Social-Energy-Aware User Clustering for Content Sharing Based on D2D Multicast Communications

LIANXIN YANG, DAN WU, SHIMING XU, GUANGCHUN ZHANG, AND YUEMING CAI, (Senior Member, IEEE)
Institute of Communications Engineering, Army Engineering University of PLA, Nanjing 210007, China
Corresponding author: Dan Wu (wujing1958725@126.com)

This work was supported in part by the Natural Science Foundations of China under Grant 61671474 and in part by the Excellent Young Scholar Foundation of Jiangsu Province under Grant BK20170089.

ABSTRACT With the ever-increasing demands for content sharing, device-to-device (D2D) multicast content sharing is becoming a promising technology to improve the quality of local area services. In order to guarantee the implementation of D2D multicast content sharing, the main concern is the user clustering, i.e., cluster head (CH) selection and cluster formation. Most of the existing works fail to design a distributed scheme with a concern of the incentive to stimulate cooperation. In this paper, we model a novel user clustering problem with the target of maximizing the energy efficiency of the D2D multicast network, where both the social tie information and the pricing scheme are adopted to stimulate cooperation. Due to its NP-hard property, it is decomposed into two subproblems. Specifically, a CH selection algorithm based on social maximum weight is first proposed to discriminate the proper CHs from multitudinous candidates and restrict the upper bound of the number of the selected CHs. After that, we model the cluster formation process as a non-transferable utility coalition formation game, and a distributed coalition formation algorithm for cluster formation is proposed based on preference relationship and switch operations. Importantly, the final coalition structure is proved to be Nash stable. Numerical results show that our proposed scheme outperforms other three baseline schemes.

INDEX TERMS Content sharing, D2D multicast communication, CH selection, cluster formation, coalition formation game.

I. INTRODUCTION
Nowadays, the local area services, such as high-quality wireless video, become increasingly popular with the rise of mobile smart terminals, which leads to an explosion of mobile traffic [1]. In particular, D2D content sharing is recognized as an effective solution to improve the quality of local area services [2], [3]. Its core is that some users cache the target contents from the base station (BS), and then deliver them to other users through D2D communications, where the transmitters turn to be content providers (CPs) and the receivers are content requesters (CRs).

In general, D2D content sharing can be divided into two categories depending on different communication modes. One is D2D unicast content sharing [4], where one CP just serves one CR based on D2D unicast communication. The other is D2D multicast content sharing [5], [6], where one CP sends the same content to multiple CRs via D2D multicast communication. In this way, this CP and the CRs which it serves form a cluster, and this CP becomes the representative of the cluster, denoted as the cluster head (CH). Note that, there exist a portion of contents which are repeatedly requested and transmitted, which will trigger the majority of the total data traffic [1]. When multiple CRs request the same content, one of the most efficient ways is that the neighboring CP which has the target content provides these CRs with this content via D2D multicast communications. Fortunately, D2D multicast content sharing can reduce the redundancy transmissions of the BS for the same contents [7], [8], and the number of the users accessing the network can be greatly improved, making it possible to achieve the large-scale parallel implementation of content sharing. Accordingly, D2D multicast content sharing is gradually attracting more and more attention.

In a typical D2D multicast content sharing scenario, the main concern is the user clustering by determining which users are selected as the CHs and which users are grouped into
clusters [9]. That is to say, for the original users who require the target content, some users should first cache the target content from the BS and turn to be the CHs [10], and then the rest, denoted as the CRs, will decide which CH they prefer to form a cluster with. Obviously, the optimal approach for user clustering is to search exhaustively to find the optimal CHs and clusters. However, such a method leads to unbearable complexity with the user scale increasing. Thus, a distributed method should be designed. As such, how to determine the CHs and then how to form the corresponding clusters are of great importance to achieve the advantages mentioned above for D2D multicast content sharing [11].

Specifically, the CHs should firstly be selected from the initial users, denoted as the CH selection process. It is desired that the selected CHs are able to connect as many CRs as possible with high achievable rates while consuming as less energy as possible. Note that, the selected CHs serve as the bridge between the BS and some CRs which are provided with the requested services. Thus, the physical transmission conditions among the selected CHs, the BS, and the serving CRs should be taken into account when selecting the CHs. However, the physical transmission conditions between the selected CHs and the BS are always ignored, because the consumed resources (e.g., energy) of the BS for the content sharing are not considered in most existing works [6], [9], [12]. However, the consumed resources of the BS exert great impact on the performance of the D2D multicast network. Consequently, we should consider all factors comprehensively to select the proper CHs during the CH selection process.

Once the CHs are decided, the cluster formation process should be considered. The existing literatures mainly focus on the centralized cluster formation schemes. Reference [7] proposes a distance-based cluster formation algorithm. The authors in [12] model the cluster formation process as a Chinese restaurant process (CRP), and the authors in [13] formulate the cluster formation as a non-convex problem based on the outage capacity of the cluster. Note that, their common point is to need a central node to collect the global channel state information. In general, the task is undertaken by the BS, which increases its traffic. Moreover, the centralized scheme contradicts with the original intention of the distributed parallel implementation for D2D multicast content sharing. As such, a practical method is to develop a decentralized, self-organizing scheme to study the cluster formation process.

Importantly, coalition formation game can model the cooperation among communication nodes, and strike a balance between the benefits and costs generated from forming coalitions. Potential applications of coalition formation game in communication networks are diverse but limited, e.g., resource allocation in D2D communications [14], interference management in ultra-dense networks [15]. In particular, the essence of the above cluster formation process is that the CH and the CRs cooperate to share the target content. Such cooperation can bring the benefits (e.g., the improvement of the energy efficiency) and the costs (e.g., the reward that the CRs have to pay the CH), and we need to make a tradeoff between these benefits and costs to form the clusters. Thus, coalition formation game is very suitable to model the cluster formation process. However, it is impossible that these aforementioned results are directly extended to the cluster formation process. That is because what we care about in the cluster formation process is the performance of the CRs in the coalition, rather than all the players in this coalition. Hence, coalition formation game theory could open up a new avenue for modeling the cluster formation process in the D2D multicast content sharing.

