoff in MANETs. Further, we develop a scheduling policy to achieve the upper bound up to a logarithmic factor.

We summarize the main observations as follows: 1) The network generally performs worse in cluster sparse regime (strong correlation of node mobility) than that of the i.i.d. mobility model. Because the strong correlation of node mobility has destroyed the network connectivity. 2) The optimal capacity-delay tradeoff in cluster dense regime and in cluster critical regime both perform better than that in i.i.d. mobility model, which is also better than the results in previous works on correlated mobility. 3) The correlated mobility in cluster critical regime can achieve the best performance of the optimal capacity-delay tradeoff among the three sub-case. It indicates that the medium correlation of node mobility can greatly benefit the network performance. The main contribution of this paper is that we are the first to demonstrate a whole picture of how correlated mobility impact the capacity-delay tradeoff in MANETs.
The rest of this paper is organized as follows. In section II, we introduce our system model. In section III, we briefly illustrate the optimal capacity-delay tradeoff in cluster sparse regime. In section IV, we establish the upper bound of the optimal capacity-delay tradeoff in cluster dense regime, accordingly the lower bound is derived in Section V. Section VI discusses the results and Section VII concludes the paper.

II. SYSTEM MODEL

In the subsequent analysis throughout this paper, we apply the correlated mobility model to depict the motion of nodes. Specifically, we consider \( n \) nodes moving over an extended square of area \( n \). All nodes are divided into \( m = \Theta(n^v) \) groups, where \( 0 \leq v < 1 \). Each group covers a circular area of radius \( R = \Theta(n^\beta) \), where \( 0 \leq \beta \leq 1/2 \). In particular, we call each group as a cluster. Note that each cluster on average contains \( q = n/m \) nodes and the result won’t change even if the value of \( q \) differs in clusters but remains \( \Theta(n/m) \) unchanged. The related notations are shown in Table I.

<table>
<thead>
<tr>
<th>TABLE I: Notations</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
</tr>
<tr>
<td>( m )</td>
</tr>
<tr>
<td>( v )</td>
</tr>
<tr>
<td>( q )</td>
</tr>
<tr>
<td>( R )</td>
</tr>
<tr>
<td>( \beta )</td>
</tr>
</tbody>
</table>

The Motion of Cluster Center: At the end of each slot, the network scheduler decides the position of each cluster center \( j \) in the next slot. In each slot, the position of cluster center \( j \) is randomly and uniformly chosen within the entire network area, independently from other cluster centers. After receiving the decision, all the cluster centers move to the scheduled positions in the next slot.

The Motion of Cluster Member: Once the new position of cluster center \( j \) is selected, all nodes in this cluster move to the new region close to \( j \), i.e., a circular area of \( \Theta(R^2) \) that the cluster of \( j \) covers. Then the position of cluster member \( i \) is randomly and uniformly chosen within the new region, independently from other nodes in this region.

We observe that either reducing the number of clusters or the area each cluster covers will achieve strong correlation of node mobility. Depending on the values of \( \beta \) and \( v \), we can divide our analysis into three different regimes: 1) cluster sparse regime \((v + 2\beta < 1)\): the total area \( mR^2 \) that all clusters cover is \( o(n) \), which indicates strong correlation of node mobility; 2) cluster dense regime \((v + 2\beta > 1)\): the total area \( mR^2 \) that all clusters cover is \( \omega(n) \), which indicates weak correlation of node mobility; 3) cluster critical regime \((v + 2\beta = 1)\): the total area \( mR^2 \) that all clusters cover is \( \Theta(n) \), which indicates medium correlation of node mobility.\(^3\)

\(^3\)The degree of correlation of node mobility can be adjusted by changing the values of \( \beta \) and \( v \).

Traffic Pattern: We assume that each node is a source node associated with one destination, which is randomly and independently chosen among all the other nodes in the network. We also assume that the destination is uniformly chosen among all the clusters excluding the cluster of the source. Then the source send packets to the corresponding destination via a common wireless channel and we utilize the protocol model [1] to reduce the interference.

Definitions of Asymptotic Throughput and Packet Delay:

Let \( \lambda_i (i = 1, \ldots, n) \) denotes the sustainable rate of data flow for node \( i \) and \( D_b (b = 1, \ldots, \lambda nT) \) represents the delay for packet \( b \). Assume that \( \lambda = \min\{\lambda_1, \lambda_2, \ldots, \lambda_{n-1}, \lambda_n\} \) and \( \bar{D} = \sum_{b=1}^{\lambda nT} D_b/\lambda nT \). Then \( \lambda = \Theta(f(n)) \) is defined as the asymptotic capacity if there exist two constants \( c \) and \( c' \), where \( c > c' > 0 \) that

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

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

Similarly, \( \bar{D} = \Theta(g(n)) \) is defined as the asymptotic delay.

III. CLUSTER SPARSE REGIME

The analysis of capacity-delay tradeoff in cluster sparse regime \((v + 2\beta < 1)\) was addressed in our previous work [19]. Here we show the main results for completeness.

A. Scheduling Policy

For a traffic stream \( s \to d \), we denote \( s \) and \( d \) as the source and its destination respectively. In addition, We denote \( C_s \)
and $C_d$ as two clusters who containing $s$ and $d$ respectively, where $C_s \neq C_d$. Opportunistic broadcasting scheme (nodes only broadcast messages when there exist a large number of nodes around) is applied here to fully utilize the correlation of node mobility. We show the scheduling policy as follow:

1) When $s$ meets a cluster $C_k$ ($k=1,\ldots,R^*_s$, where $R^*_s$ is the maximum number of clusters who contain messages of $s$) who contains no messages of $s$, a relay will be created in $C_k$ via one-hop unicast. We call this process as \textit{inter-cluster duplication} and this process will not end until one of the relays meets $C_d$.

2) If one of the relays meets $C_d$, a new relay will be created in $C_d$ via one-hop unicast. If not, back to step 1).

3) The newly created relay in $C_d$ will create relays within $C_d$ via broadcast ($R^*_d$ denotes the total number of relays created in $C_d$). We call this process as \textit{intra-cluster duplication}.

4) If one of the relays in $C_d$ is captured\(^2\) by the destination within range $l^*$, the message of $s$ will be transmitted to the destination via $h^*$-hop unicast transmission. If not, back to step 3).

\textbf{B. Upper Bound of Capacity-Delay Tradeoff}

Given the brief introduction of the scheduling scheme, we directly show the upper bound of capacity-delay tradeoff in cluster sparse regime. Please refer to [19] for detailed proofs.

\textbf{Theorem 1.} In cluster sparse regime, let $\bar{D}^s = \Theta(n^\beta)$ denote the mean delay averaged over all bits and let $\lambda^s$ be the capacity of each source-destination pair. The following upper bound holds,

\[
\begin{align*}
(\lambda^s)^3 \leq O\left(\frac{mR^2}{n}\log^3 n\right) \quad &\bar{d} \geq \frac{5}{2} - v - 6\beta \\
\lambda^s \leq O\left(\frac{mR^2}{n}\log^3 n\right) \quad &\bar{d} < \frac{5}{2} - v - 6\beta,
\end{align*}
\]

where $\lambda^s \leq mR^2/n$ and $\bar{D}^s \geq n/(mR^2)$.

\textbf{C. Achievable Lower Bound}

We divide unit slot into three sub-slot. The operation of each sub-slot is shown below:

1) Nodes (source and relays) create inter-cluster duplications and $C_d$ receives messages from one of the inter-cluster duplications via one hop unicast transmission. Each hop exploits the transmission range of $r^*$.

2) $R^*_d$ intra-cluster duplications are created via broadcasting within $C_d$.

3) If one of the intra-cluster duplications is captured by the destination within range $l^*$, then the message of $s$ will be delivered to the destination via $h^*$-hop unicast transmission. Each hop exploits the transmission range of $r^*$.

The detailed analysis of the capacity-delay tradeoff achieving scheme as well as the corresponding results were shown in our work [19], which are omitted here for simplicity.

\(^2\)We define a event that node A falls into a certain area that node B covers and the message of node A can be delivered to node B via multihop transmission in one slot. If the event happens, we say node A is captured by node B.

