

Dealing with Link Blockage in mmWave Networks: D2D Relaying or Multi-beam Reflection?

Mingjie Feng, Shiwen Mao

Dept. Electrical & Computer Engineering
Auburn University, Auburn, AL 36849-5201, U.S.A.

Tao Jiang

School of Electronic Information & Communications
Huazhong Univ. Science & Technology, Wuhan 430074 China

Abstract—Device to device (D2D) relaying and multi-beam reflection are two effective approaches to deal with the blockage problem in millimeter-wave (mmWave) communication, each with its own limitations when serving a large number of user equipments (UE). A combination of D2D relaying and multi-beam reflection is expected to enhance the performance, but the selection of UEs to be served by each approach remains a challenge. In this paper, we consider adaptive mode selection between D2D relaying and multi-beam reflection in a time division duplex (TDD) mmWave network. We formulate a joint mode selection and resource sharing problem with the objective of maximizing the sum logarithm rate, and propose a two-stage solution algorithm. In the first stage, we derive the optimal resource sharing solution under the case that all UEs are served by D2D relaying. In the second stage, an adaptive algorithm is proposed to determine the set of UEs that switch from D2D relaying to multi-beam reflection. Simulation results demonstrate that the proposed scheme achieves considerable performance gain compared to several benchmark schemes.

I. INTRODUCTION

With the growing popularity of data-intensive applications, the next generation (i.e., 5G) wireless network is expected to provide 1000x data rate [1]. Millimeter-wave (mmWave) communication is one of the key enabling technologies of 5G wireless to meet such challenges [2], [3], along with massive MIMO [4] and small cells [5], [6]. Operating in the higher end of spectrum ranging from 30GHz to 300GHz, a large bandwidth is available (e.g., a 7 GHz license-free spectrum between 57 GHz and 64 GHz was approved by FCC), resulting in significantly improved data rate.

Despite such great potential, a major challenge of mmWave communications is to overcome blockages. Due to the short wave length, the mmWave cannot penetrate obstacles such as walls and human bodies. Thus, the line-of-sight (LOS) path between a base station (BS) and a user equipment (UE) may be easily blocked due to the mobility of the UE or other UEs. To maintain connectivity, alternative links have to be found and used. To this end, there are three possible approaches, including (i) device to device (D2D) relaying by other UEs, (ii) reflection of beams [8], [9], (iii) handover to other BSs. The D2D enabled mmWave network was considered in [10], [11], in which effective MAC protocols were proposed. A multi-beam reflection architecture was recently proposed in [12], where the concept of beamspace MIMO was introduced and a multiplexing gain was achieved. In [13], an inter-BS coordina-

tion mechanism is designed to deal with the NLOS challenge by optimizing the set of BSs serving each UE.

Although these solutions are effective in dealing with blockage, the performance gain is limited by some inherent factors of these approaches. For D2D relaying, a relaying UE must share part of its resource with other UEs. Besides, multi-hop relaying increases the delay since the packets have to be forwarded multiple times. The major challenge of multi-beam reflection is the large path loss exponent of NLOS links. Compared to LOS links with a typical path loss exponent of 2, the path loss exponent of NLOS can be as large as 4 [14]. Hence, only a certain set of UEs are suitable for multi-beam reflection. In addition, in case of serving a large number of UEs, the transmission power allocated to each UE would be limited, resulting in degraded quality of service (QoS). The handover approach, which seems easy to implement in traditional cellular network, requires intensive coordination among BSs in an mmWave network. Due to the narrow beam of mmWave transmissions, discovering an alternative BS and tracking roaming UEs may incur additional overhead, such as that caused by beam sweeping [15]. To avoid these drawbacks and fully harness the benefits of these approaches, a combination of multiple approaches with a proper integral design has the potential to enhance the system performance. As the UEs are served by multiple approaches, the number of UEs served by each approach is reduced, the resource allocated to each UE can be increased, resulting in improved performance. Then, how to select the sets of UEs served by different approaches is a key factor that impacts the system performance, but has not been investigated in previous works.