In addition, the users who act as the CHs will undertake some computation tasks and transmit the target content to the CRs in the cluster, which increases the energy consumption of these users. Due to their energy scarce character, their energy consumption needs to be considered [7]. Indeed, it is desired for a D2D multicast content sharing scenario that the total energy consumption (e.g., the consumed energy of the selected CHs as well as the BS) can be minimized on the premise that all users can get the target content with data transmission rate as high as possible. Thus, the energy efficiency of the whole network is a much more practical metric to guide the D2D multicast content sharing. Moreover, some incentive should be made to motivate the CH to cooperate with the CRs [16]. Inspired by [17] and [18], the following two factors can be recognized as the incentive mechanisms: 1) social tie information. Actually, these users are not always altruistic or selfish, and at least, they are willing to share their cached content with family members, neighbors, colleagues, or friends in the vicinity. With this insight, the social tie information can be leveraged to inspirit the CHs by aligning their selfishness and selflessness, 2) pricing scheme. The CHs can ask the CRs for some “money” as the reward due to their contributions, and the charged “money” can be realized by pricing the transmit power that the CH spends.

Motivated by the above observations, we propose a novel user clustering scheme to maximize the energy efficiency of the D2D multicast network for content sharing. The main contributions are summarized as follows.

- We formulate a novel user clustering problem for a D2D multicast content sharing scenario, which aims to maximize the energy efficiency of the whole network. Due to its intractability, it is divided into two subproblems, i.e., CH selection and cluster formation. In particular, both the physical and social tie information between the CH and the CRs are considered to ensure the QoS of content sharing services and create incentives for cooperation.
- We propose a CH selection algorithm based on the social maximum weight (SMW) to select the CHs from the initial users. The proposed algorithm takes into account the knowledge of the practical physical transmission conditions between the selected CHs and the BS, as well as the upper bound of the number of the selected CHs.
The cluster formation process is modeled as a coalition formation game, where the CHs and the CRs act as the players. Then, we develop a distributed coalition formation algorithm based on preference relationship and switch operations. Moreover, we prove that it converges to a Nash-stable coalition structure by allowing the CRs to flexibly join and leave a coalition with a lower number of iterations and less information exchange among the players.

The remainder of the paper is organized as follows. The system model and the problem formulation are presented in Section II. In Section III, we design an algorithm based on SMW to select the CHs from the original users, and we model the cluster formation process as a non-transferable utility coalition formation game in Section IV. Section V presents some numerical results and discussions. Finally, Section VI concludes the key findings of the paper.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. SYSTEM MODEL

We focus on a typical D2D multicast content sharing scenario. Specifically, there exist $N$ original users which form the set $M = \{m_1, \ldots, m_i, \ldots, m_N\}$ in the network as shown in Figure 1, and they are interested in some contents. Moreover, we assume that neighbor discovery has already been achieved [19]. Importantly, we exploit the overlay D2D communication with the coexistence of cellular and dedicated modes. Then, we can utilize the existing resource management schemes, e.g., [20], to properly determine the type of D2D operation, schedule the separate frequency resources, and control their transmit power, so as to ensure that all the initial users have been allocated orthogonal in-band spectrum, and the QoS requirement constraint is satisfied. Besides, a low mobility scenario is adopted. Thus, it can be assumed that the channel conditions remain approximately constant during the user clustering process.

In general, all the original users are bound to get their target contents from the BS through cellular links. That is because the BS can afford any required content by connecting to the content servers in the Internet. However, such method will increase the resource consumption of the BS (e.g., energy, spectrum). To this end, some users are selected as the relays between the BS and the other users. That is, each of them firstly obtains the target content from the BS. Then, it caches this content, and further selectively sends this content to some other users which also want to acquire this one based on D2D multicast communications. In this way, this user which is selected to act as the relay and its receivers form a group to share the same content via D2D multicast communications.

Then, this group can be regarded as a cluster [7], [12], and all the selected users turn to be the CHs, which form the set $CH$. The remaining users, denoted as the CRs, form the set $CR$. $N_g$ is the number of the selected CHs (i.e., $|CH| = N_g$), and $N_r$ is the number of the remaining CRs (i.e., $|CR| = N_r$). Obviously, $M = CR \cup CH$, and $N = N_g + N_r$. Moreover, the set of the formed clusters is denoted by $Q$, and $N_c$ is the number of the clusters (i.e., $|Q| = N_c$). Note that $|\cdot|$ denotes the cardinality of a set. In this way, each CH can be regarded as a representative of the cluster which it is in.

Figure 1 presents an example of the selected CHs and the final formed clusters. Users $m_9$ and $m_{10}$ are selected as the relays between the BS and the other users, and then turn to be the CHs. Moreover, CH $m_9$ sends the target content to CRs $m_1$, $m_2$ and $m_3$ based on D2D multicast communications. Accordingly, they form a cluster. Similarly, CRs $m_5$, $m_6$, $m_7$ and $m_{10}$ construct a cluster with $m_{10}$ as a CH. It means that CH $m_{10}$ firstly obtains the target content from the BS and then delivers it to CRs $m_5$, $m_6$, and $m_7$ via D2D multicast communications. In addition, users $m_4$ and $m_8$ communicate directly with the BS, and obtain the target content from the BS. They may be initially considered as CHs, but cannot determine proper CRs to multicast the contents to them, or they do not enter into any cluster so as only to ask the BS for help.

In this way, how to determine the CHs and how to form the clusters are of great importance to achieve the advantages of D2D multicast content sharing. Firstly, the physical information, e.g., the channel quality, needs to be considered. Importantly, the user which is selected as the CH should own good-quality link between itself and the BS, so that the BS can deliver the target content to it effectively. Moreover, this CH should serve as many CRs as possible through D2D links with the lowest possible costs for its resources (e.g., energy). Secondly, the selected CHs are not totally altruistic or selfish. That is, on one hand, such sharing will consume their resources, which means that they would not willingly make contributions. How to stimulate an efficient cooperation is important for D2D multicast content sharing. On the other hand, they would rather share their contents with the friends and acquaintances than help the strangers, and we try to introduce the knowledge of social tie to encourage the initial user to act as a CH and to motivate the content sharing between the CHs and the CRs. Intuitively, the users often befriend others who have similar interests or perform similar actions. Hence, the knowledge of social tie is generally characterized as the social similarity [17]. In this way, the more similar the interest is, the more intimate they are, and then, the more the initial
users are willing to cooperate with each other by playing the role of CH and sharing contents based on D2D multicast communications. Consequently, the physical link and social tie information should be integrated so as to motivate the initial users to cooperate and improve the performance of the D2D multicast network.