\textbf{IV. Cluster Dense Regime}

In cluster dense regime ($v + 2\beta > 1$) where the node mobility show weak correlation, we can observe that either the area each cluster covers or the number of clusters becomes larger compared with the situation in cluster sparse regime. Consequently, clusters overlap each other with high probability (w.h.p.). The previous work [20] identifies each point in the network being covered by $\Theta(mR^2/n) = \Theta(n^{1+2\beta-1})$ clusters w.h.p., which indicates that nodes are almost distributed uniformly over the whole area of the network. We recall the scheduling scheme in cluster sparse regime that two mechanisms (inter-cluster duplication and intra-cluster duplication) of duplicating messages are proposed to deliver the source packet to its destination as quickly as possible. Intuitively, we can still apply these two mechanisms to the performance analysis in cluster dense region. However, we have observed some special phenomena during the analysis, which suggests adapting the message duplicating mechanism to the cluster dense regime. Thus, the deducing process of the case of cluster dense regime is very different from that of the case of cluster sparse regime, which is more complicated.

\textbf{A. Topology Analysis and Some Observations}

In cluster sparse regime, the correlation of node mobility is strong where either the number of clusters or the area that each cluster covers is relatively small. Thus, all clusters are distributed sparsely over the network and rarely overlap each other. One cluster should move over slots to meet another cluster and duplicate messages of the source. However in cluster dense regime, as the correlation degree of node mobility has changed, the traditional scheduling scheme used in cluster sparse regime may be inefficient. In the following, we first take a look at some interesting observations.

We consider the case where a source cluster (or a cluster containing relays) meets another cluster which does’t contain messages of the source. Intuitively, the larger number of message holders\(^3\) in one cluster, the higher probability we get to successfully hand over messages to another cluster who are not containing any message holder. However, the following lemma barely supports the above common view.

\textbf{Lemma 1.} For a particular source and its messages to be transmitted, define a cluster containing messages of the source as holder-cluster and a cluster who contains no messages of the source as empty-cluster. Then the probability of successful delivery of messages from holder-cluster to empty-cluster is independent of the number of relays who hold messages of the source in holder-cluster. We assume the transmission range as $r = o(R)$.

\textbf{Proof:} The logic of proof is outlined as follows: first, we demonstrate the best case where all nodes in holder-cluster contain messages of the source, after which we show the worst case where only one node in holder-cluster holds messages of the source. Then we compare the probabilities of successful

\(^3\)Message holders represent the relay nodes who contain messages of the source.
delivery of messages from holder-cluster to empty-cluster in above two cases and make a conclusion.

1) **Best Case:** Assume all nodes in holder-cluster contain messages of the source. For messages being successfully delivered to the empty-cluster, we consider a critical situation where the center of holder-cluster \( A \) is \( 2R + r \) apart from the center of empty-cluster \( B \). In this case, the delivery of messages can be operated on borders of the two clusters (We assume nodes are uniformly distributed in each cluster and there exist nodes near the boundary w.h.p.). Thus we concludes that the delivery of messages is successful w.h.p. if the distance between centers \( A \) and \( B \) is less than \( 2R + r \). Then we derive the probability of successful message delivery from holder-cluster to empty-cluster as 

\[
\Pr \left( X < \frac{(2R + r)^2}{n} \right) = \Theta \left( \frac{R^2}{n} \right).
\]

Thus, the probability of successful message delivery from holder-cluster to empty cluster are the same (in order sense) in above two extreme cases.

2) **Worst Case:** Assume only one node in holder-cluster contains messages of the source. As illustrated in Fig.3, even when two clusters overlap with each other (the distance between \( A \) and \( B \) is \( s \), where \( s < 2R \)), we cannot guarantee that the node conveying messages of the source exists in or near the overlapping region. Thus we formulate the problem of successful message delivery as a conditional probability of \( Pr[Y | X] Pr[X] \), where \( X \) denotes the event that center \( B \) falls into the circular area of radius \( 2R + r \) centered at \( A \) and \( Y \) denotes the event that the only node carrying messages of the source falls into the overlapping area. Then the probability of successful message delivery from holder-cluster to empty cluster can be calculated as:

\[
P = \int P_r[Y | X] P_r[X] dX
\]

\[
= \int_0^{2R} \frac{\pi R^2 \text{arccos} \left( \frac{s^2}{2R} \right) - \frac{s}{2} \sqrt{R^2 - \frac{s^2}{4}} \pi (2R + r)^2}{n} ds
\]

\[
= \int_0^{2R} \frac{4}{n} \left[ \pi R^2 \text{arccos} \left( \frac{s}{2R} \right) - \frac{s}{2} \sqrt{R^2 - \frac{s^2}{4}} \right] ds
\]

\[
= \int_0^{1} \frac{4}{n} \left[ \pi R^2 \text{arccos} (t) - R^2 t \sqrt{1 - t^2} \right] dt
\]

\[
= \left( \pi - \frac{1}{2} \right) \frac{4R^2}{n} = \Theta \left( \frac{R^2}{n} \right).
\]

Based on Lemma 1, we know that it is nonsense to add intra-cluster duplications for increasing the probability of successful message delivery from holder-cluster to empty-cluster. Next, the following two lemmas identify the disadvantages of creating inter-cluster duplications via traditional broadcast and one-hop unicast scheme.

**Lemma 2.** In cluster dense regime, an area of \( \Theta(R^2) \) is covered by \( \Theta(mR^2/n) \) clusters.

**Proof:** We assume that \( X_i^{oc} \) denotes the event that a particular cluster overlap a certain area of \( \Theta(R^2) \) within the network (If the event happens, then we have \( X_i^{oc} = 1 \). Otherwise \( X_i^{oc} = 0 \)). From Fig. 2, we can obtain

\[
P[X_i^{oc} = 1] = \frac{(2R + r)^2}{n} \leq \frac{9R^2}{n}
\]

\[
= \frac{(2R + r)^2}{n} \geq \frac{4R^2}{n}.
\]

We further denote \( X^{oc} \) as the total number of clusters which overlap a certain area of \( \Theta(R^2) \), where \( X^{oc} = \sum_{i=1}^{n} X_i^{oc} \).
Using the multiplicative form of Chernoff bound, we have

\[ P[X_{\text{min}}^\text{oc}] = P[X^\text{oc} > \frac{18mR^2}{n}] < \left( \frac{e^{\frac{9mn^2}{4}}}{n} \right) < O\left( \frac{1}{n} \right), \]

\[ P[X_{\text{max}}^\text{oc}] = P[X^\text{oc} < \frac{2mR^2}{n}] < e^{-\frac{mR^2}{2n}} < O\left( \frac{1}{n} \right). \]

Since \( 0 \leq X^\text{oc} \leq m \), we have

\[ E[X^\text{oc}] = E\left[ X^\text{oc}\mathbb{1}_{\{X^\text{oc}\leq m\text{min}\}} \right] + E\left[ X^\text{oc}\mathbb{1}_{\{X^\text{oc}\leq m\text{max}\}} \right] \]
\[ \leq \frac{18mR^2}{n} + \frac{m}{n} \leq 19mR^2, \]

\[ E[X^\text{oc}] = E\left[ X^\text{oc}\mathbb{1}_{\{X^\text{oc}\leq m\text{max}\}} \right] + E\left[ X^\text{oc}\mathbb{1}_{\{X^\text{oc}\leq m\text{max}\}} \right] \]
\[ \geq \frac{mR^2}{n} + \frac{1}{n} = \frac{mR^2}{n}. \]

Then

\[ \frac{mR^2}{n} \leq E[X^\text{oc}] \leq 19mR^2. \]

So an area of \( \Theta(R^2) \) is covered by \( \Theta(mR^2/n) \) clusters.

**Lemma 3.** If we have already created \( R_x \) inter-cluster duplications, where \( R_x \leq \Theta(m) \), each point will be covered by at least \( \Theta(mR^2/n) \) empty-clusters.