In this paper, we consider a combination of D2D relaying and multi-beam reflection with an adaptive selection between the two approaches in a single cell time division duplex (TDD) mmWave network. We formulate a joint mode selection and resource sharing problem with the objective of maximizing the sum logarithm rate, and propose a two-stage solution algorithm. In the first stage, we consider the case that all NLOS UEs are served by D2D relaying, and derive the optimal resource sharing solution. Based on the solution, an adaptive mode selection algorithm is then proposed in the second stage to determine the sets of NLOS UEs served by D2D relaying and multi-beam reflection. The proposed scheme is evaluated with simulations and compared with several benchmark schemes, where considerable performance gains are achieved.

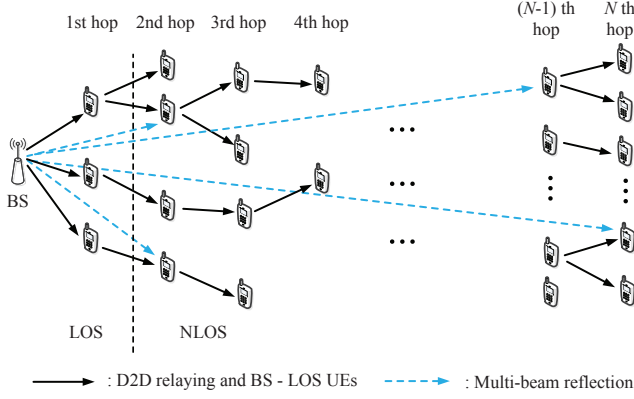


Fig. 1. System model of a D2D and multi-beam enabled multi-hop mmWave cellular network.

In the remainder of this paper, we present the system model and problem formulation in Section II. The solution algorithm is given in Section III. The simulation results are discussed in Section IV. We conclude this paper in Section V.

II. PROBLEM FORMULATION

We consider an mmWave cellular network that supports directional transmissions with a small beamwidth, e.g., less than 10° . Due to the “pseudo-wired” property and large propagation loss, the performance gain achieved from inter-BS interference coordination is limited. Thus, we focus on link scheduling of *one BS* and the UEs it serves. Assume the set of UEs served by each BS is predetermined, we focus on a tagged BS serving K UEs indexed by $k = 1, 2, \dots, K$. The UEs are subject to random blockage and can relay data for other UEs in case of blockage. The blocked UEs are served by either multi-hop D2D relaying or multi-beam reflection.

We assume the system operate in the TDD mode, so that channel reciprocity can be exploited for efficient channel estimation. As shown in Fig. 1, we consider link scheduling for downlink transmissions. Each NLOS UE receives its packets from a single relaying UE. Then, the network architecture of multi-hop D2D relaying can be viewed as a tree, while multi-beam reflection can be viewed as an alternative (or, augment) connection from the root to a node that corresponds to an NLOS UE.

Define x_k and y_k as indicators for multi-hop D2D relaying and multi-beam reflection for user k , respectively, given as

$$x_k \doteq \begin{cases} 1, & \text{user } k \text{ is served by D2D relaying or LOS link} \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

$$y_k \doteq \begin{cases} 1, & \text{user } k \text{ is served by multi-beam reflection} \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

We assign $x_k = 1$ to LOS UEs since they are part of the D2D relaying architecture. We then use a *descendent matrix* to indicate the relaying route of each NLOS UEs, defined as

$$a_{k,k'} \doteq \begin{cases} 1, & \text{user } k' \text{ is a descendent of user } k \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