B. ANALYSIS FOR THE UTILITIES

Each original user \( m_i \) is likely to play one of the following three roles in the network.

**Role 1:** User \( m_i \) is selected as a CH. For ease of description, it is denoted as CH \( g_k \), and the cluster which it belongs to is denoted as cluster \( \mathcal{S}_k \). Then, it will obtain the target content from the BS directly. Note that in order to save energy, the BS will adjust its transmit power to just meet the signal to noise ratio (SNR) threshold of CH \( g_k \). In this regard, the transmit power of the BS for CH \( g_k \) can be given by

\[
P_{B,g_k} = \frac{SNR_{th,g_k} N_0 B}{h_{B,g_k}},
\]

where \( SNR_{th,g_k} \) is the SNR threshold of CH \( g_k \), \( N_0 \) is the power spectral density of the additive white Gaussian noise (AWGN), \( B \) is the bandwidth of transmission channel, and \( h_{B,g_k} \) is the channel power gain between the BS and CH \( g_k \).

In this case, the achievable rate of CH \( g_k \) is defined as

\[
R_{B,g_k} = \log_2 (1 + SNR_{th}).
\]

Moreover, we define the utility of the initial users to measure the revenue as well as the cost in the D2D multicast communications. On one hand, CH \( g_k \) should spend some resources transmitting the target content to the CRs in the same cluster. On the other hand, the CRs in the cluster give some payment to CH \( g_k \) for his contribution, which will be explained later. Thus, the cost of CH \( g_k \) is considered to be zero, and the utility of CH \( g_k \) is just its achievable rate, that is

\[
u_{B,g_k} = R_{B,g_k}.
\]

**Role 2:** User \( m_i \) joins in cluster \( \mathcal{S}_k \) which has CH \( g_k \). We denote the set of the CRs in cluster \( \mathcal{S}_k \) as \( \mathcal{R}_k \), i.e., \( \mathcal{R}_k = \mathcal{S}_k \setminus \{g_k\} \), and \( |\mathcal{R}_k| = N_{R_k} \). CH \( g_k \) needs to guarantee the QoS of all the CRs in cluster \( \mathcal{S}_k \). Thus, the transmit power of CH \( g_k \) depends on the worst physical link in the cluster, and can be given by

\[
p_{g_k} = \max_{m_i \in \mathcal{R}_k} \frac{SNR_{th,m_i} N_0 B}{h_{g_k,m_i}},
\]

where \( SNR_{th,m_i} \) is the SNR threshold of CR \( m_i \) which is in cluster \( \mathcal{S}_k \), and \( h_{g_k,m_i} \) is the channel power gain between CH \( g_k \) and CR \( m_i \) in \( \mathcal{R}_k \).

Hence, the achievable rate of the link between CH \( g_k \) and CR \( m_i \) in \( \mathcal{R}_k \) can be given by

\[
R_{g_k,m_i} = \log_2 \left( 1 + \frac{p_{g_k} h_{g_k,m_i}}{B N_0} \right).
\]

In addition, we leverage the knowledge of social tie information in terms of the social similarity [17] to promote the efficient cooperation between the selected CH \( g_k \) and CR \( m_i \). Its essence is to calculate the social tie strength based on the interest similarity between the selected CH \( g_k \) and CR \( m_i \) so as to make the users with strong social similarity be able to cooperate with each other. Specifically, the BS defines the interest list as the \( M \)-dimensional space, and broadcasts this information to the original users in the cell. Here, \( M \) is the total number of some specific keywords, such as “music”, “sport”, etc. According to the interest list, \( g_k \) and \( m_i \) determine their own interest profiles to show the normalized degree of interest, denoted by \( I_{g_k}, I_{m_i} \in [0, 1]^M \), respectively. When \( m_i \) sends a request to \( g_k \) with its own interest profile, \( g_k \) calculates the social similarity with \( m_i \) based on its own interest profile. Inspired by [21], we adopt the cosine similarity to calculate the social similarity. That is,

\[
e_{g_k,m_i} = \cos \left( \angle I_{g_k} I_{m_i} \right) = \frac{I_{g_k} I_{m_i}}{|I_{g_k}| |I_{m_i}|},
\]

where \( | \cdot | \) is the length of a vector. In fact, \( M \) determines the accuracy of the calculated social similarity. The larger \( M \) is, the more accurate the calculated social similarity will be.

Furthermore, CR \( m_i \in \mathcal{R}_k \) needs to pay CH \( g_k \) for his contribution, which can also be regarded as the incentive for CH \( g_k \). Inspired by [18], such payment can be characterized by pricing the power that CH \( g_k \) needs to consume. In particular, due to the different social similarities among users, the prices need to be distinguished [22]. Intuitively, the larger the social similarity between CH \( g_k \) and CR \( m_i \), the more it is reinforced in cooperation, and then, the lower unit power price CR \( m_i \) should pay CH \( g_k \). Thus, the unit power price that CR \( m_i \) needs to pay CH \( g_k \), is defined as

\[
\lambda_{g_k,m_i} = C \frac{1}{e_{g_k,m_i}},
\]

where \( C \) is the basic unit price which is determined in advance by the BS. In essence, it is a constant to show the reserved bid of consumed power from the perspective of the whole efficiency of resource management. Hence, the utility of CR \( m_i \) in cluster \( \mathcal{S}_k \) which has CH \( g_k \) is defined as

\[
u_{g_k,m_i} = R_{g_k,m_i} - \lambda_{g_k,m_i} p_{g_k}.
\]

**Role 3:** User \( m_i \) is neither selected as a CH, nor joins in any cluster. In this case, it needs to obtain the target content from the BS directly. Similar to **Role 1**, the transmit power of the BS for user \( m_i \) can be given by

\[
p_{B,m_i} = \frac{SNR_{th,m_i} N_0 B}{h_{B,m_i}},
\]

where \( SNR_{th,m_i} \) is the SNR threshold of CR \( m_i \), and \( h_{B,m_i} \) is the channel power gain between the BS and user \( m_i \).