**Proof:** We assume that \( X_i^\text{ec} \) denotes the event that a certain point within the network is covered by an empty-cluster. (If the event happens, then we have \( X_i^\text{ec} = 1 \). Otherwise \( X_i^\text{ec} = 0 \).) Similarly, define \( X^\text{ec} \) as the total number of empty-clusters which cover a certain point within the network, where \( X^\text{ec} = \sum_{i=1}^{mR^2/n} X_i^\text{ec} \). As there still exist \( m - R_x = \Theta(m) \) empty-clusters, we have \( P[X_i^\text{ec} = 1] = \Theta(1) \). According to Chernoff bound, we have

\[ P[X_{\text{min}}^\text{ec}] = P\left[ X^\text{ec} < \Theta\left( \frac{mR^2}{n} \right) \right] < e^{-\frac{mn^2}{4}} < O\left( \frac{1}{n} \right). \]

Since \( 0 \leq X^\text{ec} \leq m \), we have

\[ E[X^\text{ec}] = E\left[ X^\text{ec}\mathbb{1}_{\{X^\text{ec}\leq m\text{min}\}} \right] + E\left[ X^\text{ec}\mathbb{1}_{\{X^\text{ec}\leq m\text{max}\}} \right] \]
\[ \geq \Theta\left( \frac{mR^2}{n} \right) + 0 \frac{1}{n} = \Theta\left( \frac{mR^2}{n} \right). \]

Lemma 2 tells us that clusters overlap with each other w.h.p. in cluster dense regime, so messages can be simultaneously transmitted to several empty-clusters via broadcasting and it is much more efficient than the traditional one-hop unicast delivery when two clusters meet. Intuitively, we can still conduct the traditional broadcast scheme, which could be performed even better due to the highly overlapping feature of clusters in cluster dense regime. However, Lemma 3 proves that when the number of inter-cluster duplications is less than \( \Theta(m) \), each node will still be covered by at least \( \Theta\left( \frac{mR^2}{n} \right) \) clusters which contain no duplications of the source message. This observation indicates that we should restrict the transmission power to some extents during each time of broadcast for making the \( \Theta\left( \frac{mR^2}{n} \right) \) empty clusters receive the duplications. Otherwise, the overhigh transmission power not only does no good to the message dissemination, but also causes extra interference to other concurrent transmissions.

So we apply \( u \) times broadcast scheme with broadcast area of \( A_d \in [1, \Theta(mR^2/n)] \). First, the source node will broadcast messages to the surrounding relay nodes. Then in the second time of broadcast, the source node and the relay nodes who receive messages in the first time of broadcast simultaneously broadcast messages to other relay nodes. This process will carry on until the message is captured by the cluster of the targeting destination or each cluster holds the duplication of the source message.

Now we introduce the general causal scheduling policy A in cluster dense regime, which is illustrated in Fig. 4. Opportunistic broadcast scheme is still applied here. For a particular message,

1. Nodes containing a certain message create relays via \( k \)th broadcast with broadcast area \( A_d \in [1, \Theta(u)] \). The total number of inter-cluster duplications is \( \Theta(u) \).
2. If one of the relays is captured by a node in \( C_d \) within range \( d^k \), the message will be transmitted to the node via \( h^k \)-hop unicast transmission. If not, we come back to step 1.
3. The captured relay in \( C_d \) create new intra-cluster duplications via broadcast (\( R^d_c \) denotes the overall number of intra-cluster duplications within \( C_d \)).
4. If one of the intra-cluster duplications is captured by its destination within range \( d^k \), the message will be transmitted to the destination via \( h^k \)-hop unicast transmission. If not, we come back to step 3.

This general causal scheduling policy A only performs well when the correlation of node mobility is relatively strong. When the correlation of node mobility becomes extremely strong, we shall re-design the scheduling policy to achieve the optimal network performance, which will be analyzed in Step 2)

\[ \text{Step 1) } \]

\[ \text{Step 2) } \]

\[ \text{Step 3) } \]

\[ \text{Step 4) } \]
Section IV-H. In the following analysis, we divide the general causal scheduling policy A into two parts based on the four steps in Fig. 4. One is step 1-2) (Part I) and the other is step 3-4) (Part II).

**Remark 1:** Before studying the capacity-delay tradeoff, we briefly outline the logic sequence of the derivation of the tradeoff. First we explore the fundamental relationship between the delay, capacity and the scheduling parameters including the number of relays, the size of capture region, the number of hops, etc. These scheduling parameters correlate closely to the network performance. Then, we establish formulas to depict the quantitative relationship between them, which can be utilized to derive the upper bound of capacity-delay tradeoff later. In addition, the deducing process is complex with much mathematical tools being involved. We try hard to simplify it and emphasize key issues.

### B. Tradeoff for Delay of Part I

We denote $D^d_{I3}$ as the delay of creating $R^d_{ch}$ inter-cluster duplications until one of the relay nodes is captured by a particular node in $C_d$. In addition, we denote $D^d_{II1}$ as the delay of delivering the message from the captured relay node to the node in $C_d$.

Based on scheduling policy A, $D^d_{II1}$ is equivalent to the time consumption of $u$ times broadcast. Each time of broadcast, at least $\Theta\left(\frac{mR^d_{ch}}{n}\right)$ empty clusters will receive duplications of the source message from one relay node. Assume $R^d_{ch} = \omega(1) = n^\hat{a}$ and $A_d = n^\alpha$, where $\hat{a}$ and $\alpha$ are two constants greater than 0. We can easily get the equation of $(A_d)^u = R^d_{ch}$, then we have:

$$u = \frac{\log n^\hat{a}}{\log n^\alpha} = \frac{\hat{a}}{\alpha} = \Theta(1).$$

Specifically, for the case of $R^d_{ch} = \Theta(1)$ or the cluster critical regime, the delay $D^d_{I1}$ can still be bounded by up to $\Theta(\log n)$, which is neglectable in order sense.

Next, we calculate the delay $D^d_{II1}$ as follow:

$$D^d_{II1} = \frac{1}{\left(1 - \left(1 - \frac{R^d_{ch}}{R^d}ight)^2\right)} \frac{i^2c^2n/m}{R^d} \geq \frac{m}{R^d|d|^2}.$$  

According to the analysis of $D^d_{I3}$ and $D^d_{II1}$, we show the whole delay of Part I through the lemma below.

**Lemma 4.** Under scheduling policy A in cluster dense regime, the delay for a particular bit $b$ of Part I $D^d_{ch}$ and the associated scheduling parameters comply with the following inequality.

$$c^d_1 \log n E[D^d_{ch}] \geq \frac{m}{E[R^d_{ch}]} E\left[l^d_{ch} + \frac{1}{\hat{a}}\right]^2,$$  

where $c^d_1$ is a positive constant. We use $l^d_{ch}$ here to denote $l^d_{h}$ of a particular bit $b$, which is convenient for the subsequent analysis. The similar rule holds for the other variables.

The proof of Lemma 4 is similar to that in Appendix A of our work [19], which is omitted here.

### C. Tradeoff for Radio Resource, Half Duplex and Multihop of Part I

Based on Lemma 4, we can find that the delay can be reduced if we increase either the number of relays or the capture range. Because a larger number of relays results in a higher probability of the incidence that the packet is captured by the destination. This reason also holds for increasing the capture range.

However, more relays generated, more radio resources consumed, which would decrease the network capacity. In addition, as the capture range increases, the number of concurrent transmissions within the area that the capture region covers is reduced, which poses a negative impact on the capacity performance. The following Lemma verifies the aforementioned idea and presents us the fundamental relationship between the capacity and the related scheduling parameters.

**Lemma 5.** Under scheduling policy A in cluster dense regime, we have the following inequality hold to tradeoff the capacity and the associated scheduling parameters.

$$\sum_{b=1}^{\lambda nT} \frac{\Delta^2 E[R^d_{ch}] - 1}{n} \sum_{h=1}^{\lambda nT} \frac{h^d_{ch} \pi \beta^2 r^h_b}{4n} \leq c^d_2 WT \log n,$$

where $c^d_2$ is a positive number, $h^d_{ch}$ is the number of hops to reach $C_d$ after being captured, and $r^h_b$ is the transmission range of each hop.

**Proof:** The proof is similar to that in [19], which is omitted here for simplicity.

Based on the half duplex transmission mechanism and the nature of multihop scheme, we have the following inequalities hold.