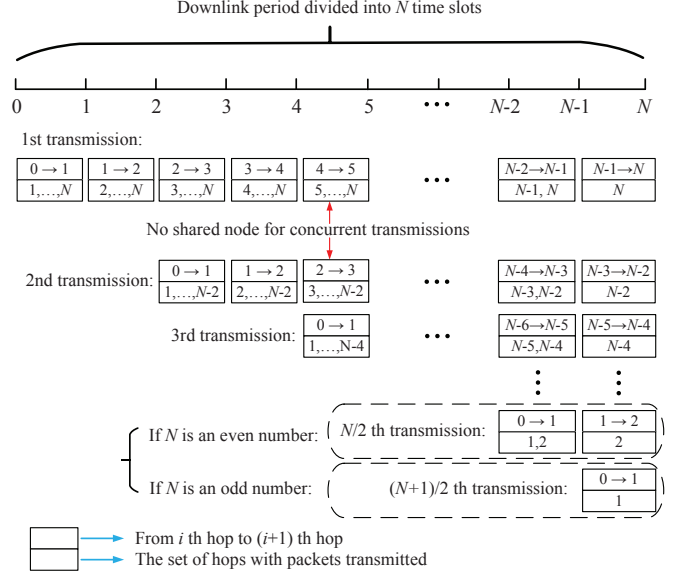


Fig. 2. Transmission pattern of a TDD-based multi-hop D2D mmWave cellular network.

Since $a_{k,k'}$ is defined for UEs involved in the D2D relaying process, we have $a_{k,k'} \leq x_k$ and $a_{k,k'} \leq x_{k'}$. Let ρ_k be the depth of UE k in the D2D relaying tree, given by

$$\rho_k = \sum_{k' \neq k} a_{k',k} + 1. \quad (4)$$

That is, ρ_k is also the number of hops required to transmit the packets of UE k . Let N be the maximum value of ρ_k . To accommodate N -hop D2D relaying and guarantee that the packets of UEs at the N th hop can be delivered at the end of each downlink transmission, we divide the downlink transmission period into N time slots as shown in Fig. 2. Each time slot is used for transmission from UEs in one hop to UEs in the next hop. For a UE served with multi-beam reflection, the entire downlink period is allocated to the UE.

We assume that all UEs can only perform half duplex relaying; two adjacent links that share the same UE cannot operate concurrently [16]. Under this constraint, a total number of $\lceil \frac{N}{2} \rceil$ transmissions can be implemented, as shown in Fig. 2. For example, after the 2nd stage of the 1st transmission, which is from UEs with $\rho_k = 1$ to UEs with $\rho_k = 2$, the 2nd transmission can be initiated since UEs with $\rho_k = 1$ have finished transmission and are free for reception. Note that, as the index of transmissions increases, the packets of UEs with larger ρ_k are not contained since there is not enough time for the transmissions of these UEs. In particular, the packets of UEs with $\rho_k = N - 1$ and $\rho_k = N$ are not contained in the 2nd transmission, the packets of UEs with $\rho_k = N - 3, \dots, N$ are not contained in the 3rd transmission, and so on. Then, the total number of transmissions for UE k , θ_k , is given by

$$\theta_k = \begin{cases} \frac{N}{2} - \lceil \frac{\rho_k}{2} \rceil + 1, & N \text{ is even} \\ \frac{N+1}{2} - \lceil \frac{\rho_k}{2} \rceil, & N \text{ is odd,} \end{cases} \quad (5)$$

The BS is able to serve multiple users on the same time-frequency resource block with advanced techniques, such as

hybrid beamforming [7]. Due to hardware constraints, when a UE relays the packets of multiple UEs, resource sharing is required among these UEs. Without loss of generality, we assume time division multiple access (TDMA) is applied when a UE forwards the packets to multiple UEs. Specifically, a fraction of time in each time slot is allocated to each D2D link between the relaying UE and a UE at the next hop. We define $t_k^{i,j}$ as the fraction of time allocated to UE k on its i th hop during the j th transmission. From the perspective of outflow, $t_k^{i,j}$ should satisfy

$$\sum_{k' \neq k} a_{k,k'} t_{k'}^{\rho_k+1,j} \leq 1, k \in \{k | \rho_k \leq N-1\}, j=1, \dots, \theta_k + \left\lceil \frac{\rho_k}{2} \right\rceil. \quad (6)$$

For UEs with $\rho_k = 1$, i.e., the LOS UEs, the inflow constraint is given as

$$\sum_{k' \neq k} a_{k,k'} t_{k'}^{1,j} + t_k^{1,j} \leq 1, k \in \{k | \rho_k = 1\}, j=1, \dots, \left\lceil \frac{N}{2} \right\rceil. \quad (7)$$

