Received January 18, 2017, accepted February 12, 2017, date of publication February 23, 2017, date of current version May 17, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2672996

Cooperative Content Delivery in Multicast Multihop Device-to-Device Networks

ZICHENG XIA, JIAWEI YAN, AND YUAN LIU, (Member, IEEE)
School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China

Corresponding author: Yuan Liu (eeyliu@scut.edu.cn)

This work was supported in part by the Natural Science Foundation of China under Grant 61401159 and in part by the Pearl River S&T Nova Program of Guangzhou.

ABSTRACT The increasing demand of mobile devices (MDs) for data services brings tremendous pressure to cellular networks. It has become a great challenge for traditional offloading techniques to balance the energy efficiency and quality of service. The concept of device-to-device (D2D) communication shows a huge potential in cellular offloading. In this paper, we investigate the scenario where MDs have the same demand for a common content and they cooperate to download the content by multihop relaying. We aim to minimize the total power consumption by grouping MDs in multihop D2D networks, while satisfying the minimum rate required by each MD. As the problem is NP-complete and the optimal solution cannot be found in polynomial time, we propose three greedy algorithms with different grouping strategies to trade off the performance and complexity. Simulation results demonstrate that the total power consumption can be saved significantly in the content delivery situation using cooperative D2D communication, and the proposed algorithms are suitable for static and dynamic networks with different advantages.

INDEX TERMS Content delivery, device-to-device (D2D) communication, multihop relaying, multicast.

I. INTRODUCTION

Nowadays, explosive growth in mobile applications brings great data traffic pressure to wireless communication systems. People acquire data service more frequently using their mobile devices (MDs) in various scenarios such as connecting with others, following social media, watching live shows, etc. Consequently, the rapidly increasing data demands are requiring more efficient cellular technologies to remain users' quality of service. Traditional cellular network technologies are no longer capable of meeting data service needs in the foreseeable future.

Highly attracting attention in recent years, offloading cellular networks have been developed by either migrating to new network topologies or enhancing techniques of current cellular networks to accommodate more subscribers with higher data rates [2]–[9]. A series of cellular offloading techniques have been studied in the literature including switching to small cells [10], [11], Wi-Fi networks [4], [12], and caching [13]–[15]. These existing techniques show various advantages such as low cost, standardized interface and high data rates. Among these directions, a promising basic technique is widely applied to improve the network efficiency, which is multicasting popular content to MDs on the cellular network [6], [16]. This technique shows its considerable offloading capacity as well as energy-saving advantage in high density of MDs requiring the same content. Nevertheless, a major issue of multicast is that the download rates of all MDs are limited by an identical rate decided by the worst channel among all MDs, which inevitably sacrifices the quality of service for MDs in better channel conditions. Another preferable technique is to download content cooperatively and MDs can act as relays and connect each other with no congestion [2], [7], [17]–[22].

The concept of device-to-device (D2D) communications enables MDs to connect each other directly and bypass the base station (BS). This property determines that cooperative D2D networks have great application potential in content delivery scenarios (e.g., files, videos, live shows, etc.) with their good performances [4], [17]–[21]. From the knowledge of recent studies, the D2D-enabled cellular can offload data traffic by following ways. One of the ideas is that the undelivered content is divided into different blocks and BS delivers them to different MDs in a group. Then MDs exchange the blocks via an ad hoc manner until the complete content is received by every MD. By doing so, the cellular network only needs to deliver a few copies of the content to MDs instead of the entire one [8]. Group-based multicast in multihop D2D networks is another direction, that is, MDs are divided into different groups and the content is delivered group by group using multicast via multihop relaying. However, this direction has been much less touched yet.

There are many advantages of using multicast in multihop networks. Firstly, the traffic pressure of BS is considerably offloaded as much more content can be delivered by MDs. Most importantly, cooperative relaying and group multicast conduce to power-saving in high rate data transmission significantly when compared to a traditional scheme. Moreover, MDs on the cell edge can obtain a much better channel from their close multicast groups for data transmission. However, the implementation of multicast multihop transmission is based on how to group MDs and establish the multicast links. Nevertheless, finding an optimal group solution usually falls into NP-complete problem due to the combinatorial nature. In addition, the network becomes dynamic in practice as MDs may join or exit from the D2D network at different time. Hence, content delivery in dynamic networks is also important but has not been considered yet.

Several works investigated multicast D2D networks. For example, in [23], Wi-Fi cooperation and D2D-based multicast content distribution are compared in terms of time-saving and power-saving with a pre-grouped cooperative network model. In [24], the authors considered grouping multicast in order to deal with delay and throughput problems. Similarly, the groups were assumed to be fixed. The authors in [8] involved fairness issue to user grouping for optimizing channel usage.

To our best knowledge, total power minimization with group multicast in cooperative multihop D2D networks has not been investigated in the literature. The goal of this paper is to minimize total power consumption by finding progressive schemes to form an efficient group-based multicast network, while maintaining rate requirements of all MDs. The multicast groups are connected via multihop relaying fashion. The problem cannot be solved optimally with polynomial complexity. Thus alternatively, we propose three heuristic algorithms to balance the performance and complexity. Each algorithm has its own advantages and disadvantages. In addition, the three algorithms can be generally extended to dynamic networks, where MDs randomly join or exit from the cooperative D2D network. The effectiveness of the proposed algorithms is verified by simulation results.

The remainder of this paper is organized as follows. Section II describes the system model and the optimization problem. Section III proposes three heuristic algorithms. Section IV extends the proposed algorithms in dynamic networks. Comprehensive numerical results and discussions are provided in Section V. Finally, Section VI concludes the paper.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a cellular network shown in Fig. 1, where MDs tend to obtain the same content from the single BS, denoted as set $\mathcal{C} = \{1, \ldots, C\}$. Note that no matter services and wireless environments, BS needs to be ready for transmission. Hence, the load of BS is heavy if mass MDs are active in requesting services. In addition, when a group of MDs request for a large content, they will consume much wireless resources. This kind of data demands is getting popular in modern communications such as file downloading and sharing. Therefore, it becomes a great challenge for any cellular communication system to maintain sufficient power output for the heavy data traffic in such applications. To tackle this issue, a way that MDs cooperate with each other using D2D communications to obtain the same content from BS is considered. That is, the content can be delivered to a MD either from BS or other MDs through multiple hops, which is illustrated by Fig. 1.

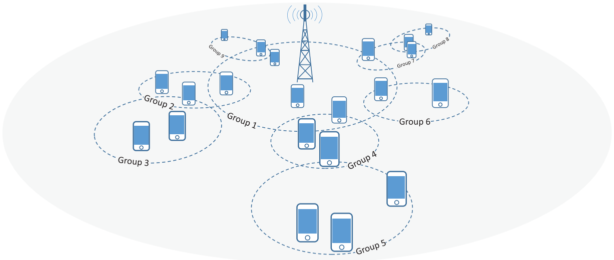


FIGURE 1. **Cooperative content delivery in multicast multihop D2D network.**

The considered content delivery process in D2D networks is carried out as follows: When a new content is available in BS and downloadable, BS chooses some appropriate MDs and multicasts the content to them. The required service rates of the selected MDs must be satisfied. Subsequently, the rest of MDs can obtain the same content from those MDs received content from BS. Some certain grouping rules help these MDs to establish a number of subsets for multicast and each subset regards the relay as its multicast transmitter. This procedure iterates until all MDs are grouped as well as their required rates are maintained. In another word, the content delivery process is executed group by group using multihop relaying.

A. PARAMETERS AND VARIABLES

We present the main parameters and variables in Table I.