**Lemma 6.** The following inequality holds,

$$\sum_{b=1}^{\lambda nT} \sum_{h=1}^{h^d_{ch}} 1 \leq \frac{WT}{2} n.$$  

**Lemma 7.** The following inequality holds,

$$\sum_{b=1}^{\lambda nT} \sum_{h=1}^{h^d_{ch}} r^h_b \geq l^d_{ch}.$$  

---

### Table II: The optimal values of scheduling parameters under scheduling policy A of part I ($d = \frac{3-\alpha-6\tilde{d}}{2}$).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^d_{ch}$</td>
<td>$\Theta(n^{\frac{\alpha-\hat{a}}{\alpha}})$</td>
</tr>
<tr>
<td>$l^d_{ch}$</td>
<td>$\Theta(n^{\frac{\alpha-\hat{a}}{\alpha}})$</td>
</tr>
<tr>
<td>$h^d_{ch}$</td>
<td>$\Theta(n^{\frac{\alpha-\hat{a}}{\alpha}})$</td>
</tr>
<tr>
<td>$r^h_b$</td>
<td>$\Theta(\log^\frac{\alpha}{\hat{a}} n)$</td>
</tr>
</tbody>
</table>
D. Detailed Upper Bound on Capacity-Delay Tradeoff of Part I

In this section, we mainly derive the capacity-delay tradeoff on the basis of fundamental tradeoffs established in Section IV-B and IV-C. Then we achieve the optimal values of scheduling parameters by making the upper bound tight.

From Lemma 4, we have

\[ \sum_{b=1}^{\lambda_{d}^{r} n T} \mathbb{E} [R_{cb}^{d}] \geq \frac{1}{c_{1}^{d} \log n} \sum_{b=1}^{m} \left( \mathbb{E} \left[ l_{cb}^{d} \right] + \frac{1}{n} \right)^{2} \mathbb{E} \left[ D_{1b}^{d} \right] \]

\[ \geq \frac{m}{c_{1}^{d} \log n} \sum_{b=1}^{\lambda_{d}^{r} n T} \frac{1}{(\sum_{b=1}^{\lambda_{d}^{r} n T})^{3}} \left( \sum_{b=1}^{\lambda_{d}^{r} n T} \left( \mathbb{E} \left[ l_{cb}^{d} \right] + \frac{1}{n} \right) \right)^{2} \]

\[ = \frac{m}{c_{1}^{d} \log n} D_{1} \left( \sum_{b=1}^{\lambda_{d}^{r} n T} \left( \mathbb{E} \left[ l_{cb}^{d} \right] + \frac{1}{n} \right) \right)^{2} \]  

Inequality (5) is deduced by using Jensen’s Inequality and Hölder’s Inequality. From Lemma 5 and Cauchy-Schwartz inequality, we obtain

\[ \sum_{b=1}^{\lambda_{d}^{r} n T} \frac{\Delta^{2} \mathbb{E} \left[ R_{cb}^{d} \right]}{n} - 1 + \frac{\pi \Delta^{2}}{2W T n^{2}} \left( \sum_{b=1}^{\lambda_{d}^{r} n T} \mathbb{E} \left[ l_{cb}^{d} \right] \right)^{2} \leq c_{2}^{d} W T \log n \]

\[ \frac{\Delta^{2} m}{4 c_{1}^{d} n \log n} D_{1} \left( \sum_{b=1}^{\lambda_{d}^{r} n T} \left( \mathbb{E} \left[ l_{cb}^{d} \right] + \frac{1}{n} \right) \right)^{2} + \frac{\pi \Delta^{2}}{2W T n^{2}} \left( \sum_{b=1}^{\lambda_{d}^{r} n T} \mathbb{E} \left[ l_{cb}^{d} \right] \right)^{2} - \frac{\Delta^{2}}{4} \lambda_{d}^{r} T \leq c_{2}^{d} W T \log n \]

If \( \sum_{b=1}^{\lambda_{d}^{r} n T} \left[ l_{cb}^{d} \right] < \lambda_{d}^{r} T \),

\[ \frac{\Delta^{2} m}{4 c_{1}^{d} n \log n} \left( D_{1} \left( \lambda_{d}^{r} T \right)^{3} \right) n^{2} \leq c_{2}^{d} W T \log n \]

\[ \lambda_{d}^{r} \leq \frac{4 \pi c_{1}^{d} W T \tilde{D}_{1} \log n}{\Delta^{2} n^{3}} \]  

If \( \sum_{b=1}^{\lambda_{d}^{r} n T} \left[ l_{cb}^{d} \right] \geq \lambda_{d}^{r} T \),

\[ \frac{\Delta^{2} m}{4 c_{1}^{d} n \log n} \left( \sum_{b=1}^{\lambda_{d}^{r} n T} \left( \mathbb{E} \left[ l_{cb}^{d} \right] \right)^{2} \right) + \frac{\pi \Delta^{2}}{2W T n^{2}} \left( \sum_{b=1}^{\lambda_{d}^{r} n T} \mathbb{E} \left[ l_{cb}^{d} \right] \right)^{2} \leq c_{2}^{d} W T \log n \]

\[ \sqrt{\frac{\pi \Delta^{2} n^{2}}{8 c_{1}^{d} W \log n}} \left( \lambda_{d}^{r} \right)^{m} D_{1} \leq c_{2}^{d} W T \log n \]  

TABLE III: The optimal values of scheduling parameters under scheduling policy A of part I (\( \tilde{d} < \frac{3-v-6\beta}{2} \)).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_{cb}^{d} )</td>
<td>( \Theta(n^{1-2\beta-d}/\log n) )</td>
</tr>
<tr>
<td>( l_{cb}^{d} )</td>
<td>( \Theta(n^{v+2\beta-1}/\log^{2} n) )</td>
</tr>
<tr>
<td>( l_{cb}^{d} )</td>
<td>( \Theta(n^{v+2\beta-1}/\log n) )</td>
</tr>
<tr>
<td>( c_{b}^{d} )</td>
<td>( \Theta(\log^{2} n) )</td>
</tr>
</tbody>
</table>

(\( \lambda_{d}^{r} \))

In order to obtain the tight upper bound of the capacity-delay tradeoff, the four inequalities (1), (3), (4) and (7) should hold. Denote the delay of Part I as \( \Theta(n^{d}) \), then we can achieve the optimal values of scheduling parameters, which are shown in Table II.

Remark: According to the scheduling policy A, the optimal scheduling parameters should meet the two constraints \( 1 \leq R_{cb}^{d} \leq m \) and \( 1 \leq l_{cb}^{d} \leq \sqrt{m R^{2}/n} \). Thus, we have \( \tilde{d} \geq (3-v-6\beta)/2 \) in regardless of the logarithmic factor. If further investigating the case of \( \tilde{d} < (3-v-6\beta)/2 \), which indicates the smaller delay, we should enlarge the capture range \( l_{cb}^{d} \). Note that the above two constraints should still be satisfied, so we set the capture range as \( l_{cb}^{d} = \sqrt{m R^{2}/n} \). Then we recalculate the upper bound of capacity-delay tradeoff on condition that \( l_{cb}^{d} = \sqrt{m R^{2}/n} \), which follows the same deducing logic as in the case of \( \tilde{d} \geq (3-v-6\beta)/2 \).

\[ \lambda_{d}^{r} \leq O \left( \frac{R^{2} \tilde{D}_{1}}{n} \log^{3} n \right) \]

The corresponding optimal values of scheduling parameters are shown in Table III.

Theorem 2. Under scheduling policy A in cluster dense regime, let \( \tilde{D}_{1}^{r} \) denote the mean delay averaged over all bits and let \( \lambda_{d}^{r} \) be the capacity of each source-destination pair. In Part I, the following upper bound holds,

\[ \left\{ \begin{array}{l}
(\lambda_{d}^{r})^{3} \leq O \left( \frac{\tilde{D}_{1}^{r} \log^{3} n}{\log n} \right) \quad \tilde{d} \geq \frac{3-v-6\beta}{2}

\lambda_{d}^{r} \leq O \left( \frac{R^{2} \tilde{D}_{1}^{r}}{n} \log^{3} n \right) \quad \tilde{d} < \frac{3-v-6\beta}{2}
\end{array} \right. \]
E. Tradeoff for Delay of Part II

Denote the delay of delivering the message from the captured node to the destination within $C_d$ as $D^d_{12}$, then we have
\[
D^d_{12} = \frac{1}{\left(1 - (1 - \frac{n^2}{R^2})^{R^2}ight)} \geq \frac{R^2}{R^d D^d_{12}}.
\]
Similar to the analysis in Section IV-B, the delay of creating intra-cluster duplications can be bounded as a constant, which is omitted here.