The inflow constraint (7) is only required for the 1st hop transmission, since the inflow of other hops would only be part of a time slot due to TDMA at the previous hop; the left-hand side of (7) would always be no larger than 1. In a multi-hop transmission, the data rate of a link at the current hop should be no less than the data rate of the link at the next hop. Thus, $\{t_k^{i,j}\}$ should also satisfy

$$C_k^{i,j} \geq C_k^{i+1} t_k^{i+1,j}, k \in \{k | \rho_k \geq 2\}, i=1, \dots, \rho_k-1, \quad (8)$$

where C_k^i is the link capacity of hop i of UE k , given by

$$C_k^i = B \log(1 + \gamma_k^i), \quad (9)$$

where B is the system bandwidth and γ_k^i is the SINR of hop i of UE k . Here, the SINR is constant over all θ_k transmissions, since the duration of uplink and downlink periods is less than the coherence interval in a TDD system.

Next, we derive the SINR expressions under different transmission schemes, employing the baseline SINR model presented in [14]. For UEs served by D2D relaying, we have

$$\gamma_k^i = \frac{p_k^i h_k^i G_k^i (d_k^i)^{-2}}{\sigma^2}, \quad (10)$$

where h_k^i is the small scale fading, which is a normalized Gamma random variable [14]; G_k^i is antenna array gain; d_k^i is the distance of the link; and p_k^i is the transmission power of the BS or UE for the i th hop of UE k . The noise power is σ^2 , and we neglect the impact of interference. In (10), we assume that the LOS path loss exponent is 2.

For UEs served by BS with multi-beam reflection and UEs served by BS with an LOS link, we assume the power of BS is equally allocated to these UEs. Then, the transmission power allocated to each UE is

$$\tilde{p}_k = \frac{P_0}{\sum_k y_k + N_{\text{LOS}}}, \quad (11)$$

where P_0 is the BS power and N_{LOS} is the number of LOS

UEs. Similar to (10), the SINR of an LOS UE is given by

$$\gamma_k^1 = \frac{\tilde{p}_k h_k^1 G_k^1 (d_k^1)^{-2}}{\sigma^2}. \quad (12)$$

Suppose UE k is served with the reflection of M_k beams indexed by $m = 1, \dots, M_k$. Let $\tilde{\gamma}_k^m$ be the SINR of the path corresponding to the m th beam of UE k , given as

$$\tilde{\gamma}_k^m = \frac{\tilde{p}_k^m h_k^m G_k^m (d_k^m)^{-4}}{\sigma^2}, \quad (13)$$

where \tilde{p}_k^m , h_k^m , G_k^m , and d_k^m are the power, small-scale fading, antenna array gain, and distance of the m th path, respectively. We neglect the interference caused by side lobes. Without loss of generality, we assume \tilde{p}_k is equally allocated to the M_k beams. The path loss factor for NLOS reflection is 4.

For UE k served by multi-hop D2D relaying, its downlink data rate is the sum of all θ_k transmissions. In each multi-hop transmission, the effective data rate is determined by data rate of the *final hop* divided by the number of hops. Then, the data rate of UE k when served by multi-hop D2D relaying in the downlink period is given by

$$R_k = \sum_{j=1}^{\theta_k} \frac{C_k^{\rho_k} t_k^{\rho_k,j}}{\rho_k}. \quad (14)$$

The data rate of LOS UEs is a special case of (14) with $\rho_k = 1$.

When UE k is served by multi-beam reflection in the downlink period, its data rate is given by

$$\tilde{R}_k = \sum_{m=1}^{M_k} B \log(1 + \tilde{\gamma}_k^m). \quad (15)$$

In this paper, we aim to maximize the logarithm rate sum of all UEs with adaptive selection between D2D relaying and multi-beam reflection for all NLOS UEs. Let \mathbf{x} , \mathbf{y} , \mathbf{a} , and \mathbf{t} be the vector/matrix forms of $\{x_k\}$, $\{y_k\}$, $\{a_{k,k'}\}$, and $\{t_k^{i,j}\}$. The problem is formulated as

$$\mathbf{P1:} \max_{\{\mathbf{x}, \mathbf{y}, \mathbf{a}, \mathbf{t}\}} \left\{ \sum_{k=1}^K x_k \log R_k + \sum_{k=1}^K y_k \log \tilde{R}_k \right\} \quad (16)$$