In this paper, the wireless channel gain between MD m and MD n is determined in the following model:

$$\begin{aligned} h_{mn}(dB) &= \frac{P_{r,mn}(dB)}{P_{t,mn}(dB)} \\ &= \underbrace{10\log_{10} K - 10\beta \log_{10}\frac{d}{d_0}}_{pathloss} - \underbrace{\varphi_{dB}}_{shadowfading} \end{aligned} \quad (1)$$

where $P_{t,mn}$ and $P_{r,mn}$ denote the transmit and received power of MD m, respectively; K is a constant which is determined by the characteristics of antenna and the attenuation of average channels; d_0 is a reference distance (1-10 meters indoors and 10-100 meters outdoors) of the antenna far-field; β is the path loss exponent (2-4 in most cases); φ_{dB} represents the shadow fading which is Gauss-distributed random variable with mean

TABLE 1. Main Parameters and Variables

Parameters	
$\mathcal{C}$	The set of MDs $\mathcal{C} = \{1, 2, ..., C\}$
$P_{t,mn}$	Transmit power of transmitter m to n
$P_{r,mn}$	Receive power at MD n from transmitter m
β	The path loss exponent
N_0	The noise power density
$\mathcal{K}_{s_j}$	The multicast group consists of MDs receiving content from transmitter s_j
$R_g(\mathcal{K}_{s_j})$	The multicast rate of group $\mathcal{K}_{s_j}$
d_0	A reference distance of the antenna far-field
h_{mn}	The channel quality exponent between MD m and MD n
x_{mn}	A binary variable that indicates whether the condition of channel between MD m and MD n is the worst in the multicast group transmited by MD m
y_{mn}^h	A binary variable indicates whether MD m transmits content to MD n on hop h
R_{min}	The required rate of all MDs
H_{max}	The maximum number of transmit hops
$\mathcal{R}$	The set of receivers have not yet get content
$\mathcal{S}$	The set of potential transmitters containing content
$\mathcal{K}$	The set of disjoint multicast among multiple hops
N	The upper bound of the number of requesting MDs joining the networks on a time moment.
f	The probability, that the transmitters in set $\mathcal{S}$ without transmit task may exit from the D2D network.

zero and variance $\sigma^2_{\varphi dB}$; d is the distance from a transmitter to a receiver.

For transmitter m, the achievable rate of MD n is given by

$$R_n = \log_2\left(1 + \gamma_{mn} P_{t,mn}\right), \tag{2}$$

where $\gamma_{mn} = h_{mn}/N_0$ is the channel-to-noise ratio (CNR) and N_0 is the power of noise. The multicast rate R_n should follow the Short Slab theory, i.e., the multicast rate is decided by the worst channel. Assume that a multicast group $\mathcal{K}_{s_j}$ consists more than one MD and needs multicast to serve all MDs. The highest achievable multicast rate $R_g(\mathcal{K}_{s_j})$ is given by

$$R_g(\mathcal{K}_{s_j}) = \log_2\left(1 + \gamma_{mw'} P_{t,mw'}\right), \tag{3}$$

where w' is the MD whose CNR is the worst in multicast group $\mathcal{K}_{s_j}$, so there exists $R_g(\mathcal{K}_{s_j}) = \min R_n$, $n \in \mathcal{K}_{s_j}$. On the other hand, the required transmit power depends on the decodable rate, which means, for determined quality of service demand R_{min}, the required transmit power is

$$P_{t,mn} = (2^{R_{min}} - 1)/\gamma_{mn}. \tag{4}$$

B. OPTIMIZATION PROBLEM FORMULATION

In this paper, we aim to minimize the total power consumption of the entire network by grouping MDs with multiple hops, while maintaining the minimum decodable rate of each MD. The problem can be mathematically formulated as

$$\min_{P_t, x_{mn}, y_{mn}^h} \sum_{m=1}^{C} \sum_{n=1}^{C} P_{t,mn} x_{mn} + P_{(BS)} \tag{5}$$

$$s.t.\ x_{mn} \le y_{mn}^h, \quad \forall m,\ \forall n \in \mathcal{C}, 2 \le h \tag{6}$$

$$y_{mn}^h \le y_{(BS)m}^1, \quad \forall m,\ \forall n \in \mathcal{C}, 2 \le h \tag{7}$$

$$y_{ji}^h \le \sum_{k=1}^{C} y_{kj}^{h'}, \quad h' = 1, .., h-1 \tag{8}$$

$$R_{min} \le R_g(\mathcal{K}_{s_j}), \quad \forall \mathcal{K}_{s_j} \subseteq \mathcal{K} \tag{9}$$

$$\sum_{h=2}^{H_{max}} \sum_{m \ne n}^{C} y_{mn}^h + y_{(BS)n}^1 = 1, \quad \forall n \tag{10}$$

$$h \le H_{max}, \quad \forall h \tag{11}$$

where x_{mn} is a binary variable representing whether the channel condition between MDs m and n is the worst in the multicast group with MD m as the transmitter; y_{mn}^h is a binary variable representing whether MD m transmits content to MD n on hop h; R_{min} is a pre-defined minimum rate requirement of all MDs with respect to quality of service; $R_g(\mathcal{K}_{s_j})$ is the multicast rate achieved by multicast group $\mathcal{K}_{s_j}$; H_{max} is the pre-defined maximum tolerated hops to some extent delay considerations.

Specifically, objective (5) minimizes total transmit power of BS and all MDs. Constraint (6) guarantees that the transmission rate determined by the worst channel condition in each multicast group. Constraint (7) ensures that MD n can transmit content on the next hop only if it receives the same content from BS. Constraint (8) ensures that MD n can transmit content on next hop only if it receives the same content on previous hop. Constraint (9) ensures that each multicast group $\mathcal{K}_{s_j}$ should be no less than the minimum rate requirement R_{min} to ensure quality of service. Constraint (10) ensures that each MD can receive the content once among the total H_{max} hops. Constraint (11) is the maximum hop tolerance related to the delay problem in practice.

C. COMPLEXITY

The considered optimization problem is a mixed integer programming (MIP) problem, which is NP-complete due to the combinatorial nature with the binary variables [25]. The number of binary variables in (5)-(11) is $C^2 + H_{max}C^2$, thus, the worst complexity of determining the optimal multicast group division of our problem is $\mathcal{O}(2^{(C^2 + H_{max}C^2)})$.

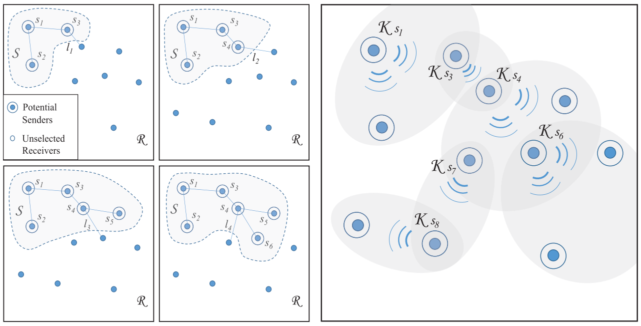
Finding the optimal solution will exponentially increase the complexity with the number of MDs and transmit hops, which is prohibitive in practice. Hence, we turn to propose practical suboptimal methods with lower complexity, which are detailed in next section.

III. PROPOSED ALGORITHMS IN QUASI-STATIC NETWORKS

In this section, three heuristic algorithms in quasi-static D2D networks are proposed. In general, the core of these heuristic algorithms is to devise certain mechanisms of building up D2D connections and dividing the MDs into different multicast groups in multihop networks. We design three heuristic algorithms according to diverse greedy strategies, which are presented in detail by the following three subsections respectively.