**Lemma 8.** Under scheduling policy $A$ in cluster dense regime, the delay for a particular bit $b$ of Part II and the associated scheduling parameters comply with the following inequality
\[
\begin{align*}
\frac{c^d_3}{\log n} & \leq \frac{R^2}{\mathbb{E} \left[ R^d_{db} \right] \mathbb{E} \left[ t^d_{2b} \right] + \frac{1}{n}}. \\
\end{align*}
\]

The proof of Lemma 8 is similar to the proof in Appendix A of our work [19], so we omit it for simplification.

F. Tradeoff for Radio Resource, Half Duplex and Multihop of Part II

Similar to Part I, we obtain the following three lemmas corresponding to three fundamental tradeoffs of radio resource, half duplex and multihop with respect to the scheduling parameters.

**Lemma 9.** Under scheduling policy $A$ in cluster dense regime, we have the following inequality hold to tradeoff the capacity and the associated scheduling parameters.
\[
\begin{align*}
\frac{c^d_3}{\log n} & \leq \frac{\Delta^2 m R^2}{\mathbb{E} \left[ R^d_{db} \right] - 1} + \frac{\pi \Delta^2}{\mathbb{E} \left[ t^d_{2b} \right]} \frac{R^2}{\mathbb{E} \left[ D^d_{12} \right]} \left( \sum_{b=1}^{\lambda^d_3 n T} \frac{1}{n} \right)^3 \\
& \leq \frac{c^d_4 WT \log n}{\mathbb{E} \left[ D^d_{12} \right] \mathbb{E} \left[ t^d_{2b} \right] + \frac{1}{n}}. \\
\end{align*}
\]

Based on Lemma 9 and Cauchy-Schwartz inequality, we have
\[
\begin{align*}
\frac{c^d_3}{\log n} & \leq \frac{\Delta^2 m R^2}{\mathbb{E} \left[ R^d_{db} \right] - 1} + \frac{\pi \Delta^2}{2WT n^2} \left( \sum_{b=1}^{\lambda^d_3 n T} \mathbb{E} \left[ t^d_{2b} \right] \right)^2 \leq \frac{c^d_4 WT \log n}{\mathbb{E} \left[ D^d_{12} \right] \mathbb{E} \left[ t^d_{2b} \right] + \frac{1}{n}}. \\
\end{align*}
\]

**Lemma 10.** The following inequality holds,
\[
\begin{align*}
\sum_{b=1}^{\lambda^d_3 n T} \frac{h^d_{2b}}{n} & \leq \frac{WT^4}{2n}. \\
\end{align*}
\]

**Lemma 11.** The following inequality holds,
\[
\begin{align*}
\sum_{b=1}^{\lambda^d_3 n T} \frac{r^b}{n} & \geq \frac{t^d_{2b}}{2n}. \\
\end{align*}
\]

G. Detailed Upper Bound on Capacity-Delay Tradeoff of Part II

In this section, we derive the capacity-delay tradeoff within the cluster of destination. Following the same logical sequence for deriving the capacity-delay tradeoff of part I, we also calculate the capacity-delay tradeoff of part II on the basis of fundamental tradeoffs established in Section IV-E and IV-F. At last, we obtain the optimal values of scheduling parameters by making the upper bound tight.

**TABLE IV:** The optimal values of scheduling parameters under scheduling policy $A$ of part II ($d \leq 2 - 2\nu - 2\beta$).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^d_{db}$</td>
<td>$\Theta n^2$</td>
</tr>
<tr>
<td>$t^d_{2b}$</td>
<td>$\Theta \left( \frac{\nu + \Delta - 1 - d}{3} / \log n \right)$</td>
</tr>
<tr>
<td>$h^d_{2b}$</td>
<td>$\Theta \left( \frac{\nu + \Delta - 1 - d}{3} / \log n \right)$</td>
</tr>
<tr>
<td>$r^b$</td>
<td>$\Theta \left( \log n^2 \right)$</td>
</tr>
</tbody>
</table>

From Lemma 8, we have
\[
\begin{align*}
\sum_{b=1}^{\lambda^d_3 n T} \mathbb{E} \left[ R^d_{db} \right] & \geq \frac{1}{c^d_3 \log n} \sum_{b=1}^{\lambda^d_3 n T} \mathbb{E} \left[ R^d_{db} \right] \mathbb{E} \left[ t^d_{2b} \right] + \frac{1}{n} \mathbb{E} \left[ D^d_{12} \right] \\
& \geq \frac{R^2}{c^d_3 \log n} \mathbb{E} \left[ D^d_{12} \right] \left( \sum_{b=1}^{\lambda^d_3 n T} \frac{1}{n} \right)^3 \\
& \geq \frac{R^2}{c^d_4 \log n} \mathbb{E} \left[ D^d_{12} \right] \left( \sum_{b=1}^{\lambda^d_3 n T} \mathbb{E} \left[ t^d_{2b} \right] + \frac{1}{n} \right)^2. \\
\end{align*}
\]

If $\sum_{b=1}^{\lambda^d_3 n T} \mathbb{E} \left[ t^d_{2b} \right] < \lambda^d_3 T$,
\[
\begin{align*}
\frac{\Delta^2 m R^4}{4c^d_3 n^2 \log n} \frac{\left( \lambda^d_3 n T \right)^3 n^2}{D^d_{12} \left( \lambda^d T \right)^2} & \leq \frac{c^d_4 WT \log n}{\mathbb{E} \left[ D^d_{12} \right] \mathbb{E} \left[ t^d_{2b} \right] + \frac{1}{n}}. \\
\end{align*}
\]

\[
\begin{align*}
\lambda^d_3 & \leq \frac{4c^d_4 WT D^d_{12} \log^2 n}{\Delta^2 n^4 m R^2 T}. \\
\end{align*}
\]
If $\sum_{b=1}^{\lambda_2^d} [t_{2b}^d] \geq \lambda_2^d T$, then
\[
\frac{\Delta^2 m R^4}{4c_1^d n^2 \log n} \sum_{b=1}^{\lambda_2^d} (E[t_{2b}^d])^2 + \frac{\pi \Delta^2}{2WTn^2} \sum_{b=1}^{\lambda_2^d} E[t_{2b}^d] \leq c_4^d WT \log n 
\]
\[
\frac{\pi \Delta^2 T^2}{8c_4^d W \log n} \frac{(\lambda_2^d)^3 m R^4}{nD_2^4} \leq c_4^d WT \log n 
\]
\[
(\lambda_2^d)^3 \leq \frac{8c_4^d W^3 n D_2^4 \log^3 n}{\pi \Delta^2 m R^4}.
\]

Compare the two inequalities (13) and (15), we obtain the upper bound as follow
\[
(\lambda_2^d)^3 \leq O \left( \frac{nD_2^4 \log^3 n}{m R^4} \right).
\]

In order to obtain the tight upper bound of the tradeoff, inequalities (9), (11), (12) and (14) should hold. Denote the delay as $\Theta(n^d)$, then we can achieve the optimal values of scheduling parameters shown in Table IV.

**Remark:** As the average number of nodes in each cluster is $\Theta \left( \frac{n}{\bar{d}} \right)$, thus the constraint $1 \leq R_{db}^d \leq n/\bar{d}$ should be satisfied. In other words, the number of intra-cluster duplications should not exceed the number of total nodes in one cluster. Thus we obtain $\bar{d} \leq 2 - 2\nu - 2\beta$ regardless of the logarithmic factor. If we would like to increase the capacity at the expense of degrading the delay performance ($\bar{d} > 2 - 2\nu - 2\beta$), we should either reduce the number of intra-cluster duplications or the capture range. Note that the aforementioned constraint should still be satisfied, thus we set $R_{db}^d$ as 1 and recalculate the upper bound following the same logic as in the case of $\bar{d} \leq 2 - 2\nu - 2\beta$.

\[
(\lambda_2^d)^3 \leq O \left( \frac{D_2^4 \log^3 n}{R^2} \right).
\]

The corresponding optimal values of scheduling parameters are shown in Table V.

**Theorem 3.** Under scheduling policy A in cluster dense regime, let $D_2^d$ denote the mean delay averaged over all bits and let $\lambda_2^d$ be the capacity of each source-destination pair. In Part II, the following upper bound holds,
\[
\begin{align*}
(\lambda_2^d)^2 &\leq O \left( \frac{D_2^4}{n R^2} \log^3 n \right) \quad \bar{d} > 2 - 2\nu - 2\beta \\
(\lambda_2^d)^3 &\leq O \left( \frac{nD_2^4}{m R^4} \log^3 n \right) \quad \bar{d} \leq 2 - 2\nu - 2\beta.
\end{align*}
\]

Theorem 2 and 3 show us the tradeoff of Part I and Part II of scheduling policy A respectively. Then the whole upper bound of capacity-delay tradeoff of scheduling policy A can be derived easily.