subject to:

$$x_k + y_k \leq 1, \forall k \quad (17)$$

$$\sum_{k' \neq k} a_{k,k'} t_{k'}^{\rho_k+1,j} \leq 1, k \in \{k | \rho_k \leq N-1\}, j=1, \dots, \theta_k + \left\lceil \frac{\rho_k}{2} \right\rceil \quad (18)$$

$$\sum_{k' \neq k} a_{k,k'} t_{k'}^{1,j} + t_k^{1,j} \leq 1, k \in \{k | \rho_k = 1\}, j=1, \dots, \left\lceil \frac{N}{2} \right\rceil \quad (19)$$

$$C_k^{i-1} t_k^{i-1,j} \geq C_k^i t_k^{i,j}, k \in \{k | \rho_k \geq 2\}, i=1, \dots, \rho_k \quad (20)$$

$$\sum_{k' \neq k} a_{k',k} \leq 1, k \in \{k | \rho_k \leq 2\}, \quad (21)$$

$$a_{k,k'} \leq x_k, \forall k. \quad (22)$$

$$t_k^{i,j} \leq x_k, 0 \leq t_k^{i,j} \leq 1, \forall k, i, j. \quad (23)$$

$$x_k, y_k, a_{k,k'} \in \{0, 1\}, \forall k. \quad (24)$$

Constraint (20) is due to the fact that each node in the D2D relaying tree can only have one parent node.

III. SOLUTION ALGORITHM

Problem **P1** is a mixed integer programming problem with multiple sets of variables, which cannot be solved with standard techniques. To derive an effective solution, we propose a two-stage algorithm for adaptive selection between D2D relaying and multi-beam reflection. In the first stage, we consider the case that all NLOS UEs are served with D2D relaying and derive the optimal resource allocation solution. Based on the solution, the UEs that can be served by multi-beam reflection are evaluated in the second stage, then the set of UEs to be served by multi-beam is determined.

A. First Stage

We assume that path selection for D2D relaying is predetermined with a routing approach, e.g., adopting Dijkstra's Algorithm by setting the weight of each link as the inverse of its channel gain [17]. With all NLOS UEs served by D2D relaying, problem **P1** is reduced to the following resource allocation problem.

$$\mathbf{P2}: \max_{\{\mathbf{t}\}} \sum_{k=1}^K \log \left(\frac{\sum_{j=1}^{\theta_k} C_k^{\rho_k} t_k^{\rho_k, j}}{\rho_k} \right) \quad (25)$$

subject to: (18) – (20)

It can be verified that problem **P2** is a convex optimization problem and strong duality holds. The Lagrangian dual method can be applied to obtain the optimal solution.

B. Second Stage

Based on the optimal solution of problem **P2**, we then evaluate the performance gain of switching a UE from D2D to multi-beam, and determine the set of UEs to be served with multi-beam reflection. We define $\Delta_k^{[\tau]}$ as the performance gain of all UEs by selecting UE k to switch from D2D relaying to multi-beam reflection at the τ th round. For a UE with descendent UE(s) in the D2D tree, we assume that all of its descendent UEs would also switch to multi-beam reflection since re-constructing the D2D tree is time-consuming, which should not be carried out frequently. Then, the performance gain is given as: $\Delta_k^{[\tau]} = \log \tilde{R}_k^{[\tau]} - \log R_k^{[\tau]} + \sum_{k' \neq k} a_{k, k'} (\log \tilde{R}_{k'}^{[\tau]} - \log R_{k'}^{[\tau]})$.