A. RECEIVER-ORIENTED ALGORITHM

In this algorithm, we suppose $H_{max} = C$, which is applicable for networks with low-density MDs or delay-tolerant applications. The initial sets of transmitters and receivers are denoted by $\mathcal{S} = \{s_1, s_2, \ldots, s_j\}$ and $\mathcal{R} = \{r_1, r_2, \ldots, r_i\}$, respectively, where s_j and r_i represent the jth transmitter and the ith receiver, respectively. Denote $\mathcal{K} = \{\mathcal{K}_{s_1}, \mathcal{K}_{s_2}, \ldots, \mathcal{K}_{s_j}\}$ as the set of disjoint subsets which represent the final multicast groups with s_1, s_2,…, s_j as transmitters. At first, BS is the only element in $\mathcal{S}$ and $\mathcal{R}$ contains all MDs in total as the receivers. That is, the content delivery starts at BS. The algorithm realizes the grouping of MDs by a iterative procedure: every time there is only one MD with the largest CNR $\gamma_{s_j r_i}$ to be chosen from the set $\mathcal{R}$ as the receiver for s_j. Such a process continues until all MDs in $\mathcal{R}$ are linked. Specifically, assume that the link between ith MD in $\mathcal{R}$ and the jth MD in $\mathcal{S}$ has the best channel condition among all the possible links. Then, the algorithm sets the jth MD to be the transmitter for the ith MD, and the ith MD shifts from $\mathcal{R}$ into $\mathcal{S}$ and becomes a potential transmitter for the next hop. That is, the potential transmitter always chooses its best receiver in the proposed receiver-oriented algorithm. At last, the MDs with a common transmitter are grouped together as the same multicast group.



(a) (b)

FIGURE 2. (a) An example of the grouping process of Algorithm 1. $s_1, s_2, \ldots, s_n$ are MDs in $\mathcal{S}$ acting as potential transmitters. $l_1, l_2, \ldots, l_n$ are the links with the maximum CNR. (b) The final formed multicast multihop D2D cooperative network.

Algorithm 1 Receiver-Oriented Solution

```
1: Initialize $\mathcal{S} = \{BS\}$, $\mathcal{R} = \{\mathcal{C}\}$, $\mathcal{K} = \emptyset$.
2: while $\mathcal{R} \neq \emptyset$ do
3:     Select $s_j \in \mathcal{S}$ and $r_i \in \mathcal{R}$ that have the maximum $\gamma_{s_j r_i}$
4:     Let MD $s_j$ be the transmitter for MD $r_i$
5:         $\mathcal{K}_{s_j} \leftarrow \mathcal{K}_{s_j} \cup r_i$
6:     Update the sets of transmitters and receivers as:
7:         $\mathcal{S} \leftarrow \mathcal{S} \cup r_i$
8:         $\mathcal{R} \leftarrow \mathcal{R} \backslash r_i$
9: end while
10: Calculate power consumption of each multicast group as
    (4) to satisfy the rate constraint.
11: Calculate total power consumption of all groups.
```

An example of the grouping process is shown in Fig. 2. In the figure, $\mathcal{S}$ contains 3 MDs at the beginning. The MDs in $\mathcal{R}$ are selected one by one according to the rule of choosing the link with the largest CNR. In this example, links $l_1, \ldots, l_4$ are successively established. After all D2D connections are established as in Fig. 2(a), MDs with the same transmitter are allocated in a multicast group as shown in Fig. 2(b)

The power consumption of each multicast group can be calculated by (4), and the total power consumption is the sum power of all multicast groups. Formally, the whole solution is described in Algorithm 1. In this algorithm, the transmitter set $\mathcal{S}$ contains at most $(C + 1)$ elements and the receiver set $\mathcal{R}$ contains at most C elements. The computationally complexity of line 3 in Algorithm 1 is $\mathcal{O}(C^2)$ for finding the MD with the maximum CNR. This process is executed for C times and thus, the whole complexity of Algorithm 1 is $\mathcal{O}(C^3)$.

Note that, there would be numerous relays generated by the receiver-oriented algorithm if the network involved a high density of MDs, which may trigger the delay problem. In addition, the complexity of this algorithm is $\mathcal{O}(C^3)$, which may be high for a network with a large number of MDs. In following, we propose a cluster-oriented algorithm with lower complexity.

B. CLUSTER-ORIENTED ALGORITHM

In this subsection, we propose a cluster-oriented algorithm which aims at reducing total power consumption by minimizing the number of multicast groups, or equivalently, minimizing the number of transmitters. Nevertheless, this problem is NP-complete even in a single hop.

Theorem 1: Determining the minimum number of transmitters on a single hop is a NP-complete problem.

Proof: For example, denote $\mathcal{C} = \{\{1, 3, 4\}, \{2, 3, 5\}, \{2, 4\}, \{1, 4\}\}$, where the MDs in boldface represent transmitters in their own subsets. It is easy to choose $\{1, 4\}$ and $\{2, 3, 5\}$ as the result, because they have the minimum number of subsets meanwhile cover all elements. This problem is the classical set cover problem which is defined as follow: Given a set $\mathcal{A}$ and disjoint subsets $\mathcal{A}_{s_1}, \mathcal{A}_{s_2}, \ldots \mathcal{A}_{s_j} \subseteq \mathcal{A}$, the problem is to select some of these disjoint subsets which have the minimum number and contain all elements, i.e., $\mathcal{A}_{s1} \cup \mathcal{A}_{s2}, \ldots \cup \mathcal{A}_{sj} = \mathcal{A}$. The set cover problem is NP-complete by definition [26]. ■

Since the set cover problem cannot be optimally solved in polynomial time, a commonly adopted method is heuristic greedy algorithm. Here, we apply the clustering idea to solve this problem. The idea is to involve receivers as many as possible for multicast in each group if the CNR is greater than or equals to a predefined threshold γ_0 to satisfy the minimum rate R_{min}. Because in some sense, minimizing the number of multicast groups can result in reducing total transmission power consumption. We define that $\mathcal{A} = \{\mathcal{A}_{s_1}, \mathcal{A}_{s_2}, \ldots \mathcal{A}_{s_j}\}$, where the subset $\mathcal{A}_{s_j}$ contains element r_i in $\mathcal{R}$ (i.e., all uncovered MDs) if $\gamma_{s_j r_i} \geq \gamma_0$. It should be noted that a MD can be included by more than one subset so it may needed a further selection and $\mathcal{A}$ is not the final grouping result.

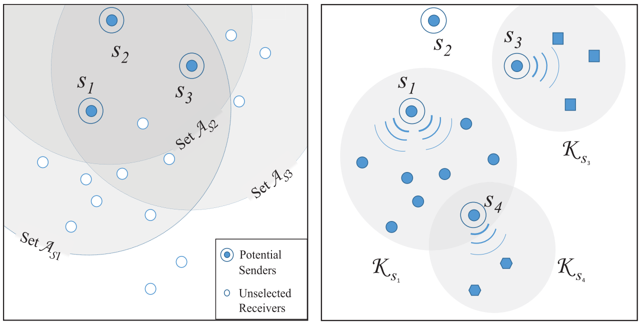
Algorithm 2.1 presents the greedy set cover method to deal with the grouping problem on a specific hop, where inputs