The achievable rate of user \( m_i \) in this case is defined as

\[
R_{B,m_i} = \log_2 (1 + SNR_{th}).
\]
Similar to Case 1, the utility of user $m_i$ is defined as

$$u_{B,m_i} = R_{B,m_i}.$$  

(11)

Note that the defined utility measures the revenue as well as the cost. In Role 1 and Role 3, user $m_i$ gets the target content from the BS directly, the cost is regarded as 0, and his utility is just the achievable rate. As for Role 2, CR $m_i$ needs to pay CH $g_k$ for his contribution, and his utility needs to subtract the cost as shown in (8).

**C. PROBLEM FORMULATION**

With the goal of improving the performance of D2D multicast content sharing, we need to select the CHs from the initial users and determine the CRs with which each CH jointly form a cluster. Based on the analysis above, we define a matrix $x = (x)_{(N+1)\times N}$ to denote the transmission relationship among the users in the D2D multicast network. That is,

$$x = \begin{pmatrix} x_{B,m_1}, \ldots, x_{B,m_N} \\ x_{m_1,m_1}, \ldots, x_{m_1,m_N} \\ \vdots \\ x_{m_N,m_1}, \ldots, x_{m_N,m_N} \end{pmatrix},$$

(12)

where its element $x$ is binary, where $x_{B,m_i} = 1$ if user $m_i$ sets the target content from the BS directly, otherwise, $x_{B,m_i} = 0$, and $x_{m_i,m_j} = 1$ if user $m_i$ delivers the target content to user $m_j$, otherwise $x_{m_i,m_j} = 0$. When the binary variable matrix $x$ is determined, the CH and the final formed cluster structure and the transmission relationship are both decided. For example, if the number of the initial users is 4, and

$$x = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

(13)

it illustrates that user $m_1$ forms a cluster with users $m_2$ and $m_3$. user $m_1$ gets the target content from the BS, and then multicasts it to users $m_2$ and $m_3$. As for user $m_4$, it gets the target content from the BS directly. Moreover, our objective of CH selection and cluster formation is to maximize the ratio of the total utilities of all the users to the total power consumption for the whole network. In fact, this ratio shows the utility value unit power. Importantly, the essence of this ratio can be regarded as the energy efficiency for the whole network. Thus, the optimization problem can be formulated as

$$\max_x \sum_{i=1}^N x_{B,m_i} u_{B,m_i} + \sum_{i=1}^N \sum_{j=1}^N x_{m_i,m_j} u_{m_i,m_j}$$

$$s.t. \quad 1 - x_{B,m_i} \sum_{j=1}^N x_{m_i,m_j} \leq 0, \quad \forall m_i,$$

(14c)

$$y_{m_i} p_{m_i} \leq p_{\max}, \quad \forall m_i,$$

(14d)

$$x_{m_i,m_j} d_{m_i,m_j} \leq d_1, \quad \forall m_i, m_j,$$

(14e)

$$y_{m_i} d_{m_i} \leq d_2, \quad \forall m_i,$$

(14f)

where $y_{m_i} = 1$ if $\sum x_{m_i,m_j} \geq 1$, otherwise $y_{m_i} = 0$.

Here, the constraint in (14b) ensures that one user can get the target content at last, and it can either join in at most one cluster or get the target content from the BS. The constraint in (14c) restricts that only the user who is selected as the CH can deliver the target content to other CRs, and the CR who gets the target content from the CH is forbidden to deliver the target content to other CRs. The constraint in (14d) ensures the transmit power threshold for these devices, where $p_{\max}$ denotes the maximum power that a device can afford. The constraint in (14e) ensures that the proximity between user $m_i$ and user $m_j$ is no larger than the proximity threshold, denoted as $d_1$, for establishing a D2D link. Moreover, the constraint in (14f) guarantees that the BS can deliver the target content to the selected CHs effectively, where $d_2$ denotes the proximity threshold between the BS and user $m_i$.

In essence, the optimization problem in (14) is a non-linear 0-1 programming problem. As shown in [23], it is typically NP-hard, which usually leads to computational intractability. Moreover, the centralized approach to find the optimal solution usually results in a lot of pressure on the control center, e.g., the BS. It can be calculated that the total number of the possible clusters in (14) is $N \cdot (2^{N-1} - 1)$, which increases with $N$ exponentially. Thus, in practical terms, it is more preferable to search some algorithms to strike a balance between the solution performance and the implementation complexity.

In fact, the optimization problem in (14) involves two issues. Firstly, the CHs should be selected from the initial users with the metric which is related to not only the channel quality with the BS, but also the physical connections and social ties with some other users. In this way, each of the selected CH can act as the representative of a cluster, and some other CRs which require the same content as this CH make decisions on whether they join in this cluster or not, until a stable cluster is formed. Once the cluster is formed, this CH will multicast the target content to the CRs in this cluster. Note that, some original users may be initially selected as the CHs, but there exists no CR acquiring the target content from them. In this context, they just communicate with the BS. Importantly, we still regard them as the CHs, and form a singleton cluster which only involves itself. Moreover, there exist some CRs which do not enter into any cluster. In this context, they have to ask the BS for help. As a result, the optimization problem in (14) can be transformed into two sub-problems. The first one is the CH selection process, and the second one is the cluster formation process.
III. CH SELECTION PROCESS BASED ON SOCIAL MAXIMUM WEIGHT

We always hope that the selected CHs are likely to make the utilities of the CRs in the clusters as large as possible while consuming as less energy as possible. The calculation for the utilities of the CRs in the clusters considers both physical link and social tie information so that we provide a CH selection algorithm based on social maximum weight (SMW) to discriminate the proper CHs from the initial users \( M = \{m_1, \ldots, m_i, \ldots, m_N\} \). Its core is to define and calculate the weight for every user, and then select the users with large weight as the CHs.