Since the objective function is to maximize the sum logarithm rate, the value of $t_k^{\rho_k, j}$ would be close to each other among different UEs. Thus, when an NLOS UE l switches from D2D relaying to multi-beam reflection, we approximate the values of $t_k^{\rho_k, j}$ of other UEs to be increased by $\frac{t_l^{\rho_l, j}}{\sum_k x_k}$ or $\frac{t_l^{\rho_l, j} + \sum_{l' \neq l} a_{l, l'} t_{l'}^{\rho_{l'}, j}}{\sum_k x_k}$, depending on whether UE l has any descendent in the D2D tree. Since we assume equal power allocation, the transmission power allocated to each LOS UE and multi-beam NLOS UE is decreased by a factor of $\frac{\sum_k y_k + N_{\text{LOS}}}{\sum_k y_k + 1 + N_{\text{LOS}}}$ or $\frac{\sum_k y_k + N_{\text{LOS}}}{\sum_k y_k + 1 + \sum_{l' \neq l} a_{l, l'} + N_{\text{LOS}}}$, depending on whether UE l has any descendent in the D2D tree.

Let $\eta_k = \{k' | a_{k, k'} = 1\}$ and Φ be the set of UEs served by D2D relaying. The adaptive algorithm for selection between

Algorithm 1: Adaptive Mode Selection Algorithm

```

1 Initialize:  $x_k = 1, y_k = 0, \forall k, \Phi = \{1, \dots, K\}$ ;
2  $\tau = 1, l^{[1]} = \arg \max_{\{k | x_k = 1\}} \Delta_k^{[1]}$ ;
3 while  $\Delta_l^{[\tau]} > 0$  do
4   if  $\rho_l = 1$  then
5     Set  $y_l = 1, x_l = 0, \Phi = \Phi - \{l\}$ ;
6     for  $k = 1 : K$  do
7       Update  $R_{k'}^{[\tau]}$  with (14) by adding  $t_l^{\rho_l, j} / \sum_k x_k$  to
          each  $t_k^{\rho_k, j}, k \in \Phi$ ;
8       Update  $\tilde{R}_k^{[\tau]}$  with (15) by multiplying  $\tilde{p}_k$  by
           $(\sum_k y_k + N_{\text{LOS}}) / (\sum_k y_k + 1 + N_{\text{LOS}})$ ;
9       Calculate  $\Delta_k^{[\tau]}$ ;
10    end
11  else
12    Set  $y_l = 1, x_l = 0; y_{l'} = 1, x_{l'} = 0, \text{ for } l' \in \eta_l$ ;
13     $\Phi = \Phi - \{l\} - \eta_l$ ;
14    for  $k = 1 : K$  do
15      Update  $R_{k'}^{[\tau]}$  with (14) by adding
           $(t_l^{\rho_l, j} + \sum_{l' \neq l} a_{l, l'} t_{l'}^{\rho_{l'}, j}) / \sum_k x_k$  to each  $t_k^{\rho_k, j},$ 
           $k \in \Phi$ ;
16      Update  $\tilde{R}_k^{[\tau]}$  with (15) by multiplying
           $(\sum_k y_k + N_{\text{LOS}}) / (\sum_k y_k + 1 + \sum_{l' \neq l} a_{l, l'} + N_{\text{LOS}})$ 
          to each  $\tilde{p}_k^m, k \notin \Phi$ ;
17      Calculate  $\Delta_k^{[\tau]}$ ;
18    end
19  end
20   $l^{[\tau]} = \arg \max_{\{k \in \Phi\}} \Delta_k^{[\tau]}$ ;
21   $\tau = \tau + 1$ ;
22 end
```

D2D and multi-beam is summarized in Algorithm 1. In each round of Algorithm 1, the UE or the set of UEs that brings the largest performance gain is selected to switch from D2D relaying to multi-beam reflection. Such process terminates until no positive gain can be achieved.

IV. SIMULATION STUDY

We validate the performance of the proposed scheme through MATLAB simulations. We consider one BS serving multiple UEs with coverage range of 50 m. The system bandwidth is 1 GHz. The UEs are randomly distributed in the coverage area. Each UE is subject to random blockage with probability P_k^b . We assume that the blockage probability is proportional to the distance between the UE and the BS with coefficient κ , given as $P_k^b = \min\{\kappa \cdot D_k, 1\}$. As an example, when $\kappa = 0.04$, a UE that is 10 m away from the BS has a probability of 0.4 to be blocked; for UEs that are more than 25 m away, they are always blocked. We compare with a heuristic scheme. In each round of the heuristic scheme, the UE with the largest value of \tilde{R}_k is selected to be served with multi-beam reflection. Such process terminates until $\sum_k y_k \tilde{R}_k$ decreases. We also consider the case of setting the objective function of Problem **P2** as sum rate maximization, which serves as an *upper bound* for the sum rate performance.

