IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 13, 2017, accepted November 9, 2017, date of publication November 22, 2017,
date of current version December 22, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2776289

Resource Allocation Schemes for Revenue Maximization in Multicast D2D Networks

JAGADEESHA R. BHAT[1], (Student Member, IEEE), JANG-PING SHEU[1,2], (Fellow, IEEE), AND WING-KAI HON[1]

[1]Department of Computer Science, National Tsing Hua University, Hsinchu 30013, Taiwan
[2]Institute of Communication Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan

Corresponding author: Jang-Ping Sheu (sheujp@cs.nthu.edu.tw)

ABSTRACT Mobile users' service satisfaction plays a vital role in improving the revenue of telecom companies. When user data demands are uniform, satisfying the user requests is much easier than when demands are diverse. Yet, in a multicast scenario, each user may wish to view the same multicast data at different bit rates based on their own device capacities or resolutions. In this paper, we consider device-to-device (D2D) multicast users who may demand the multicast data at various rates, and in return, they offer different profits (revenue) to the telecom operator. Moreover, users may have different channel qualities from the base station, which will also affect the data rates. We show that satisfying the user requests to maximize the profit becomes NP-hard when the resource blocks are limited, and propose a greedy heuristic and two approximation algorithms to solve this problem. Besides, we consider an alternative objective of maximizing the number of satisfied users and propose a greedy heuristic algorithm for this variant. Our simulation results demonstrate that the proposed algorithms offer higher profit, throughput, and satisfy more users than the other candidate algorithms.

INDEX TERMS Approximation algorithm, device-to-device multicast, profit, satisfied user, service satisfaction.

I. INTRODUCTION

Recently, due to the unprecedented rise in mobile generated traffic, there has been a tremendous demand for spectrum resources. The recent survey by Cisco has anticipated that by 2021, with the introduction of the fifth-generation (5G) mobile networks, there will be 24.5 million mobile devices [1]. In addition, by the next decade, mobile traffic will increase by 100 folds [2]. The recent progress in the device-to-device (D2D) communications has brought some hope to meet this enormous spectrum demand [3]. In a D2D scenario, users reuse the cellular user's (CU's) spectrum for data transmission and thus conserve the resource blocks (RBs) [4]. Here, the resource block (RB) corresponds to the time and frequency component of an orthogonal frequency division multiple access (OFDMA) subframe. As two nearby devices can communicate directly, the burden on the base station (e-NB) can be greatly reduced. However, in practice, there is interference challenge between the e-NB, the CUs, and D2D transmitter-receiver pair [4]. In an underlay based traditional D2D communication, the e-NB serves the CUs in the downlink by allocating RBs. The CUs utilize the uplink frequencies for their data uplink while the D2D user (DUs) reuse the same uplink frequencies to communicate with other DUs. In such a scenario, interference between the users will be the main issue to be addressed. The interference in the said underlay model can be mitigated by suitable power regulation, spectrum and efficient radio resource management schemes [5]–[7].

Nowadays, multimedia broadcast multicast service (e-MBMS) based video service has attained much more attention, as it can efficiently support many users located in a small geographical region at a high data rate [8], [9]. As an example, the e-MBMS in conjunction with D2D will be implemented in 2020 Olympics, to provide live video broadcast to thousands of viewers in the stadium [10]. In this scenario, we can imagine that there will be a need for transmission and resource allocation by considering user's video quality demand. Few subscribers with lower tariff plans and value added services might request the data at lower rates based on their device capacity or resolution. On the other hand, users with better tariff plans may demand the data at a higher rate. In this situation, satisfying a user request also depends on their channel quality (CQI) and RB requirements. Therefore, the operator needs to make a wise decision in choosing the correct set of users to serve to maximize the profit. Our present work is mainly motivated by this requirement.