In order to make full use of the proximity gain of D2D links, the selected CHs are expected to be connected with as many CRs as possible through D2D links. Based on the constraint in (14d), we firstly define a binary variable \( z_{m_i, m_j} \) to indicate whether user \( m_i \) is able to transmit the target content to user \( m_j \), that is
\[
z_{m_i, m_j} = \begin{cases} 1 & \text{if } d_{m_i, m_j} \leq d_1 \\ 0 & \text{if } d_{m_i, m_j} > d_1. \end{cases}
\]  

(15)

Secondly, based on (14e) which guarantees that the BS can deliver the target content to the selected CHs effectively, we define the distance factor \( DF_{m_i} \) for user \( m_i \) as
\[
DF_{m_i} = \frac{1}{d_{m_i}}.
\]  

(16)

in which \( d_{m_i} \) is the distance between the BS and user \( m_i \). The larger the value of \( d_{m_i} \), the closer user \( m_i \) stays from the BS. Then, the transmit power of the BS can be decreased and its total energy consumption can be reduced.

Thirdly, we also introduce the social tie based on the social similarity in (6) to stimulate the cooperation between the selected CH and the CRs. The underlying rationale lies in that the selected CH, which has the strong social similarity with as many CRs as possible, is likely to serve as many CRs as possible. It implies that the social similarity should be added into the weights for these users. Taking into all the parameters into consideration, the social weight (SW) \( w_{m_i} \) of user \( m_i \) is defined as
\[
w_{m_i} = DF_{m_i} \sum_{m_j \in M \setminus m_i} z_{m_i, m_j} \varepsilon_{m_i, m_j},
\]  

(17)

where the users with large SWs are likely to be selected as the CHs.

After the CHs are decided, the next step is to study how the CRs form clusters with them. Obviously, one optimal approach is to search all the possible clusters exhaustively and find the one with the highest energy efficiency. However, the method cannot work well due to its complexity, especially in a large-scale network. Hence, its necessary to find a distributed cluster formation framework to strike a balance between the solution performance and the implementation complexity.

When a CR joins in a cluster, it is likely to bring the benefits from the following three aspects: \( i) \) the achievable rate of the CR can be improved due to the good channel quality within short distance, \( ii) \) the traffic of the BS can be efficiently offloaded, \( iii) \) for both the CH and the BS, the D2D multicast content sharing can save more energy than that based on D2D unicast or cellular communication. Following the entrance into this cluster, this CR is natural to pay for the devotion of the CH in this cluster. However, it cannot always succeed in joining in any satisfying cluster. That is due to the following facts: once a cluster forms, the CH serves all the CRs in the same cluster simultaneously with a guarantee that the QoS of the CR with the worst channel in the cluster should be satisfied. When the channel quality of the CR who is going to join in the cluster is worse than the channel qualities of the CRs who have already existed in the cluster, the CH needs to raise its power to satisfy the QoS of the CR. In this way, the CRs who have already existed in the cluster need to pay the CH more, and their utilities may be decreased after this
Coalition formation game (CFG) [26], [27] involves a set of players who want to cooperate by forming coalitions in order to improve the performance of the individual user and the whole network.

Here, we model the cluster formation process as a coalition formation game, which is defined by a triple \((M, Q, v)\), and the formulation of this game is as follows:

- **Players**. \(M\) is the set of the original users, which is the game players.
- **Coalition structure**. A coalition structure (coalition partition) is defined as the set \(Q = \{ S_1, \ldots, S_i, \ldots, S_L \}\) where \(S_i\) is disjoint coalition verifying \(\bigcup_{i=1}^{L} S_i = M\). The coalition structure is initialized with each player as a singleton coalition, denoted as \(Q_0 = \{ Q_1, Q_2 \}\), where \(Q_1\) is the partition of \(CR\), and \(Q_2\) is the partition of \(CH\).
- **Coalition Value**. \(v(S_k)\) denotes the coalition value of coalition \(S_k\). Our defined game is with non-transferable utility (NTU) where \(v(S_k)\) is a set of payoff vectors, where each element of the vector represents the utility of the players in coalition \(S_k\) [24], e.g. \(v_m(S_k)\). Since the coalition value depends on the members in the coalition, the proposed coalition formation game here is in characteristic form [24], [25].

A transfer of a CR from a coalition in \(Q_1\) to a coalition in \(Q_2\) is called a switch operation:

**Definition 2**: A switch operation \(\sigma_{m,k}(m_i)\) is defined as the transfer of player \(m_i\) from \(S_m \in Q_1\) to \(S_k \in Q_2\), \(\sigma_{m,k}(m_i)\): \(S_m \mapsto S_m \backslash m_i\), and \(S_k \mapsto S_k \cup m_i\).

To judge whether a switch operation can improve the performance of the coalition structure or not, we will define the gain of the switch operation as follows, which is defined based on the utility of singleton CR.

**Definition 3**: The gain \(g(\sigma_{m,k}(m_i))\) of the switch operation associated with \(\sigma_{m,k}(m_i)\) is defined through the following two cases.

**Case 1**: When \(m_i\) escapes from \(S_m\) and joins in \(S_k\) which has \(CH\ g_k\) and the channel quality between \(m_i\) and \(g_k\) is not worse than the existing channel qualities in \(S_k\), \(CH\ g_k\) does not raise its transmit power, and the utilities of the CRs who have already existed in \(S_k\) don’t change. What we need to care about is the utility of \(m_i\) after it joins in \(S_k\). Based on (8), the utility of \(m_i\) when it gets the target content from \(g_k\) directly is given as

\[
v_{m_i}(S_k \cup m_i) = u_{g_k,m_i} = \log_2 \left( 1 + \frac{p_{g_k} h_{g_k,m_i}}{B N_0} \right) - \lambda_{g_k,m_i} p_{g_k}.
\]

The gain \(g(\sigma_{m,k}(m_i))\) of the switch operation \(\sigma_{m,k}(m_i)\) is defined as

\[
g(\sigma_{m,k}(m_i)) = v_{m_i}(S_k \cup m_i).
\]

**Case 2**: When \(m_i\) escapes from \(S_m\) and joins in \(S_k\) and the channel quality between \(m_i\) and \(CH\ g_k\) is worse than the worst channel quality in \(S_k\), \(CH\ g_k\) needs to raise its transmit power to satisfy the QoS of \(m_i\). We denote \(p'_{g_k}\) as the increased transmit power of \(CH\ g_k\) after \(m_i\) joins in \(S_k\). If \(p'_{g_k} > p_{\text{max}}\), which \(CH\ g_k\) is not able to afford, \(g(\sigma_{m,k}(m_i))\) is defined as \(-\infty\). Otherwise, we should consider the utilities of both the CRs who have already existed in \(S_k\) and \(m_i\) when defining the gain.