**Theorem 4.** Under scheduling policy A in cluster dense regime, let $D_2^d_A$ denote the mean delay averaged over all bits and let $\lambda_2^d_A$ be the capacity of each source-destination pair. Assume $D_1^d = D_2^d = D^d_A$. Then we have,
\[
\lambda_2^d_A = \min \{ \lambda_2^d, \lambda_2^d \}.
\]

**H. Scheduling Policy B and Overall Upper Bound of Capacity-Delay Tradeoff**

Previously we derived the upper bound of capacity-delay tradeoff in the case of relatively weak node correlation. The scheduling policy A was designed for better utilizing the characteristics of weak node correlation. However, as the node correlation is getting weaker (i.e., either the number of clusters or the area of each cluster tends to be even larger), the scheduling policy A no longer performs well. Because the weaker node correlation results in the stronger overlapping among clusters. Thus, the scheduling policy A fails to take advantage of the intra-cluster transmission (mainly refers to the transmission within $C_d$) which can save time and radio resources. Based on the new situation where nodes show extremely weak correlation, we have modified the scheduling policy A and newly developed the scheduling policy B, which is illustrated in Fig 5.

1. Nodes containing a certain message create relays via $k$th broadcast with area $A_d (k = 1, ..., \Theta(n^d); R_b^d$ denotes the overall number of relays).
2. If one relay is captured by the destination within range $l_{2b}^d$, the message will be transmitted to the destination via $h_{2b}^d$-hop unicast transmission. If not, we come back to step 1).

![Fig. 5: Scheduling policy B in cluster dense regime.](image-url)
can be similarly derived as that in [6]. In the following, we only show the basic tradeoff of delay and directly give the upper bound of capacity-delay tradeoff of scheduling policy B.

**Lemma 12.** Under scheduling policy B in cluster dense regime, the delay for a particular bit b and the associated scheduling parameters comply with the following inequality

\[ c_B^d \log n \mathbb{E} \left[ D_B^d \right] \geq \frac{n}{\mathbb{E} [R_B^d] \mathbb{E} \left[ l_B^d + \frac{1}{n} \right]^2}, \]  

(16)

where \( c_B^d \) is a positive constant.

**Proof:** Here we only give the intuitive proof. The detailed proof is similar to previous lemmas.

We assume that there are \( R_B^d \) relays in the i-th cluster, where \( i = 1, 2, \ldots, m \). Denote \( X^{cap} \) as the event that at least one relay is captured by the destination, then when \( l_B^d \leq \Theta(R) \):

\[
P[X^{cap}] \leq 1 - \prod_{i=1}^{m} \left( 1 - \frac{(R - l_B^d)^2}{n} \left( \frac{l_B^d}{R} \right)^2 \right) \leq 1 - \left( 1 - \frac{c_B^d \left( l_B^d \right)^2}{n} \right)^{R_B^d} \leq \frac{c_B^d \left( l_B^d \right)^2}{n} R_B^d.
\]

When \( l_B^d > \Theta(R) \): we treat relays as \( \Theta(R_B^d) \) inter-cluster duplications, because they are created by \( u \) times broadcast.

\[
P[X^{cap}] \leq 1 - \left( 1 - \frac{c_B^d \left( l_B^d \right)^2}{n} \right)^{R_B^d} \leq \frac{c_B^d \left( l_B^d \right)^2}{n} R_B^d.
\]

Then the average packet delay is:

\[ D_B^d = \frac{1}{P[X^{cap}]} \geq \frac{n}{c_B^d R_B^d \left( l_B^d \right)^2}. \]

By Hölder’s Inequality and Jensen’s Inequality, we have

\[ c_B^d \mathbb{E} \left[ D_B^d \right] \geq \frac{n}{\mathbb{E} [R_B^d] \mathbb{E} \left[ l_B^d \right]^2}. \]

If the number of needed relays is larger than the number of clusters \( m \) when \( l_B^d > \Theta(R) \), this basic tradeoff (inequality (16)) will no longer hold. Fortunately, we prove that it will not happen by contradiction. If \( R_B^d > m \) when \( l_B^d > R \), we have \( (1 - d)/3 > v \) and \( 2(1 - d)/3 > 2\beta \). Combine the two equalities, we have \( 1 - d > v + 2\beta \geq 1 \) which will make sense only if \( d < 0 \). Thus the number of needed relays will not be larger than \( m \) when \( l_B^d > \Theta(R) \).

**Theorem 5.** Under scheduling policy B in cluster dense regime, let \( D_B^d \) denote the mean delay averaged over all bits and let \( \lambda_B^d \) be the capacity of each source-destination pair. The following upper bound holds,

\[ (\lambda_B^d)^3 \leq O \left( \frac{D_B^d}{n} \log^3 n \right). \]

So far we have both obtained the upper bounds of capacity-delay tradeoff of scheduling policy A and B in cluster dense regime, of which two results are presented in Theorem 4 and 5 respectively. Assume \( D_A^d = D_B^d = D^d_d \), then the overall upper bound of capacity-delay tradeoff in cluster dense regime can be derived easily.

**Theorem 6.** In cluster dense regime, let \( D^d \) denote the mean delay averaged over all bits and let \( \lambda^d \) be the capacity of each source-destination pair. Assume \( D_A^d = D_B^d = D^d \), then the following upper bound holds,

\[ \lambda^d = \max \{ \lambda_A^d, \lambda_B^d \}. \]

Note that scheduling policy A and B are both applicable to the cluster dense regime. Under what circumstance we should choose A or B to operate highly depends on the correlation degree of node mobility. Generally speaking, If \( \lambda_A^d \geq \lambda_B^d \), we use scheduling policy A. Otherwise, we use scheduling policy B.

V. LOWER BOUND OF THE CLUSTER DENSE REGIME

We mainly investigate the achievable lower bound for cluster dense regime in this section. Similar to the derivation of the upper bound, we calculate the lower bound under scheduling policy A and B respectively. Given certain network condition \( m = n^e, R = n^3, D^d = n^5, \) we choose lower bound A, if \( \lambda_A^d \geq \lambda_B^d \). Otherwise, we choose lower bound B.

A. Lower Bound of Scheduling Policy A

Under scheduling policy A, we divide unit time slot into four sub-slot. The basic operation of each sub-slot are shown below:

1. Nodes (source node and relays) create inter-cluster duplications via \( u \) times broadcast.
2. One of the inter-cluster duplications is captured by a node in \( C_d \) within range \( l_A^d \) and is transmitted to the node via \( h_A^d \)-hop unicest transmission. Each hop exploits the transmission range of \( r_A^d \).
3. Nodes in \( C_d \) create \( R_A^d \) intra-cluster duplications via broadcast within \( C_d \).
4. One of the intra-cluster duplications is captured by the destination within range \( l_A^d \) and is transmitted to the destination via \( h_A^d \)-hopunicest transmission. Each hop exploits the transmission range of \( r_A^d \).

The value of scheduling parameters in the above scheme for achieving the lower bound are selected from Table II, III, IV, and V on different delay-tolerant conditions and system parameters. Each cell owns a constant fraction \( 1/c_B^d \) of time to transmit by utilizing the time division multiple access (TDMA) mechanism. In the following, we in detail describe the lower bound achieving scheme.

1. In sub-slot 1, we divide the whole network into \( \cap_1^d = [n/(A_d \log n)] \) cells of equal area, where \( A_d = \min \{m R^2/n, R_A^d \} \) denotes the broadcast area. Nodes in each cell take turns to broadcast for at least \( \lambda^d/(c_B^d \log n) \) fraction of one sub-slot, where \( \lambda^d/(c_B^d \log n) \leq 1/(c_B^d A_d \log n) \). If
any cell contains more than \( c_d^5 \log n / \lambda^d \) nodes, we call the operation as \( \text{Error}^{IA}_d \).