When users have different CQIs, one may consider throughput maximization by serving all the users simply at a single robust CQI [11]. Alternatively, one may serve the users with different CQIs by forming groups based on the best CQI at a time [12]. However, these traditional methods suffer from poor quality of service (QoS) or lower service fairness issues. Militano *et al.* [15] discussed the throughput maximization problem, where users have varying CQI in a two-hop single frequency D2D networks by efficiently enabling all the relay nodes that maximize the throughput. They investigated downlink (DL) and uplink (UL) resource allocation by assuming the time division duplex (TDD) model, and distributed all the RBs to the eligible users by a single frequency. Therefore, all the users who reuse the same RBs will be compelled to use the same CQI. However, as per this method, an immediate requirement for the second hop D2D users is to have a channel quality at least same as the first hop CUs that serve them, which may not be possible in a random scenario.

In this paper, we consider a multicast scenario in a TDD based two-hop D2D network, where each user possesses distinct CQI and data request rates, while in return each user may offer different profits to the operator. We divide the transmission into downlink (DL) and uplink (UL) slots, where CUs are served in the DL slots, and the underlaid DUs are served in the UL slots by reusing the RBs. Our two primary objectives are (i) to maximize the profit of the telecom operator and (ii) to maximize the number of satisfied users when RBs are limited. However, these two objectives are related as latter is a special case of the former. The first objective targets at maximizing the overall revenue of the operator by satisfying the users that offer the best profit. On the other hand, the second objective has the intuition that by satisfying more number of users, the operator may find it easy to attract new customers to join the network.

The main contributions of this work are as follows.

i. We are the first to consider profit maximization problem in the D2D model and we show that the problem is NP-hard as given in subsection III. B.
ii. We propose a greedy heuristic and two approximation algorithms to address the objective of maximizing the operator's profit. In the greedy heuristic, we prioritize the users based on individual user's profit to RB ratio. In the other two approximation algorithms, we assign the RBs by considering all possible CQI combinations to the users in the DL and UL slots. These two algorithms promise an approximation ratio of 2 and $2g$ respectively, where g represents the number of CQI groups we use.
iii. We propose a greedy heuristic to address the second objective of maximizing the number of satisfied users.
iv. Finally, we perform simulation whose results show that the performance of all our proposed algorithms is better than the baseline algorithms when we consider profit, data rate, fairness index and satisfied user count.

We organize rest of this paper in the following way. In Section II, we review some of the most relevant research works that have drawn our attention. In Section III, we describe our system model, which is followed by our algorithms in Section IV. Section V narrates our simulation results, and finally, in Section VI we conclude the paper.

II. RELATED WORK

In the past few years, several researchers have studied resource allocation schemes for throughput maximization in the context of multicast scenario [11]–[26]. One category of literature addresses the objective of improving the overall system throughput by jointly optimizing power and resource allocation with the aid of game theory [13] [18]–[23]. The other category aims at efficient resource allocation schemes to maximize the throughput based on transmission rate [11], [12], [24], [25].

For the first category, Lin *et al.* [13] have given an analytical model for D2D multicast to measure the network throughput based on node mobility and network assistance features. These articles have not considered the channel quality of each of the links for multicast. A few other works have shed light on the same objective of throughput improvement by a game theoretical approach to address interference mitigation with the help of suitable power regulatory mechanisms [18], [19]. In [18], the D2D users form clusters and data is multicast into the cluster. In this case, any lost packets will be replaced by the users within the cluster, not by e-NB. To distribute the resources efficiently, the authors have tried joint power management and resource distribution using game theory [19]. The different game models like cooperative, non-cooperative, and auction-based games have been suggested for resource management. A Stackelberg based game has been proposed for resource allocation between e-NB, D2D, and CUs [20]–[23]. In these works, the authors have modelled the scenario as a buyer-seller pair to solve the problem of interference and throughput maximization by resource allocation for the D2D underlay networks. In these models, e-NB offers some incentives to attract users. The authors demonstrated that the resultant utility like throughput, power, and interference achieves a stable state of equilibrium. Another main point to observe in these kinds of literature is all of them use D2D as a transceiver pair for the CUs to serve as an underlay, which is unlike our model.