Based on (8), the utility of CR \(m_j \in R_k = S_k \setminus \{g_k\}\) after \(m_i\) joins in \(S_k\) turns to be

\[
v_{m_j}(S_k \cup m_i) = u_{g_k,m_i} = \log_2 \left( 1 + \frac{p'_{g_k} h_{g_k,m_i}}{B N_0} \right) - \lambda_{g_k,m_i} p'_{g_k}.
\]

where \(h_{g_k,m_i}\) is the channel power gain between \(CH\ g_k\) and \(CR\ m_j\), and \(\lambda_{g_k,m_i}\) is the unit power price that \(CR\ m_j\) needs to pay \(CH\ g_k\).

The utility of \(CR\ m_i\) after it joins in \(S_k\) is given by

\[
v_{m_i}(S_k \cup m_i) = u_{g_k,m_i} = \log_2 \left( 1 + \frac{p'_{g_k} h_{g_k,m_i}}{B N_0} \right) - \lambda_{g_k,m_i} p'_{g_k}.
\]

**Remark**: Firstly, the utility of the CH in a coalition is not defined. That is because the CH is thought to be always willing to join in the coalition due to the incentive mechanism.

Secondly, the energy efficiency of the whole network can be improved based on the definition in (23). When a CR requests to join in the coalition, the utilities of the CRs who have already been in the coalition are not impaired, and the entrance of the CR also contributes to the utilities of the whole network. More importantly, the transmit power of D2D link...
Algorithm 2 The Cluster Formation Algorithm Based on Coalition Formation Algorithm in CFG

1. Initialize the coalition structure as \( Q_0 \) and the indicator vector \( k = (k)_1 \times N_p \), whose element is set to zero.
2. Set the current coalition structure as \( Q_0 \rightarrow Q_{\text{cur}} = \{ Q_{\text{cur},1}, Q_{\text{cur},2} \} \), where \( Q_{\text{cur},1} = Q_1 \) and \( Q_{\text{cur},2} = Q_2 \).
3. Repeat:
   4. Randomly choose a coalition \( S_k \in Q_{\text{cur},2} \), whose corresponding indicator \( k \) is increased by 1.
   5. Set \( U = \emptyset \).
   6. For each \( S_m \in Q_{\text{cur},1} \) do:
      7. if \( g(\sigma_{m,k}(m_i)) \geq 0 \) then
         8. \( U = U \cup \sigma_{m,k}(m_i) \).
      9. end if
   10. end for
   11. if \( U \neq \emptyset \) then
      12. Find \( m^* \) such that \( \sigma_{m,k}(m^*) > \sigma_{m',k}(m_i), \forall \sigma_{m',k}(m_i) \in U \).
      13. CR \( m^* \) in \( S_{m^*} \) leaves \( S_{m^*} \) and joins in \( S_k \).
      14. Update the current coalition structure as follows:
         \( Q_{\text{cur},1} = Q_{\text{cur},1} \setminus S_m \), \( Q_{\text{cur},2} = (Q_{\text{cur},2} \setminus S_k) \cup S_k \), where \( S_k = S_k \cup S_m \).
      15. end if
   16. Until \( Q_1 = \emptyset \) or all the elements in \( k \) are up to \( n^{up} \), where \( n^{up} \) is a prescribed large number.

In Algorithm 2, the coalition which CR \( m^* \) decides to participate in is chosen randomly in Step 4. In order to judge switch operation \( \sigma_{m,k}(m_i) \) operates or not, the system only needs to know the channel state information in coalition \( S_k \) and CR \( m_i \), which can be obtained from the BS. These switch operations in Algorithm 2 only make use of the information of local coalitions, which does not lead to additional overhead. The CHs are important to implement Algorithm 2. Thus, the incentive for them is necessary. Note that the proposed algorithm allows the CR flexibly to join or leave the coalition. However, the CR who has already joined in a coalition will not leave the coalition, otherwise its utility will be impaired.

C. PROPERTY ANALYSIS

Proposition 1: For a fixed network topology and starting from the initial coalition structure \( Q_0 \), the proposed cluster formation algorithm based on coalition formation algorithm maps to a sequence of switch operations which converges to a final coalition structure \( Q \) in a finite iterations.

Proof: There exist two cases when switch operations happens. One is that switch operations are with strictly positive gain, and the other is that switch operations are with zeros gain. In the first case, after any switch operation, the energy efficiency strictly increases, meaning that the same coalition structure can never be visited twice. Furthermore, since there are finite coalition structures, Algorithm 2 can always converge to a final coalition structure \( Q \) after a finite number of iterations. It is noted that when a switch operation happens, the energy efficiency of the whole network increases. Nevertheless, due to the fact that we choose the coalition where a CH exists in a random manner in Step 4 in Algorithm 2, the process for seeking the switch operation does need numerous iterations. To guarantee the convergence of Algorithm 2, we set a prescribed upper bound for CHs. If the number of CHs being chosen is up to the upper bound, it means that the CHs would not absorb any CR to join in the cluster. In the second case, some switch operations are with zero gains, which does not contribute to the performance. However, the prescribed upper bound \( n^{up} \) guarantees the convergence of Algorithm 2. In summary, Algorithm 2 tends to coverage to the final coalition partition through a finite iterations.