Algorithm 2.1 Greedy Set Cover Method

```
Step 1. $\mathcal{I} \leftarrow \mathcal{R}$.
Step 2. while $\mathcal{I} \cap \mathcal{A} \neq \emptyset$
            a: select $\mathcal{A}_{s_j} \subseteq \mathcal{A}$ that maximizes $\{\mathcal{I} \cap \mathcal{A}_{s_j}\}$
               $\mathcal{K}_{s_j} \leftarrow \mathcal{I} \cap \mathcal{A}_{s_j}$
            b: $\mathcal{K} \leftarrow \mathcal{K} \cup \mathcal{K}_{s_j}$
            c: $\mathcal{I} \leftarrow \mathcal{I} \backslash \mathcal{A}_{s_j}$
        end while
Step 3. return $\mathcal{K}$
```

are $\mathcal{R}$ and $\mathcal{A}$ and the final grouping result is recorded by $\mathcal{K}$. Initially, $\mathcal{I}$ heritages all elements from $\mathcal{R}$, which is a set of all the uncovered receivers. As a consequence, the intersection of $\mathcal{I}$ and $\mathcal{A}$ indicates the set of uncovered MDs whose CNR is greater than or equal to γ_0, i.e., the set of potential receivers on current hop. While there exist uncovered elements, the MDs already covered will connect the remainders on next hop. On each hop, the set which involves the most MDs satisfying the minimum rate R_{min} is selected to be a new multicast group $\mathcal{K}_{s_j}$ with the transmitter s_j. Meanwhile, $\mathcal{I}$ removes those the elements intersected with $\mathcal{A}_{s_j}$ since they belong to a new multicast group. It is worth mentioning that when each iteration is done, all possible sets of $\mathcal{I} \cap \mathcal{A}_{s_j}$ (i.e., uncovered MDs maintaining the minimum rate R_{min}) may be changed since some of their elements may not exist in $\mathcal{I}$. Thus, by judging the size of $\mathcal{I} \cap \mathcal{A}_{s_j}$ instead of $\mathcal{A}_{s_j}$, we can correctly select the largest set from the remainders.

We provide an example in Fig. 3 to illustrate the grouping process on a specific hop of the cluster-oriented algorithm. Initially, there are three elements in transmitter set $\mathcal{S}$, i.e., $\mathcal{S} = \{s_1, s_2, s_3\}$. Subset $\mathcal{A}_{s_1}$, involving the most potential receivers which maintain the minimum rate R_{min} with transmitter s_1, is formed and shown in Fig. 3(a). Therefore, the multicast group $\mathcal{K}_{s_1}$ is formed when the first iteration is finished. Then by removing $\mathcal{K}_{s_1}$ from the set, $\mathcal{K}_{s_3}$ is formed. After several iterations, grouping process ends and the final result is shown in Fig. 3(b).



(a) (b)

FIGURE 3. An example of Algorithm 2.1: (a) Every MD that has already been linked finds all unlinked MDs satisfying the rate constraint and form its own subset. (b) According to the greedy rule, a transmitter with the largest number of receivers is selected to form group.

Algorithm 2.2 Greedy Clustering Solution

```
1: Initialize $\mathcal{S} = \{BS\}$, $\mathcal{R} = \{\mathcal{C}\}$, $\mathcal{K} = \emptyset$, $h = 1$.
2: for $h = 1$ to $H_{max}$ do
3:     a. $\mathcal{A}_{s_j} = \mathcal{A}_{s_j} \cup r_i$ if $\gamma_{s_j r_i} \geq \gamma_0$, $\forall s_j \in \mathcal{S}$, $\forall i \in \mathcal{I}$.
4:     b. Run Algorithm 2.1 and obtain the grouping result
          $\mathcal{K}$ on current hop.
5:     c. Update the sets of transmitters and receivers as:
              $\mathcal{S} \leftarrow \mathcal{S} \cup \mathcal{K}$
              $\mathcal{R} \leftarrow \mathcal{R} \backslash \mathcal{K}$
       d. Increase the number of hops $h = h + 1$.
6: end for
7: Calculate power consumption of each multicast group as
   (4) to satisfy the rate constraint.
8: Calculate total power consumption of all groups.
```

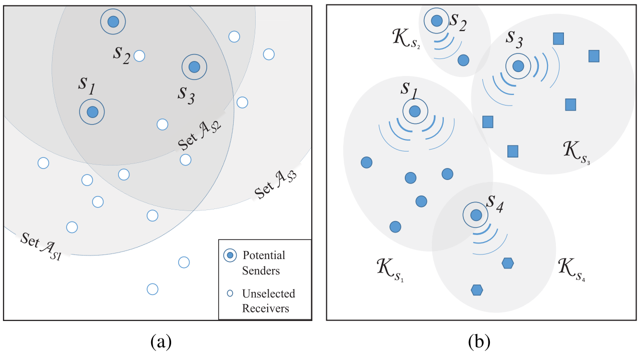
In Algorithm 2.2, we present the whole algorithm to determine the grouping result in multihop. It should be notice that some MDs may remain uncovered when a whole process is done because they cannot meet the request of γ_0 with any transmitter within H_{max} hops. When such situation occurs, the algorithm reduces the CNR threshold γ_0 so that more MDs can be involved on each hop, and accordingly increases the transmit power to maintain the minimum rate constraint. The determining of γ_0 is based on experiential simulations and we do not discuss the details in the paper.

In Algorithm 2.2, the transmitter set $\mathcal{S}$ contains at most $(C + 1)$ elements and the receiver set $\mathcal{R}$ contains at most C elements. Therefore, the process of updating set $\mathcal{A}$ has complexity of $\mathcal{O}(C^2)$. Moreover, executing line 4 is $\mathcal{O}(C^2)$ for searching the subset with maximum number of elements. Since line 3 and line 4 are iterated at most H_{max} times and thus, the total complexity of cluster-oriented algorithm is $\mathcal{O}(H_{max} C^2)$.

C. TRANSMITTER-ORIENTED ALGORITHM

The idea of the cluster-oriented algorithm is to reduce the number of transmitters so as to reduce the total power consumption. However, it may cause high power consumption on the transmitter of a multicast group since the transmitter should increase its transmit power to satisfy the rate constraint of all MDs in the group. Moreover, such a power distribution can bring the unfairness problem. So we try to take this problem into consideration and propose a new transmitter-oriented algorithm.

Firstly, like the cluster-oriented algorithm, we let $\mathcal{A} = \{\mathcal{A}_{s_1}, \mathcal{A}_{s_2}, \cdots, \mathcal{A}_{s_j}\}$, where the subset $\mathcal{A}_{s_j}$ denotes the multicast group including MD r_i that satisfies $\gamma_{s_j r_i} \geq \gamma_0$ with transmitter s_j. Then, on each hop, every MD in $\mathcal{A}$ picks up its receivers that maintain the minimum rate R_{min} with the predefined γ_0. Since a MD may be selected by multiple transmitters, each of the selected MDs reversely chooses its transmitter with the strongest CNR. Compared with the receiver-oriented algorithm, the MDs in $\mathcal{A}$ choose their transmitters in $\mathcal{S}$ so that more than one link can be determined



(a) (b)

FIGURE 4. An example for Algorithm 3. (a) Every MD that has already been linked finds all the unlinked MDs satisfying the rate constraint with predefined γ_0 and forms its own subset. (b) Every MD selects the transmitter with strongest channel for itself among all transmitters.

every time on a specific hop in the transmitter-oriented algorithm. Thus the complexity can be reduced. In addition, possibly more transmitters are chosen every time so we can suppose that the power distribution is relatively more fair by this way.

We take an example in Fig. 4 to illustrate the transmitter-oriented algorithm on a specific hop. Assume that $\mathcal{S}$ contains three transmitters, i.e., $\mathcal{S} = \{s_1, s_2, s_3\}$. Subsets $\mathcal{A}_{s_1}, \mathcal{A}_{s_2}, \mathcal{A}_{s_3}$, including MDs which surpass the predefined CNR γ_0 from transmitter s_j ($j = 1, 2, 3$) on current hop, are formed and shown in Fig. 4(a). Then every receiver chooses its transmitter with best channel for itself and the final grouping result is shown in Fig. 4(b). From Fig. 2 and Fig. 4, the grouping results of the transmitter-oriented and receiver-oriented algorithms are different in the same scenario and the transmitter-oriented algorithm has lower power consumption on a single transmitter. Like the cluster-oriented algorithm, if H_{max} hops cannot cover all MDs, the algorithm decreases the CNR threshold γ_0 to involve more MDs on each hop and correspondingly increases transmit power of each multicast group to satisfy the rate constraint.