Fig. 3 shows the sum rate performance under different number of UEs. The sum rate of the multi-beam only scheme is lower than other schemes due to the large path loss exponent of NLOS links and the increased transmission distance brought by reflection. The D2D only scheme achieves higher sum rate than the multi-beam only scheme, but its performance is

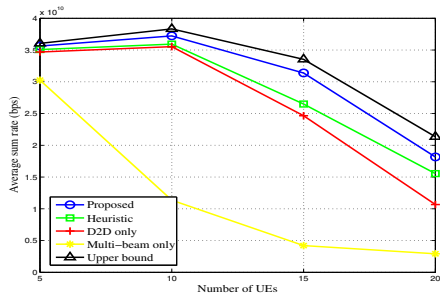


Fig. 3. Average sum rate under different numbers of UEs. $\kappa = 0.02$.

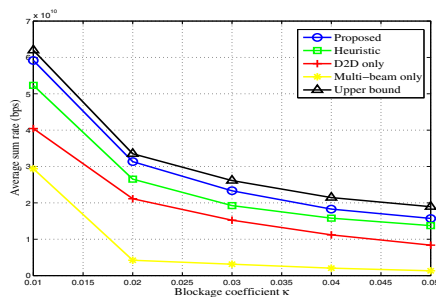


Fig. 4. Average sum rate under different values of κ . The number of UEs is 15.

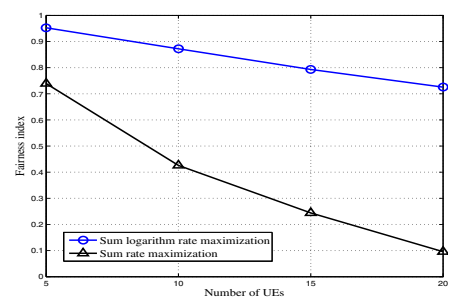


Fig. 5. Fairness with different objective functions. The number of UEs is 15 and $\kappa = 0.02$.

significantly degraded when the number of UEs is large. The degraded performance of D2D only scheme results from both resource sharing among UEs and the increased time spent on multi-hop transmission. By selecting some UEs to be served by multi-beam reflection, the heuristic scheme and the proposed scheme achieve better performance, due to the reduced number of UEs involved in D2D transmission. The proposed scheme outperforms the heuristic scheme since the proposed adaptive algorithm jointly considers the performances of UEs served by D2D and multi-beam, while the heuristic scheme is only based on UEs served by multi-beam. The performance of the proposed scheme is close to its upper bound, showing that the performance loss due to fairness concern is relatively small. The results of Fig. 3 indicate that both D2D relaying and multi-beam reflection are highly limited by the increasing traffic. A combination of the two approaches with proper UE selection can effectively improve the system performance.

The performances under different blockage coefficients, κ , is shown in Fig. 4. As κ increases, the performances of all schemes are degraded due to increased ratio of NLOS UEs. It can be seen that the performance of both D2D only and multi-beam only schemes are highly sensitive to blockage, since the time and power resources are shared by the increasing number of NLOS UEs. With adaptive selection between D2D and multi-beam, the proposed scheme achieves considerable performance gain, especially when κ is large.

Fig. 5 shows the fairness performance by setting different objective functions for Problem **P2**. We use the Jain's fairness index, $(\sum_k x_k R_k + \sum_k y_k \tilde{R}_k)^2 / K \cdot \sum_k (x_k R_k + y_k \tilde{R}_k)^2$, to measure fairness between different UEs. It can be seen from Fig. 5 that when the objective is set to sum rate maximization, the fairness is poor, since the links with high channel gain would be allocated with much more transmission time than the links with low channel gain. With the objective of sum logarithm rate maximization, the proposed scheme achieves a good tradeoff between system performance and fairness.