Next, we investigate the stability of the resulting coalition structure by using the mainstream stability notions [25], [26] from the coalition formation game theory.

Definition 5: A coalition structure \( Q := \{ S_1, \ldots, S_m, \ldots, S_M \} \) is Nash-stable if \( \forall S_m \in Q_1, \forall S_k \in Q_2, g(\sigma_{m,k}(m_i)) < 0 \).

Proposition 2: Starting from the initial coalition structure, the proposed cluster formation algorithm based on coalition formation algorithm will always converge to a Nash-stable network partition.

Proof: The proof includes two cases. The first case is that \( Q_1 = \emptyset \), which means that all CRs have joined in the
coalitions through our defined switch operations. In this case, there exists no switch operation for these CRs. The second case is that $Q_1 \neq \emptyset$, which means that there still exist some CRs who are not able to join in coalitions. In this case, suppose that the final coalition structure obtained from Algorithm 2 is not Nash-stable. Then, there exists a coalition $S_m$ in $Q_1$ and a coalition $S_k$ in $Q_2$ such that $g(\sigma_{m,k}(m_i)) \geq 0$. Based on our algorithm, the CR in coalition $S_m$ can join in coalition $S_k$, which contradicts the fact that we have gotten the final coalition structure. Thus, the final coalition structure resulting from Algorithm 2 must be Nash-stable.

The complexity of Algorithm 2 depends on the switch operations. In the worst case, the complexity is $O(N_g \cdot N_q)$, which is much lower than the exhaustive search method as discussed before.

V. NUMERICAL RESULTS

In this section, we present and analyze the simulation results to verify the performance of our proposed algorithms. Our simulations are carried out in a single cell scenario of a 200m $\times$ 200m area with one BS located in the center and some original users. Inspired by [28], to distinguish the distribution of the original CRs, the probability density function $\lambda(q)$ at the point $q$ is set as the sum of four Gaussian functions of the form $5\exp\left(-0.05 ((x-x_{cen})^2+(y-y_{cen})^2)\right)$, in which the centers $(x_{cen}, y_{cen})$ are $(50, 50)$, $(50, -50)$, $(-50, 50)$, and $(-50, -50)$. The bandwidth for each transmission is set as 1 Hz. The power spectral density of AWGN is $10^{-5}$ W/Hz [29], and the path loss exponent for cellular link is set 3 and D2D link 2. The maximum transmit power for a mobile device is set as 23 dBm [30].

We give a specific scenario to illustrate the formed clusters and the importance to consider both the physical link and social tie information when $d_1 = 50$ m and $d_2 = 40$ m. When the initial users are distributed as Figure 2(a), the selected CHs are users $m_5$ and $m_8$. Moreover, CH $m_5$ and CRs $m_1$, $m_2$, and $m_3$ form a cluster, CH $m_8$ and CR $m_9$ form a cluster. Other CRs either stay too far away from the selected CHs or own weak social similarity with the selected CHs. Thus, they are all not able to join in the clusters. Note that when CR $m_3$ moves to the location as shown in Figure 2(b), CR $m_3$ is not able to form cluster with CR $m_5$. This is because, on one hand, the distance between CR $m_3$ and CH $m_5$ exceeds the threshold for establishing a D2D link. On the other hand, in order to maintain the level of the QoS of $m_3$, $m_5$ needs to increase its transmit power, which will lead to a larger cost for CRs $m_1$ and $m_3$, and impair their utilities. In addition, when the distance is fixed, if the social similarity between CR $m_3$ and CR $m_5$ becomes weak as shown in Figure 2(c), CR $m_3$ is still not able to form cluster with CR $m_5$. This is because weak social similarity leads to high price for CR $m_3$, the utility of CR $m_3$ will be negative if it joins in the cluster.

In addition, the performance of our proposed D2D multicast scheme is closely related to the CH selection process and cluster formation process. Thus, the effects of detailed CH selection method and cluster formation method are respectively analyzed. For convenience, we denote our proposed algorithm as SMW+CFG algorithm in the following context. Firstly, the performance of CH selection and cluster formation is analyzed with respect to the energy efficiency of the whole network, the number of CHs $N_c$ ($N_1$) and the number of D2D multicast clusters $N_m$ ($N_2$) (Here, the cluster

FIGURE 2. A specific scenario of the clusters. (a) the originally formed clusters, (b) the updated clusters when the physical link information changes, (c) the updated clusters when the social tie information changes.
refers to the cluster with more than one player in order to compared with $N_1$), as well as the averaged size of D2D multicast clusters, which is denoted as the average number of members in the coalitions, and the total numbers of the CRs in and out of the formed clusters ($N_3, N_4$), considering the impacts of $N$, D2D physical distance proximity threshold $d_1$ and the physical distance proximity threshold $d_2$. Secondly, we evaluate the performance of our proposed SMW+CFG algorithm by comparing it with three algorithms, denoted as MW+CFG, SMW+Physical and MW+Physical, respectively. In the MW algorithm, social tie information is not considered in the CH selection process, and in the Physical algorithm, social tie information is not considered in the cluster formation process. Three aspects of performance are analyzed including energy efficiency of the whole network, the averaged size of D2D multicast clusters and the number of the CRs in the clusters. All the results are averaged over 100 random realizations.

In Figure 3, we evaluate the energy efficiency of the network for our proposed cluster formation schemes when $N_{th}$ is 8 under different physical proximity thresholds. In Figure 3(a), we can see that the energy efficiency decreases with $N$ increasing. Because the upper bound of the CHs is fixed, the number of the CRs who are not able to join in clusters increases with $N$ increasing, as will be shown in Figure 4(c). As a result, the BS should spend its power to

![Figure 3](image-url)

**FIGURE 3.** The energy efficiency of the network for different user scales. (a) the energy efficiency in our proposed schemes versus user scales, (b) the comparison of the energy efficiency between our proposed schemes and the cellular scheme versus user scales.