In Algorithm 3, the transmitter set $\mathcal{S}$ contains at most $(C + 1)$ elements and the receiver set $\mathcal{R}$ with C elements. The complexity of selecting links on line 5 is $\mathcal{O}(C)$ and it repeates at most C times. Since it starts from the BS to H_{max} hops, The total complexity of Algorithm 3 is $\mathcal{O}(H_{max}C^2)$.

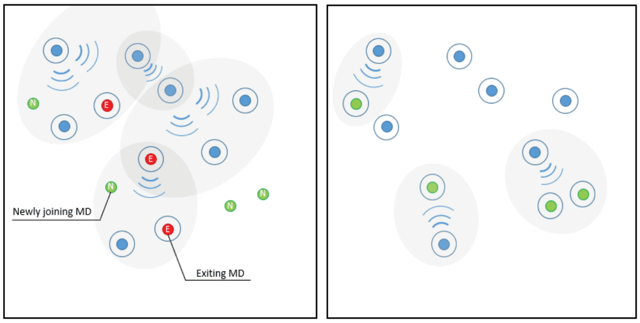
IV. EXTENSION TO DYNAMIC NETWORKS

In this section, we extend the proposed three algorithms to dynamic networks. In real scenario, the network is dynamic where every MD may join or quit the network at different time. Specifically, some MDs may quit the cooperative transmission after they obtain the content, and other new MDs having the content request may join the network. Thus, the content delivery policy of the D2D network needs to be re-optimized. It is interesting that the three proposed algorithms in above section can be easily extended to this dynamic situation. In this section, we present the details.

For ease of exposition, a time index t with finite and discrete slots is used. We set T time slots to describe a dynamic process over this time index t. In each time slot, a random number of MDs in the range from 0 to N, are newly joining the network for requesting the content. Meanwhile, the MDs which receive the content on the previous time slots may exit from the D2D cooperative transmission with a certain probability f. The influence of these parameters will be discussed in simulation. The dynamic process is shown in Fig. 5, where green dots represent newly joining MDs and the red dots represent exiting MDs in Fig. 5(a). When some transmitters exit and newly receivers emerge, MDs are regrouped and the network still ensures that the newly joining MDs can receive content. Fig. 5(b) shows the regrouped result.

Algorithm 3 Transmitter-Oriented Solution

```
1:  Initialize $\mathcal{S} = \{BS\}$, $\mathcal{R} = \{\mathcal{C}\}$, $\mathcal{K} = \emptyset$, $h = 1$, $\mathcal{I} \leftarrow \mathcal{R}$.
2:  for $h = 1$ to $H_{max}$ do
3:      $\mathcal{A}_{s_j} = \mathcal{A}_{s_j} \cup r_i$ if $\gamma_{s_j r_i} \geq \gamma_0$, $\forall s_j \in \mathcal{S}$, $\forall i \in \mathcal{I}$.
4:      while $\mathcal{I} \cap \mathcal{A} \neq \emptyset$ do
5:          Select MD $s_j \in \mathcal{S}$ and MD $r_i \in \mathcal{I}$ with maximum $\gamma_{s_j r_i}$.
6:              $\mathcal{K}_{s_j} \leftarrow \mathcal{K}_{s_j} \cup r_i$
7:              $\mathcal{I} = \mathcal{I} \backslash r_i$
8:      end while
9:      Update the sets of transmitters and receivers as:
10:             $\mathcal{S} \leftarrow \mathcal{S} \cup \mathcal{K}$
11:             $\mathcal{R} \leftarrow \mathcal{R} \backslash \mathcal{K}$
12:     Increase the number of hops $h = h + 1$.
13: end for
14: Calculate power consumption of each multicast group as
    (4) to satisfy the rate constraint.
15: Calculate total power consumption of all groups.
```



(a) (b)

FIGURE 5. An example for dynamic situation. (a) newly joining MDs and exiting MD are determined in networks. (b) a new topological network is formed for content delivery.

Specifically, in Algorithm 4, $\mathcal{S}$ and $\mathcal{R}$ still represent the sets of the transmitters and the receivers, respectively. In the end of each slot, the left MDs and the newly joined MDs are regrouped by the three proposed algorithms respectively. Although $\mathcal{S}$ and $\mathcal{R}$ vary dynamically during the process of the three proposed algorithms, $\mathcal{S}$ at least has one transmitter, i.e. BS, and thus it does not affect the algorithms. Therefore, suppose that all transmitters exit from $\mathcal{S}$ except BS in a

Algorithm 4 Content Delivery in Dynamic D2D Network

```
Initialize $T$, $\mathcal{S} = \{BS\}$, $\mathcal{R} = \{\mathcal{C}\}$.
for $t = 1$ to $T$ do
    a. Assume that requesting MDs randomly access the network with uniformly distributed integer numbers between $[0, N]$.
    b. Run Algorithm 1, Algorithm 2.2 and Algorithm 3 respectively on time index $t$.
    c. Update $\mathcal{S}$ by executing the exiting MDs and update $\mathcal{R}$ by involving the newly joining MDs.
end for
```

time slot, the problem degenerates to the first time slot, i.e., the content delivery starts with BS.

V. SIMULATION RESULTS

In this section, the performance of three proposed algorithms are evaluated by numerical results. The stimulation parameters are set up as follows. We assume a circle area with a radius of 500 meters where MDs are randomly distributed. A BS locates at the center point of this circle area. The MDs require the same content that is initially stored in BS. Considering more general cases, we set the noise N_0 to be -100dBm. The channel parameters are: the constant K equals -31.54dB, the path loss exponent is $\beta = 3$, $d_0 = 1$m, and ϕ_{dB} is a zero-mean Gaussian random variable which represents the effects of shadow fading [27]. The required rate $R_{min} = 10$ (bit/s/Hz). H_{max} is considered to be C hops. These parameters remain unchanged when discussing dynamic network.

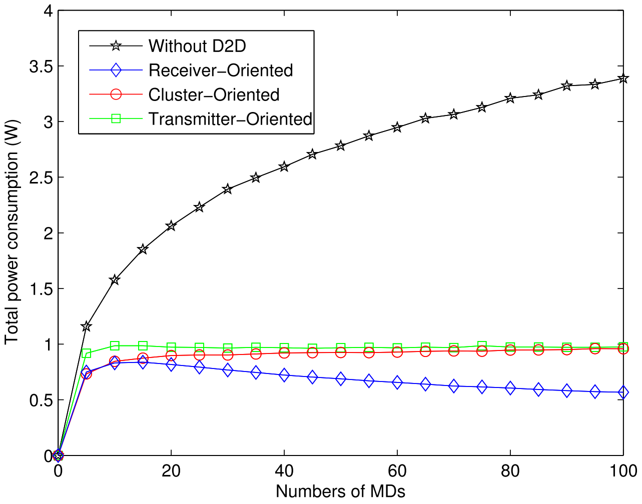


FIGURE 6. The total transmit power versus the number of MDs.

A. STATIC NETWORK

Fig. 6 shows the total transmit power consumption by the three proposed algorithms. For the purpose of performance comparison, the scheme of the BS broadcasting is also considered as a benchmark in the figure. We can see that the three proposed algorithms significantly outperform the traditional scheme without multihop D2D communications. This is because D2D connections always have much shorter distances so that they can guarantee higher transmission rates with much less power consumption. Therefore, the D2D network works more efficiently when compared to traditional cellular network with only BS to deliver the content. In addition, a trend can be observed that the power consumption goes down as the number of MDs increases if the number of MDs is greater than 10. The reason is that the increase of node density lowers the D2D distances generally and thus reduces the transmit power. For the cluster-oriented algorithm, the variation of power consumption is much smaller than the receiver-oriented algorithm. The performance of the cluster-oriented algorithm is slightly better than the transmitter-oriented algorithm and the difference is vanishing as the number of MDs increases. The receiver-oriented algorithm has the best performance in total power consumption in spite of its uncontrolled hops which may bring delay problem. Therefore, the receiver-oriented algorithm is the best in performance of power-saving if the content is delay-tolerant.