For the second category, it can be further classified into single-rate and multi-rate schemes. For single-rate schemes which serve all the users at a single CQI, Afolabi *et al.* [11] discussed a simple fairness based multicast scheme for OFDMA multicast network. Here, e-NB considers the most robust channel quality (CQI) to serve all the users at once. Although, this method promises fairness, it will not be efficient in quality of service (QoS) when the majority of the users have high CQIs while only a few users have low CQIs. Low *et al.* [12] considered another extreme, where their method serves the users with the best channel quality at every serving time interval. As a result, this scheme renders best user group to maximize user throughput at every time instant; nevertheless, one cannot expect fairness. Zhang *et al.* [24] considered only those users with SNR above a preset threshold to serve. Alexious *et al.* [25] suggested three methods to improve the spectral efficiency. All these methods, target to select the best modulation and coding scheme (MCS) that enables to reach a predefined spectral efficiency.

For multi-rate schemes, they divide the network into different groups and use different network coding to serve the users. As a result, these schemes may exploit the diversity of users' channel qualities to serve users at different rates simultaneously. Militano *et al.* [26] proposed a subgroup based service method by considering the CQI of the users and the number of users that come under the same CQI group. The method of group formation considers throughput and user fairness [27]. There are similar schemes in [28] and [29], where the authors proposed an algorithm FAST, which is a subgroup-based scheme for OFDMA network. In this method, the throughput will be the product of the number of users and the maximum CQI used for the group. It promises the best performance at reduced computational complexity. Chen *et al.* [30] proposed a group based utility improvement scheme for e-MBMS based multicast service using dynamic programming. They analyzed the utility and achieved a trade-off between the unicast and multicast users group size and throughput.

All the above works consider networks without relays. In contrast, some authors have considered the relay-based methods for OFDMA [16], [32], [33] and D2D [15], [31]. Zhang *et al.* [31] proposed a relay based D2D communication, where particular power controlled D2D users act as a relay to forward data to their counterpart DUs and optimize the achievable data rate of D2D users. This proposal outperforms the direct cellular mode in data rate maximization for D2D communications. Militano *et al.* [15] used single frequency based D2D with TDD based multicast. Initially, they enforce the channel quality of the uplink to be at least equal to the channel quality of the downlink. Then, users with the CQI that could maximize the throughput are served by allocating all the RBs in DL and reusing them to relay in UL. As a foundation for all these works, a relay based scheme in OFDMA network has been proposed in [32] and [33]. It is an approximation algorithm for frequency division duplex (FDD) and Time Division Duplex (TDD) based multicast in OFDMA relay networks. It models the problem as d-MCKP by choosing a resource from e-NB to relay node and from relay nodes to users.

III. SYSTEM MODEL

Let N be the number of users located in a geographical region who demand the multicast content from the base station (e-NB) at different data rates. The N users consist of CUs and DUs. Let N_C be the number of CUs, and N_D be the number of DUs. We use CU_n $(1 \leq n \leq N_C)$ and $DU_m(1 \leq m \leq N_D)$ to label the CUs and the DUs. The topology of the connections form a two-hop tree, rooted at e-NB. See Fig. 1 for an example. We have assumed a pre-assigned network topology, where user's role as CU or DU is pre-determined. It even includes parent-child relationship and their CQIs. In addition, for scheduling, we consider the standard CQI feedback method, where each user reports CQI to e-NB over dedicated control channels. The e-NB adjusts the modulation schemes based on the received CQI feedback [34].