![Figure 4](image-url)

**FIGURE 4.** The performance of our proposed schemes. (a) the number of CHs or clusters versus user scales, (b) the averaged size of formed clusters versus user scales, (c) the number of the users in or out clusters versus user scales, (d) the ratio of the users in clusters versus user scales.
transmit the target content to them respectively, which leads to the high power consumption and low energy efficiency. Moreover, the curves for $d_2 = 60$ m are above the curves for $d_2 = 40$ m, and the gap decreases with $N$ increasing. Because larger $d_2$ makes the selected CHs distribute much sparser in the network, which makes more CRs join in the clusters. However, with $N$ increasing, there always exist the CRs who are able to join in the clusters no matter what $d_2$ is. Thus, the advantage of $d_2 = 60$ m gradually weakens. At the same time, for a certain $d_2$ and user scale, the curve for larger $d_1$ is above the curve for smaller $d_1$. Because larger $d_1$ makes more clusters, and more original users are grouped into clusters, as is shown in Figure 4(c). Thus, the energy efficiency of the network is larger. In Figure 3(b), the curves for our proposed schemes are above the curves for cellular mode, which denotes the advantages of the cluster formation scheme.

In Figure 4, we evaluate the performance of our proposed schemes on the number of selected CHs and formed clusters, the averaged size of the formed clusters and the number of the users in and out clusters as well as the ratio of the CRs in the clusters. In Figure 4(a), firstly, $N_1$ is less than 8 with small $N$, but equals 8 with $N$ increasing. Because more original CRs meet the proximity threshold to become a CH with large $N$. Meanwhile, $d_2$ influences $N_1$ a lot, but $d_1$ a little. Secondly, $N_2$ is less than $N_1$, which is consistent with practical scenario where not all selected CHs are able to form clusters with CRs. Thirdly, with $N$ increasing, $N_2$ will stay almost unchanged due to the fixed $N_{th}$. Moreover, $d_1$ and $d_2$ both influence $N_2$, but the effect of $d_1$ is much stronger. This is because $d_1$ influences the establishment of D2D links, and $d_2$ mainly influences the selection of CHs. In Figure 4(b), the averaged size of formed clusters increases with $N$ increasing, and $d_2$ has little effect on the averaged cluster size. However, $d_1$ exerts great impact on the averaged size. In general, larger $d_1$ corresponds to larger size of formed clusters. However, we can see from Figure 4(b), the average size when $d_1 = 50$ m is larger than that when $d_1 = 70$ m. This is because when $d_1 = 70$ m, the number of formed clusters is larger than that when $d_1 = 50$ m as shown in Figure 4(a). Thus, even if the total number of users in clusters when $d_1 = 70$ m is larger as shown in Figure 4(c), the average size of formed clusters is still smaller when $d_1 = 70$ m. In Figure 4(d), the ratio of the CRs in clusters goes almost no changed. Because with $N$ increasing, the number of the CRs in the clusters also increases, which leads to almost unchanged ratio.

Figure 5 illustrates how the upper bound of the number of the CHs affects the energy efficiency of the network in the cases of different $d_1$ and $d_2$ when $N = 24$. We can see that when $N_{th}$ is small, with $N_{th}$ increasing, the energy efficiency increases. Because more original users are able to join in clusters when $N_{th}$ increasing. However, the energy efficiency decreases and stays almost unchanged when $N_{th}$ increases to a degree. On one hand, the number of the selected CHs increases with $N_{th}$ increasing, and the BS needs to spend much more energy to deliver the target content to these CHs. On the other hand, the original CRs who satisfy the proximity threshold of the CHs are fixed when $N = 24$. Therefore, even though $N_{th}$ increases, the number of selected CHs almost has no change, which leads to the almost unchanged energy efficiency. Moreover, under different proximity thresholds, the values of $N_{th}$ with maximum energy efficiency are various. Because larger proximity thresholds make more CRs join in clusters with small number of selected CHs. As a result, the advantage of forming clusters in this case is made full use of earlier, and the corresponding value of $N_{th}$ is smaller.

In Figure 6, the energy efficiency of the whole network under four cluster formation algorithms is compared for different user scales when $d_1 = 50$ m and $d_2 = 40$ m. We can see that the energy efficiency under SMW+CFG and MW+CFG algorithms is larger than SMW+Physical and MW+Physical algorithms. Because CFG process not only considers the proximity threshold for establishing D2D links, but also the gain for the whole network when a CR tries to join in clusters. Also, the energy efficiency under our proposed SMW+CFG...
The average size of formed clusters is larger under SMW+Physical and MW+Physical algorithms than that under SMW+Cfg and MW+Cfg algorithms. Because only the proximity constraint is considered in Physical algorithm, and it is much easier for the CRs to meet such constraint and join in clusters. However, large size brings challenge to the synchronization of formed clusters. In addition, Figure 7(b) compares the number of the CRs in the clusters under different cluster formation algorithms. The number of the CRs under SMW+Physical and MW+Physical algorithms is slightly larger than that under SMW+Cfg and MW+Cfg algorithms, due to the same reason as the smaller number of formed clusters. However, even if more CRs are grouped into clusters under SMW+Physical and MW+Physical algorithms, the energy efficiency is still smaller as shown in Figure 6. Because the utilities of the CRs are likely to be negative without the consideration of the gain for the whole network in Physical algorithm.

VI. CONCLUSIONS

In this paper, we have investigated the user clustering in a D2D multicast content sharing scenario, with the target of maximizing the energy efficiency of the network, while taking into consideration some incentive mechanisms to stimulate the cooperation. We firstly propose a CH selection algorithm based on social maximum weight to select the CHs from the original users. After that, the cluster formation process is formulated as a non-transferable utility coalition formation game. The CH and the CRs in the same coalition form one cluster, and the CH delivers the target content to the CRs within the coalition via D2D multicast communications. Subsequently, a coalition formation algorithm is proposed based on the preference relationship and the switch operations to acquire the final coalition structure. Moreover, we characterize the proposed game model and show that the proposed algorithm converges to the coalition structure that is Nash-stable. Simulation results have demonstrated that the proposed scheme can improve the energy efficiency of the whole network significantly compared to other existing schemes.

REFERENCES

L. Yang et al.: Social-Energy-Aware User Clustering for Content Sharing