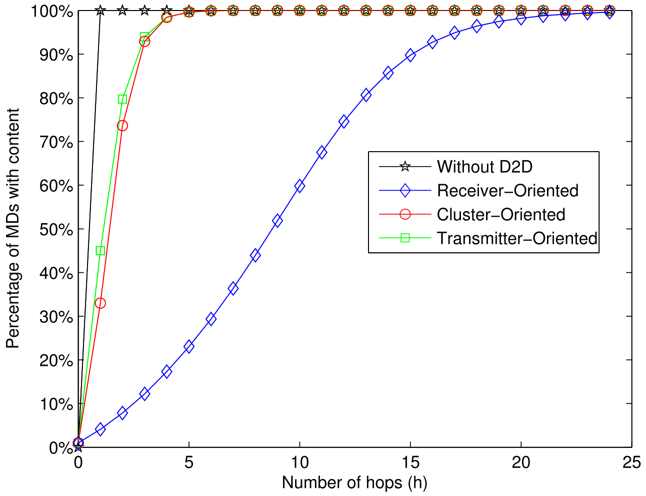


FIGURE 7. The percentage of MDs with the content versus the number of hops when there are 100 MDs in total.

Fig. 7 shows that the receiver-oriented algorithm generates more hops than the other two algorithms. If there are 100 MDs, the receiver-oriented algorithm requires up to 30 hops to make all MDs receive the content, while the other two algorithms just need about 5 hops. This also means that the remote MDs receive the content after waiting for a relatively longer time until the previous MDs have obtained the content. In general, the D2D networks formed by the transmitter- and cluster- oriented algorithms need much less hops than that by the receiver-oriented algorithm and the difference will be larger if the number of MDs increases. When below 5 hops, the transmitter-oriented algorithm has higher cover-ratio (percentage of MDs with content) than the cluster-oriented algorithm. As the number of hops increases, the stability of the network is unavoidably influenced by the multihop connections because the multihop delay may cause synchronization issues. Therefore, we reach a conclusion that the transmitter-oriented algorithm can form more favorable D2D networks for cooperative multihop multicast transmission when the networks is large and delay is regarded as an important issue.

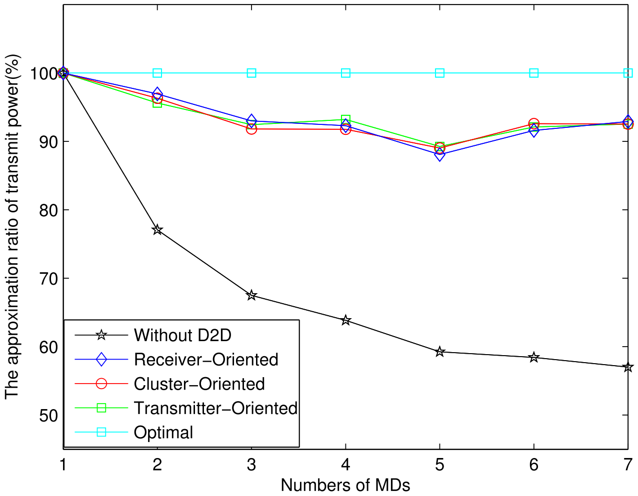


FIGURE 8. **The approximation ratio of transmit power between proposed algorithms and optimal result.**

We also compare all proposed algorithms with the optimal solution obtained by exhaustive search shown in Fig 8. The approximation ratio is defined as the transmit power of the proposed algorithms divided by the transmit power of the optimal solution. From the figure we can see that the results of the proposed algorithms are close to each other and at least 90% optimality can be achieved. When the number of MDs is small, the power consumption of the three algorithms is near optimal. Here we only simulate 7 MDs due to the exponential complexity by the exhaustive search.

Fig. 9 illustrates the relationship between the total number of requesting MDs and the number of MDs which act as relays. The figure shows a noticeable difference among the three proposed algorithms. Note that the number of transmitters is equal to the number of multicast groups. We can see that the transmitter number in the receiver-oriented algorithm is the largest and increases steadily. This is mainly attributed to that the network formed by such an algorithm delivers content via more hops than others. The transmitter-oriented algorithm does not need to choose a global strongest channel. Thus, it needs less transmitters compared to the receiver-oriented algorithm. In addition, the cluster-oriented algorithm involves more MDs in each group. So, it is reasonable that its number of transmitters is the least one and it goes up much more slowly when the requesting MDs increase.

Fig. 10 shows the relationship between average transmit power of each transmitter and the number of requesting MDs. It can reflect the distribution degree of transmission power and show the fairness of the generated network to some extent. We can see that all the average power consumptions

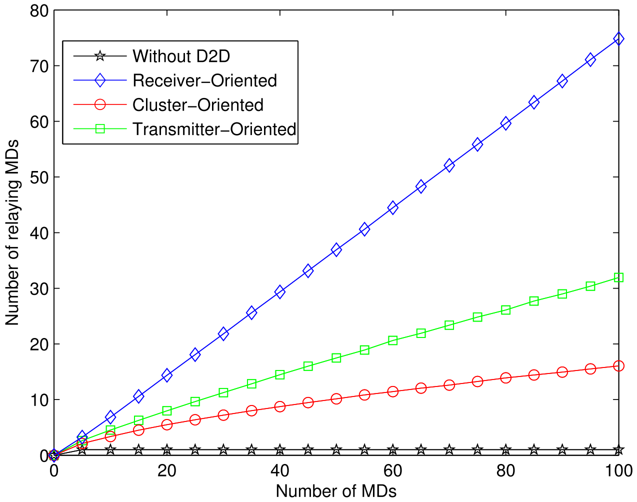


FIGURE 9. **The number of MDs act as relays in the grouping results by three proposed algorithms versus the total number of MDs.**

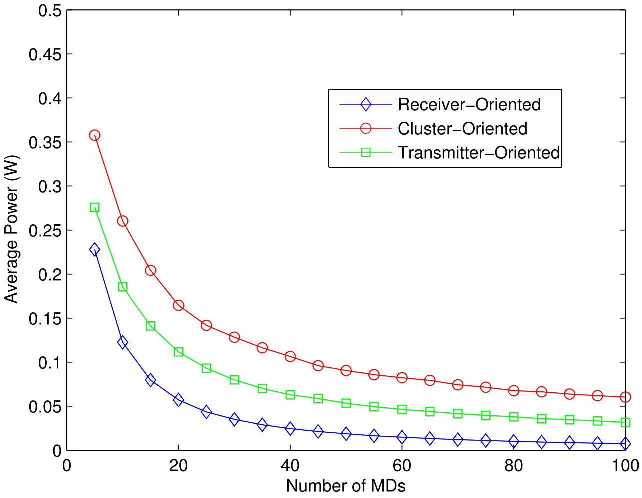


FIGURE 10. **The average power on each transmitter versus the number of MDs.**

have significant decline as the number of requesting MDs increases. This is because the distances among MDs are smaller in general and thus the transmit power can be relieved if the number of MDs increases. More specifically, the cluster-oriented algorithm tries to cover all requesting MDs with the least multicast groups at the price of a heavy burden of power consumption at each transmitter. The receiver-oriented algorithm chooses the global maximum value of γ_0 to form more multicast groups and has the lowest average transmit power among the three algorithms. However, the transmitter-oriented algorithm only chooses the local maximum value of γ_0 to form groups and its performance is between the other two algorithms.