V. CONCLUSIONS

We considered a combination of D2D relaying and multi-beam reflection to overcome blockage as well as enhance performance of a TDD mmWave network. A two-stage solution algorithm was proposed to determine the set of UEs served by D2D relaying and multi-beam reflection. The effectiveness

of the proposed scheme was validated with simulations.

ACKNOWLEDGMENT

This work is supported in part by the NSF under Grant CNS-1320664, and by the Wireless Engineering Research and Education Center (WEREC) at Auburn University.

REFERENCES

- [1] Qualcomm, "The 1000x data challenge," [online] Available: <https://www.qualcomm.com/1000x>.
- [2] J.G. Andrews, S. Buzzi, W. Choi, S.V. Hanly, A. Lozano, A.C.K. Soong, and J.C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol.32, no.6, pp.1065–1082, June 2014.
- [3] T. S. Rappaport, et al., "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access J.*, vol. 1, May 2013, pp. 335–449
- [4] M. Feng and S. Mao, "Harvest the potential of massive MIMO with multi-layer techniques," *IEEE Network*, vol.30, no.5, pp.40–45, Sept./Oct. 2016.
- [5] M. Feng, T. Jiang, D. Chen, and S. Mao, "Cooperative small cell networks: High capacity for hotspots with interference mitigation," *IEEE Wireless Commun.*, vol.21, no.6, pp.108–116, Dec. 2014.
- [6] M. Feng, S. Mao, and T. Jiang, "Joint duplex mode selection, channel allocation, and power control for full-duplex cognitive femtocell networks," *Elsevier Digital Communications and Networks Journal*, vol.1, no.1, pp.30–44, Feb. 2015.
- [7] A. Alkhateeb, G. Leus, and R.W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol.14, no.11, pp.6481–6494, Nov. 2015.
- [8] Z. He and S. Mao, "A decomposition principle for link and relay selection in dual-hop 60 GHz networks," in *Proc. IEEE INFOCOM'16*, San Francisco, CA, Apr. 2016, pp.1683–1691.
- [9] Z. He, S. Mao, S. Kompella, and A. Swami, "On link scheduling in dual-hop 60 GHz mmWave networks," *IEEE Trans. Vehicular Technology*, to appear. DOI: 10.1109/TVT.2017.2717840.
- [10] J. Qiao, X. Shen, J.W. Mark, Q. Shen, Y. He, and L. Lei, "Enabling device-to-device communications in millimeter-wave 5G cellular networks," *IEEE Commun.*, vol.53, no.1, pp.209–215, Jan. 2015.
- [11] Y. Niu, C. Gao, Y. Li, L. Su, D. Jin, and A.V. Vasilakos, "Exploiting device-to-device communications in joint scheduling of access and backhaul for mmWave small cells," *IEEE J. Sel. Areas Commun.*, vol.33, no.10, pp.2052–2069, Oct. 2015.
- [12] Q. Xue, X. Fang, and C.-X. Wang, "Beamspace SU-MIMO for future millimeter wave wireless communications," *IEEE J. Sel. Areas Commun.*, vol.35, no.7, pp.1564–1575, July 2017.
- [13] S.-C. Lin and I.F. Akyildiz, "Dynamic base station formation for solving NLOS problem in 5G millimeter-wave communication," in *Proc. IEEE INFOCOM'17*, Atlanta, GA, May 2017.
- [14] J.G. Andrews, et al., "Modeling and analyzing millimeter wave cellular systems," *IEEE Trans. Commun.*, vol.65, no.1, pp.6481–6494, Jan. 2017.
- [15] Y. Wang, S. Mao, and T.S. Rappaport, "On directional neighbor discovery in mmWave networks," in *Proc. IEEE ICDCS'17*, Atlanta, GA, June 2017, pp.1704–1713.
- [16] J. Qiao, et al., "Enabling multi-hop concurrent transmissions in 60 GHz wireless personal area networks," *IEEE Trans. Wireless Commun.*, vol.10, no.11, pp.3824–3833, Nov. 2011.
- [17] Z. He, S. Mao, and T.S. Rappaport, "On link scheduling under blockage and interference in 60 GHz ad hoc networks," *IEEE Access J.*, vol.3, pp.1437–1449, Sept. 2015.