2) In sub-slot 2, we divide the whole network into \( \Theta^d_2 = \left[ n / \left( \left( \frac{\lambda^d}{t_{db}^d} \right)^2 \log n \right) \right] \) cells of equal area. If a certain message \( b \) is not captured within \( \Theta(D^d) \) slots, we call it as \( \text{Error}^{IA}_d \). Otherwise, we further divide each cell into \( \Theta^d_3 = \left( \frac{\lambda^d}{t_{db}^d} \right)^2 \) mini cells of equal area. The captured message is transmitted to a node in \( C_d \) through mini cells via multihop transmission during sub-slot 2 (first along horizontal data path, then along vertical data path). If the network fails to transmit the captured message to a node in \( C_d \) within a certain sub-slot, we call the operation as \( \text{Error}^{IA}_d \).

3) In sub-slot 3, we divide the whole network into \( \Theta^d_4 = \left[ n / \left( m R^2 R_{db}^d \log n \right) \right] \) cells of equal area. Nodes in each cell turn to broadcast for at least \( \frac{\lambda^d}{(\frac{\lambda^d}{t_{db}^d})^2 \log n} \) fraction of one sub-slot, where \( \lambda^d / \left( c_d^5 \log n \right) \leq n / \left( c_d^3 m^2 R^2 R_{db}^d \log n \right) \). If any cell contains more than \( c_d^5 \log n / \lambda^d \) nodes, we call the operation as \( \text{Error}^{IV}_d \).

4) In the sub-slot 4, we divide the whole network into \( \Theta^d_5 = \left[ n / \left( \left( \frac{\lambda^d}{t_{db}^d} \right)^2 \log n \right) \right] \) cells of equal area. If a certain message \( b \) is not captured within \( \Theta(D^d) \) slots, we call it as \( \text{Error}^{IV}_d \). Otherwise, we further divide each cell into \( \Theta^d_6 = \left( \frac{\lambda^d}{t_{db}^d} \right)^2 \) mini cells of equal area. The captured message is transmitted to its destination through mini cells via multihop transmission during sub-slot 4 (first along horizontal data path, then along vertical data path). If the network fails to transmit the captured message to its destination within a certain sub-slot, we call the operation as \( \text{Error}^{IV}_d \).

Theorem 7. As \( n \rightarrow \infty \), \( \text{Error}^{IA}_d \rightarrow 0 \), \( \text{Error}^{IIA}_d \rightarrow 0 \), \ldots \( \text{Error}^{IV}_d \rightarrow 0 \). So we can achieve the lower bound of capacity-delay tradeoff of \( \Theta(\lambda^d / \log n) \) per-node capacity with \( \Theta(D^d \log n) \) delay.

Proof: The cases of \( \text{Error}^{IA}_d \), \( \text{Error}^{IIA}_d \) and \( \text{Error}^{IIIA}_d \) are similar to that of \( \text{Error}^{IV}_d \), \( \text{Error}^{IVA}_d \) and \( \text{Error}^{IVA}_d \), so we only prove the cases of \( \text{Error}^{IVA}_d \rightarrow 0 \), \( \text{Error}^{IVA}_d \rightarrow 0 \), and \( \text{Error}^{IVA}_d \rightarrow 0 \).

1) \( \text{Error}^{IVA}_d \): The problem can be modeled as an equivalent experiment, which is much easier to understand. We throw \( n \) balls into \( \Theta^d_4 \) urns, where the number of balls in each urn is denoted as \( X^d_{IV} \). If \( X^d_{IV} > c_d^5 \log n / \lambda^d \), \( \text{Error}^{IVA}_d \) happens,

\[
\mathbb{E}[X^d_{IV}] = \frac{n D^d \log n}{\left( t_{db}^d \right)^2 \log n} = \frac{\log n}{\lambda^d}.
\]

Using the multiplicative form of Chernoff bound,

\[
P[X^d_{IV} > \frac{2 \log n}{\lambda^d}] < \left( \frac{e}{4} \right)^{\frac{n \log n}{\lambda^d}} < O \left( \frac{1}{n} \right).
\]

As \( n \rightarrow \infty \), \( P[X^d_{IV} > \frac{2 \log n}{\lambda^d}] \rightarrow 0 \), which indicates that \( P[\text{Error}^{IVA}_d] \rightarrow 0 \) when \( n \rightarrow \infty \). 

2) \( \text{Error}^{IVA}_d \): Similar to the analysis of \( \text{Error}^{IVA}_d \), we can also model the problem as an equivalent experiment. We throw \( n \) balls into \( \left( R^2 \Theta^d_4 / n \right) \) urns, where the number of balls in each urn is denoted as \( X^d_{IV} \). If \( X^d_{IV} > c_d^5 \log n / \lambda^d \), \( \text{Error}^{IVA}_d \) happens.

\[
\mathbb{E}[X^d_{IV}] = \frac{n D^d \log n}{\left( t_{db}^d \right)^2 \log n} = \frac{\log n}{\lambda^d}.
\]

Using the multiplicative form of Chernoff bound,

\[
P[X^d_{IV} > c_d^5 \log n / \lambda^d] < \left( \frac{e}{4} \right)^{\frac{n \log n}{\lambda^d}} = O \left( \frac{1}{n} \right).
\]

With \( n \rightarrow \infty \), \( P[X^d_{IV} > c_d^5 \log n / \lambda^d] \rightarrow 0 \), which
TABLE VI: The optimal values of scheduling parameters under scheduling Policy B.

<table>
<thead>
<tr>
<th>Policy</th>
<th>Optimal Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_b^d )</td>
<td>( \Theta(n^{(1-d)/3}) )</td>
</tr>
<tr>
<td>( t_b^d )</td>
<td>( \Theta(n^{(1-d)/3} / \log^{1/2} n) )</td>
</tr>
<tr>
<td>( h_b^d )</td>
<td>( \Theta(n^{(1-d)/3} / \log n) )</td>
</tr>
<tr>
<td>( r_b^h )</td>
<td>( \Theta(\log^{2} n) )</td>
</tr>
</tbody>
</table>

indicates that \( P[Error_{V,IA}^d] \rightarrow 0 \), as \( n \rightarrow \infty \) and \( P[Error_{V,IB}^d] \rightarrow 0 \), as \( n \rightarrow \infty \).

B. Lower Bound of Scheduling Policy B

Under scheduling policy B, we divide unit time slot into two sub-slot. The basic operation of each sub-slot are shown below:

1) Nodes (source node and relays) create \( R_b^d \) relays via \( u \) times broadcast.

2) One of the relays is captured by its destination within range \( r_b^h \) and is transmitted to the destination via \( h_b^d \) hop unicast transmission. Each hop exploits the transmission range of \( r_b^h \).

The value of scheduling parameters in the above scheme are selected from Table VI on different delay-tolerant conditions and system parameters. TDMA is applied as well, so each cell can have \( 1/c_b^d \) fraction of time to transmit. In the following, we describe our lower bound achieving scheme.

1) In sub-slot 1, we divide the whole network into \( T^d_1 = [n/(A_d \log n)] \) cells of equal area. Nodes in each cell take turns to be active for at least \( \lambda^d/(c_b^d \log n) \) fraction of one sub-slot to broadcast, where \( \lambda^d/(c_b^d \log n) \leq 1/(c_b^d A_d \log n) \). If any cell contains more than \( c_b^d \log n/\lambda^d \) nodes, we call the operation as \( Error_{V,IB}^d \).

2) In sub-slot 2, we divide the whole network into \( T^d_2 = [n/(l_b^d \log n)] \) cells of equal area. If a certain message \( b \) is not captured by the cell within \( D^d \) slots, we call it as \( Error_{V,IB}^d \). Then we further divide each cell into \( T^d_3 = (h_b^d)^2 \) mini cells of equal areas. The message captured by the cell is transmitted to its destination through mini cells via multihop unicast transmissions during this sub-slot (first along horizontal data path, then along vertical data path). If the network fails to transmit the captured message to its destination within a certain sub-slot, we call the operation as \( Error_{V,IB}^d \).

Theorem 8. As \( n \rightarrow \infty \), \( Error_{V,IB}^d \rightarrow 0 \), \( Error_{V,IB}^d \rightarrow 0 \), \( Error_{V,IB}^d \rightarrow 0 \), \( Error_{V,IB}^d \rightarrow 0 \). So we can achieve the lower bound of capacity-delay tradeoff in cluster dense regime of \( \Theta(\lambda^d / \log n) \) per-node capacity with \( \Theta(D^d \log n) \) delay.

Proof: The proof is similar to that of Theorem 7, which we omit for simplification.