B. DYNAMIC NETWORK

Fig. 11 shows the transmit power consumption from the first time slot to the last time slot. N represents the upper bound of the number of joining MDs. The scenarios

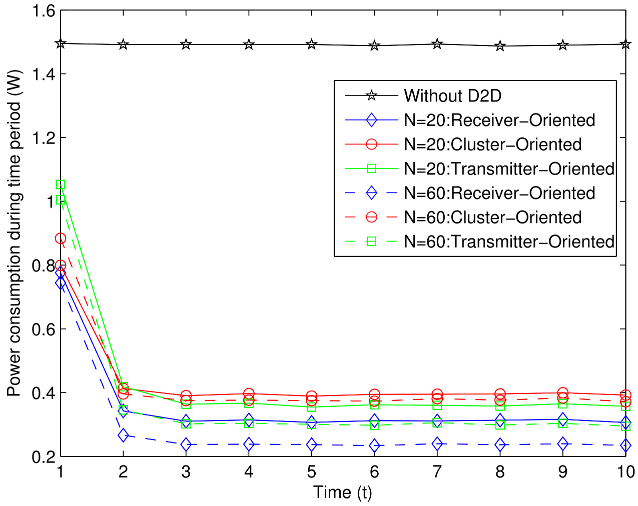


FIGURE 11. **Power consumption versus time, where $f = 100\%$.**

for $N = 20$ and $N = 60$ are considered here. One can observe that a considerable decrease occurs from the first time slot to the second time slot and then the transmit power level remains stable. This is because that the newly formed cooperative D2D network contains much more transmitters or relays than the original network which only has the BS to be the first transmitter. The average distances between transmitters and receivers decrease so the links are more likely to be of high CNR. Compare to the benchmark without D2D, the proposed algorithms significantly relieve the BS transmit power pressure and decrease total transmit power. An interesting observation is that the receiver- and transmitter-oriented algorithms have better performance in the case of $N = 60$ than $N = 20$ on the first time slot while the cluster-oriented algorithm performs in contrast. The performance of the three algorithms on the first time slot is consistent with Fig. 6. As time goes on, although all unused transmitters exit, the networks is still stable which clearly shows that the proposed algorithms are capability of power-saving to a large extent.

More specifically, the receiver-oriented algorithm still consumes the least transmit power in the whole dynamic process showing a good performance maintaining ability. Another observed interesting point is that the power consumption of transmitter-oriented algorithm is slightly less than that of the cluster-oriented algorithm after the third time slot. So it can be concluded that the transmitter-oriented algorithm is more suitable for the original scenario with multiple transmitters.

Because the cluster-oriented algorithm aims at delivering content with fewer multicast groups as possible, it ignores the advantages of multiple potential transmitters after the first time slot. The transmitter-oriented algorithm can balance the number of transmitter-receiver connections and the size of multicast groups, thus its performance is better than the cluster-oriented algorithm in a long term. In the first time slot with only BS, the receiver-oriented algorithm performs better while other two algorithms remain almost unchanged. We can conclude that increasing number of requesting MDs also promotes network performance in dynamic process as the number of relays increases and the transmit distances decrease.

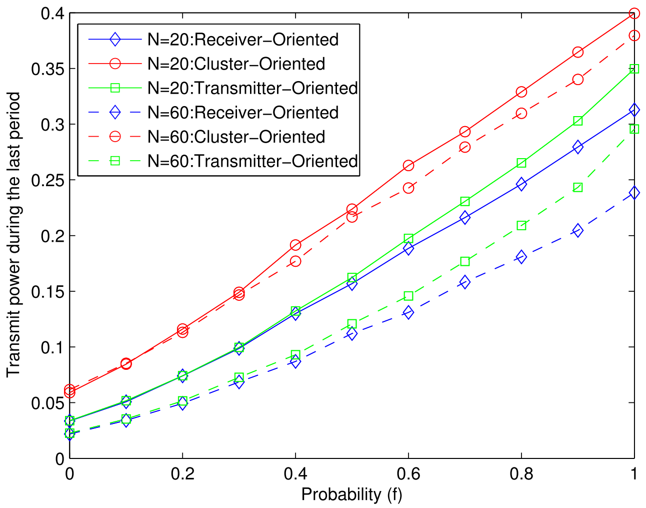


FIGURE 12. **Transmit power versus probability, where $f = 100\%$.**

We note that f is the probability represents that transmitters in $\mathcal{S}$ without transmit task may exit from the network. This probability obviously impacts the performance of whole network. To discuss the relationship between them, we investigate power consumption at the end of the whole transmission (i.e., $t = T$) with probability f in Fig. 12. We can see that the total power consumption increases as the probability f increases (i.e., more transmitters exit). We also observe that the variation of N has little influence on the cluster-oriented algorithm while it has obvious impact on the other two algorithms.

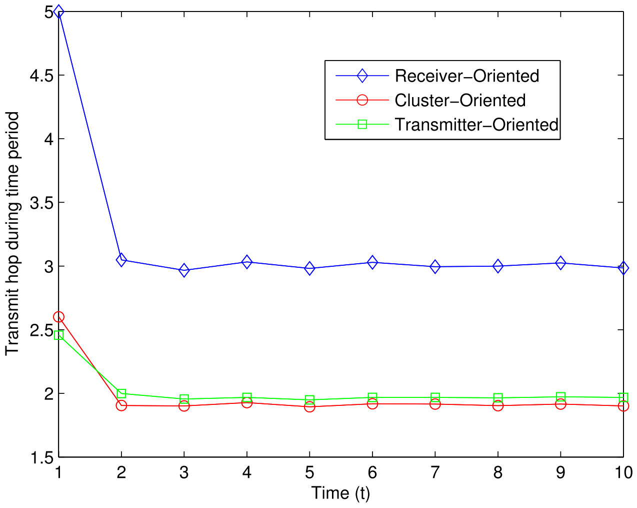


FIGURE 13. **Transmit hops versus time, where $N = 20\%$ and $f = 100\%$.**

Fig. 13 shows the number of transmit hop for three algorithms in dynamic network with $N = 20\%$ and $f = 100\%$.

Although the receiver-oriented algorithm performs best, it needs more hops to cover all MDs. The number of transmit hop in first time slot is 5 and above 3 hops with multiple transmitters in the remaining time slots. The cluster- and transmitter-oriented algorithms perform closely and need about 2 hops on average.

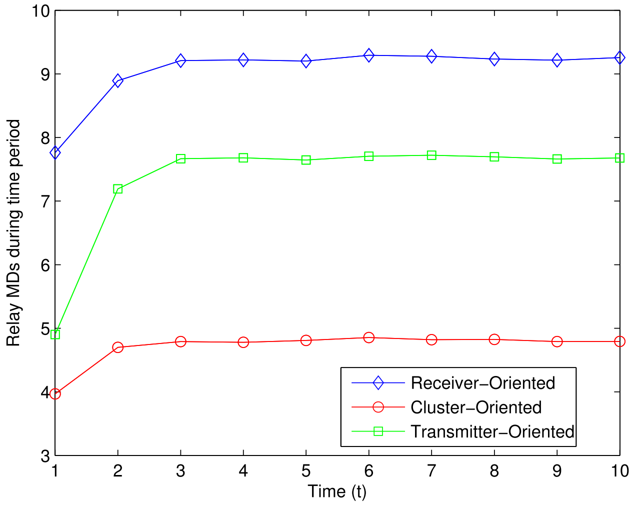


FIGURE 14. Number of relay MDs versus time, where $N = 20\%$ and $f = 100\%$.