VI. DISCUSSION

The results of the upper bound and lower bound of capacity-delay tradeoff in cluster critical regime \( (v + 2\beta = 1) \) can be obtained from the analysis of either cluster sparse regime or cluster dense regime. We will not repeat the deducing process of the cluster critical regime for simplification.

We have studied the capacity-delay tradeoff of correlated mobility for all sub-cases: the cluster sparse regime, the cluster dense regime, and the cluster critical regime. The impact of node correlation to the per-node capacity and packet delay will be discussed below.

In cluster sparse regime, the mobility of nodes performs strong correlation (i.e., the number of clusters is small or nodes in each cluster move within a small region). Thus, clusters in the network suffer a certain degree of disconnectedness, which restricts the maximum per-node capacity \( (mR^2/n) \) and minimum packet delay \( (n/(mR^2)) \). The disconnectedness due to the strong correlation of node mobility greatly degrade the performance of the network, which we should try to avoid when take into account the real application deployment. Fig. 6 demonstrates the capacity-delay tradeoff in cluster sparse regime when \( v = 5/12, \beta = 1/4 \). Although it generally performs worse than that of the i.i.d. mobility model, there still exists certain space we can explore to improve the tradeoff when high capacity becomes the major concern in real applications. Compare the two curves in Fig. 6, we can easily find that the tradeoff in cluster sparse regime is better than the optimal tradeoff in i.i.d. slow mobility model [6] when considering the high capacity region.

In cluster dense regime, the mobility of nodes show weak correlation and nodes of different clusters meet each other frequently. Obviously, network will no longer suffer the disconnectedness and thus performs better than that of the i.i.d. mobility model. Fig. 7 illustrates the capacity-delay tradeoff in cluster dense regime \( (v = 4/9, \beta = 1/3) \), which is much better than that of the i.i.d. slow mobility model.

Fig. 7 also exhibits the capacity-delay tradeoff in cluster critical regime, which is better than that in cluster dense regime. The reason is that when the node mobility show weak correlation (even extremely weak correlation), the number of clusters is large and each cluster covers a large area of the...
network, which results in the severe competition of the limited radio resources and a longer delay within each cluster. On the contrary, in cluster critical regime, the node mobility show medium correlation where the number of clusters or the area each cluster covers is neither too large nor too small. The medium correlation of node mobility benefits the network performance mainly in three aspects: 1) The connectivity is guaranteed (but not strongly connected); 2) Clusters are loosely overlapped which relieves the competition for radio resources among them compared with the case of cluster dense regime; 3) each cluster covers a relatively small area which reduces the transmission delay within the cluster. Thus, the node mobility in cluster critical regime can achieve better performance of capacity-delay tradeoff than that in cluster sparse regime and cluster dense regime.

Fig. 8 shows the capacity-delay tradeoff in cluster critical regime with various system parameters. We can easily find that either the value of $v$ or $\beta$ is too large, the network performance will be degraded. From numerous numerical experiments (we do not show in this paper), we find that when $v = 1/4$ ($\beta = 3/8$), the network can achieve the best per-node capacity of $\Theta(n^{-1/4})$ on condition that the packet delay remains to be $\Theta(1)$. When $v = 1/2$ ($\beta = 1/4$), the network can achieve the best packet delay of $\Theta(\sqrt{n})$ on condition that the per-node capacity remains to be $\Theta(1)$. Thus in cluster critical regime, we believe that it is essential to balance the number of clusters (adjust parameter $v$) and the area that each cluster covers (adjust parameter $\beta$) on condition that $v + 2\beta = 1$. Because these two system metrics greatly affect the network performance, which are also important for the system design in practice.

VII. CONCLUSION

This paper mainly focus on the impact of correlation of node mobility on the capacity-delay tradeoff in mobile ad hoc networks (MANETs). We have investigated the characteristics of correlated mobility and figured out the fundamental relationship between the capacity, delay and the associated system parameters, which afterwards provides great help to derive the capacity-delay tradeoff. Results demonstrate a whole picture of how the correlated mobility affect the network performance in different degrees of correlation. We reveal that all the three kinds of different degrees of correlated mobility can enhance the performance of capacity-delay tradeoff to some extents. In particular, the medium correlation of node mobility can better benefit the performance of capacity-delay tradeoff compared with the strong or weak correlation of node mobility. Because it can effectively control the number of clusters and area that each cluster covers, which further relieves the competition of limited radio resources and decrease the delay of the so called “last mile” transmission.

VIII. ACKNOWLEDGEMENT

This work was supported by NSF China (No.61532012, 61325012, 61271219, 61521062, 61428205)

REFERENCES


**Riheng Jia** received his B. E. degree in Electronics and Information Engineering from Huazhong University of Science and Technology, China, in 2012. He is currently pursuing the PhD degree in Electronic Engineering in Shanghai Jiao Tong University. His research interests are in the area of scaling laws analysis in wireless networks.

**Feng Yang** received his Ph.D degree in Information and Communication from Shanghai Jiao Tong University. Since 2008, he has been on the faculty at Shanghai Jiao Tong University, where he is an associate professor of electronic information. His research interest include wireless video communication, multi-hop communication. By now he is the PI of some national projects, including the National High Technology Research and Development Program of China (863 Program), National Natural Science Foundation of China.

**Shuochao Yao** received his B. E. degree in Electronic Engineering from Shanghai Jiao Tong University, China, in 2014. His research interests include the area of asymptotic analysis of capacity in wireless ad hoc networks.

**Xinhua Tian** received his B.E. and M.E. degrees in communication engineering from Northwestern Polytechnical University, Xian, China, in 2003 and 2006, respectively. He received the Ph.D. degree in the Department of Electrical and Computer Engineering (ECE), Illinois Institute of Technology (IIT), Chicago, in Dec. 2010. He is currently an Assistant Professor in Department of Electronic Engineering of Shanghai Jiao Tong University, China. He won the Highest Standards of Academic Achievement 2011 of IIT and Fieldhouse Research Fellowship 2009 of IIT, which is awarded to only one student over IIT each year. His research interests include application-oriented networking, Internet of Things and wireless networks. He serves as the guest editor of International Journal of Sensor Networks, publicity co-chair of WASA 2012. He also serves as the Technical Program Committee member for Wireless Networking Symposium, Ad Hoc and Sensor Networks Symposium of IEEE GLOBECOM 2013, Ad Hoc and Sensor Networks Symposium of IEEE ICC 2013, Wireless Networking Symposium of IEEE GLOBECOM 2012, Communications QoS, Reliability, and Modeling Symposium (CQRM) of GLOBECOM 2011, and WASA 2011.

**Xinbing Wang** (SM’12) received the B.S. degree (with hons.) in automation from Shanghai Jiao Tong University, Shanghai, China, in 1998, the M.S. degree in computer science and technology from Tsinghua University, Beijing, China, in 2001, and the Ph.D. degree with a major in electrical and computer engineering and minor in mathematics from North Carolina State University, Raleigh, in 2006. Currently, he is a professor with the Department of Electronic Engineering, Shanghai Jiao Tong University. His research interests include resource allocation and management in mobile and wireless networks, TCP asymptotics analysis, wireless capacity, cross-layer call admission control, asymptotics analysis of hybrid systems, and congestion control over wireless ad hoc and sensor networks. Dr. Wang has been a member of the Technical Program Committees of several conferences including ACM MobiCom 2012, ACM MobiHoc 2012, IEEE INFOCOM 2009-2013.

**Wenjun Zhang** received the B.S., M.S., and Ph.D. degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 1984, 1987, and 1989, respectively. In 1998, as a group leader, he was successful in charge of developing the first Chinese HDTV prototype system. He is a Changjiang Scholarship Professor in the field of communications and electronic systems with Shanghai Jiao Tong University, Shanghai, China. His current research interests include digital media processing and transmission, video coding, and wireless broadband communication systems.

**Jun/Jim Xu** received the PhD degree in computer and information science from The Ohio State University in 2000. He is a professor in the College of Computing at Georgia Institute of Technology. His current research interests include data streaming algorithms for the measurement and monitoring of computer networks and hardware algorithms and data structures for high-speed routers. He received the US National Science Foundation (NSF) CAREER Award in 2003, ACM Sigmetrics Best Student Paper Award in 2004, and IBM Faculty Awards in 2006 and 2008. He was named an ACM Distinguished Scientist in 2010. He is a member of the IEEE.