The needed number of relaying MDs in every time slot is shown in Fig 14, where $N = 20\%$ and $f = 100\%$. We can observe that the number of relays varies widely for the three algorithms. There is an upward trend for every algorithm from the first time slot to the third time slot and little variation in the remaining time slots. As the cluster-oriented algorithm tries to gather as many MDs as possible into a group, the number of relay MDs is the minimum among the three algorithms. On the other hand, the result reveals that the cluster-oriented algorithm may face more serious fairness problem as its power consumption is of highly uneven distribution throughout a dynamic process.

VI. CONCLUSION

In this paper, we studied the problem of the power minimization in multicast multihop D2D network in a content distribution scenario. The optimization problem is NP-complete and we proposed three efficient suboptimal algorithms to solve the problem in polynomial time. Simulation results showed that the proposed algorithms can significantly reduce power consumption compared with the traditional BS multicast networks. We also extended the proposed algorithm into dynamic network scenario. Several insights were obtained for practical system designs.

REFERENCES

[1] Z. Xia, J. Yan, and Y. Liu, "Energy efficiency in multicast multihop D2D networks," in *Proc. IEEE/ICC ICCC*, Sep. 2016, pp. 1–6.

[2] A. Aijaz, H. Aghvami, and M. Amani, "A survey on mobile data offloading: Technical and business perspectives," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 104–112, Apr. 2013.

[3] V. Sciancalepore, D. Giustiniano, A. Banchs, and A. Hossmann-Picu, "Offloading cellular traffic through opportunistic communications: Analysis and optimization," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 1, pp. 122–137, Jan. 2016.

[4] X. Kang, Y.-K. Chia, S. Sun, and H. F. Chong, "Mobile data offloading through a third-party WiFi access point: An operator's perspective," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5340–5351, Oct. 2014.

[5] B. Han, P. Hui, V. S. A. Kumar, M. V. Marathe, J. Shao, and A. Srinivasan, "Mobile data offloading through opportunistic communications and social participation," *IEEE Trans. Mobile Comput.*, vol. 11, no. 5, pp. 821–834, May 2012.

[6] Y. Liu, "Optimal mode selection in D2D-enabled multibase station systems," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 470–473, Mar. 2016.

[7] A. Pyattaev, K. Johnsson, A. Surak, R. Florea, S. Andreev, and Y. Koucheryavy, "Network-assisted D2D communications: Implementing a technology prototype for cellular traffic offloading," in *Proc. Wireless Commun. Netw. Conf. (WCNC)*. Sep. 2014, pp. 3266–3271.

[8] L. Al-Kanj, H. V. Poor, and Z. Dawy, "Optimal cellular offloading via device-to-device communication networks with fairness constraints," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4628–4643, Aug. 2014.

[9] M. Sheng, Y. Li, X. Wang, J. Li, and Y. Shi, "Energy efficiency and delay tradeoff in device-to-device communications underlaying cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 1, pp. 92–106, Jan. 2016.

[10] J.-H. Yun and K. G. Shin, "Distributed coordination of co-channel femtocells via inter-cell signaling with arbitrary delay," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1127–1139, Jun. 2015.

[11] D. Calin, H. Claussen, and H. Uzunalioglu, "On femto deployment architectures and macrocell offloading benefits in joint macro-femto deployments," *IEEE Commun. Mag.*, vol. 48, no. 1, pp. 26–32, Jan. 2010.

[12] C. Rossi, C. Casetti, C. F. Chiasserini, and C. Borgiattino, "Cooperative energy-efficient management of federated WiFi networks," *IEEE Trans. Mobile Comput.*, vol. 14, no. 11, pp. 2201–2215, Nov. 2015.

[13] H. Liu, Z. Chen, and L. Qian, "The three primary colors of mobile systems," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 15–21, Sep. 2016.

[14] K. Wang, Z. Chen, and H. Liu, "Push-based wireless converged networks for massive multimedia content delivery," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2894–2905, May 2014.

[15] Z. Zhao, M. Peng, Z. Ding, and W. Wang, "Cluster content caching: An energy-efficient approach to improve quality of service in cloud radio access networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1207–1220, May 2016.

[16] Y. Cao, T. Jiang, X. Chen, and J. Zhang, "Social-aware video multicast based on device-to-device communications," *IEEE Trans. Mobile Comput.*, vol. 15, no. 6, pp. 1528–1539, Jun. 2015.

[17] Y. Liu, R. Wang, and Z. Han, "Interference-constrained pricing for D2D networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 475–486, Jan. 2016.

[18] J. Whitbeck, Y. Lopez, J. Leguay, V. Conan, and M. D. D. Amorim, "Push-and-track: Saving infrastructure bandwidth through opportunistic forwarding," *Pervas. Mobile Comput.*, vol. 8, no. 5, pp. 682–697, May 2012.

[19] X. Wang, "Content dissemination by pushing and sharing in mobile cellular networks: An analytical study," in *Proc. IEEE 9th Int. Conf. Mobile Ad Hoc Sensor Syst. (MASS)*, Apr. 2012, pp. 353–361.

[20] N. Ristanovic, J. Y. L. Boudec, A. Chaintreau, and V. Erramilli, "Energy efficient offloading of 3G networks," in *Proc. 8th IEEE Int. Conf. Mobile Ad-Hoc Sensor Syst.*, Sep. 2011, pp. 202–211.

[21] L. Al Kanj and Z. Dawy, "Offloading wireless cellular networks via energy-constrained local ad hoc networks," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2011, pp. 1–6.

[22] Y. Li, G. Su, P. Hui, D. Jin, L. Su, and L. Zeng, "Multiple mobile data offloading through delay tolerant networks," in *Proc. Int. Workshop Challenged Netw.*, Las Vegas, NV, USA, vol. 13, no. 7, p. 1, Jul. 2011.

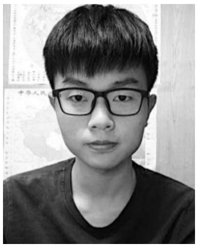
[23] L. Militano, M. Condoluci, G. Araniti, and A. Molinaro, "Wi-Fi cooperation or D2D-based multicast content distribution in LTE-A: A comparative analysis," in *Proc. Workshop Cooperat. Cognit. Mobile Netw. (ICC)*, Jun. 2014, pp. 767–776.

[24] J. Seppala, T. Koskela, T. Chen, and S. Hakola, "Network controlled device-to-device (D2D) and cluster multicast concept for LTE and LTE-A networks," in *Proc. Wireless Commun. Netw. Conf. (WCNC)*, 2011, pp. 986–991.

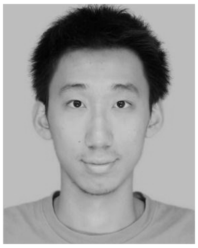
[25] M. Jünger *et al. 50 Years of Integer Programming 1958–2008: From the Early Years to the State-of-the-Art*. Springer, 2009.

[26] D. C. Kozen, *The Design and Analysis of Algorithms*. Springer, 2012.

[27] A. Goldsmith, *Wireless Communications*. Cambridge Univ. Press, 2005.



ZICHENG XIA is currently pursuing the B.Eng. degree with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China. His research interest is device-to-device communication.



JIAWEI YAN is currently pursuing the B.Eng. degree with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China. His research interest is wireless multihop networks.

YUAN LIU (S'11–M'13) received the B.S. degree from the Hunan University of Science and Technology, Xiangtan, China, the M.S. degree from the Guangdong University of Technology, Guangzhou, China, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2006, 2009, and 2013, respectively, all in electronic engineering. Since 2013, he has been an Assistant Professor with the South China University of Technology, Guangzhou, China. His research interests include wireless communications and networking.

• • •