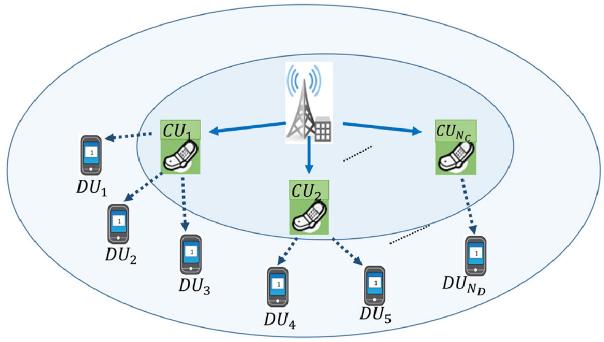


FIGURE 1. Network scenario of CUs and DUs.

Let dR_i and P_i be the data request rate (bits/sec) and profit (\$) that user $i(1 \leq i \leq N)$ offers to the operator, respectively. Besides, different users may have their own CQIs due to their varying distances from e-NB, and we use C_i to denote the CQI of user i. Let T be the number of available RBs at e-NB.

In our model, downlink (DL) time slots and uplink (UL) time slots are interleaving, where e-NB directly serves the CUs in the DL time slots, and the children DUs (underlay) reuse the RBs from their respective parent CUs in the successive UL time slots. To avoid the interference between the users of the same channel, we compel all DUs (across the whole network) at a particular UL time slot to reuse the same RBs and to adhere to the same CQI (as suggested in [35]). Furthermore, when we relay the data from DL to UL in the consecutive slot, the CQI value of UL cannot be greater than the CQI value of the DL, because we cannot transmit excess data in the UL session than what we have received in the DL session. When a CU gets multiple RBs in a DL session, all the received RBs have to be reused to the child DU in the next UL session. Multiple transmission sessions may be needed to satisfy a user's data request completely [37].

When a particular transmission session is done via a certain CQI c, all users with CQI smaller than c cannot receive any data in that session; in contrast, all users with CQI at least c receive all the RBs sent in that session. Also, the transmission rates with different CQIs could be different; a standard assumption is that the rate is proportional to the CQI value [15]. Since the rates will not play any role in our discussion (except to determine the total data a user receives), we will not define the corresponding variables explicitly for simplicity.

A. MAXIMIZING THE SATISFIED PROFIT AND MAXIMIZING THE SATISFIED USERS

Our first objective is to maximize the *satisfied profit* where satisfied profit refers to the revenue offered by a user when its data request has fulfilled. As the profit offered by the individual users may differ, the primary goal will be to pick out the users that can maximize the total collected profit without exceeding the available RBs.

Our second objective is to maximize the number (count) of the *satisfied users*, where a user is *satisfied* if and only if its data request has met. Note: this is a special case of the previous problem when the profits of all users are the same.

We represent our objectives formally by an integer program. We shall define some extra variables:

d_i^t: Data received by user i in the t^{th} DL-UL slot
R^t: Number of resource blocks used in the t^{th} DL-UL slot.
C_{DL}^t: CQI used by downlink in the t^{th} DL-UL slot
C_{UL}^t: CQI used by uplink in the t^{th} DL-UL slot
Additionally, two 0-1 indicators:
z_i: 1 if data request of user i is satisfied
y_i^t: 1 if user i receives data during the t^{th} DL-UL slot

The integer program is as follows:

In the following, the equations (1) and (2) represents the objective functions of maximizing the satisfied profit and maximizing the number of satisfied user count respectively.

$$\max \sum_{i=1}^{N} z_i P_i \tag{1}$$

$$\max \sum_{i=1}^{N} z_i \tag{2}$$

$$\text{Subject to: } \sum_{t} d_i^t \geq z_i dR_i \tag{3}$$

$$d_i^t = y_i^t C_{DL}^t R^t \quad \text{if user } i \text{ is a CU} \tag{4}$$

$$d_i^t = y_i^t C_{UL}^t R^t \quad \text{if user } i \text{ is a DU} \tag{5}$$

$$\sum_{t} R^t \leq T \tag{6}$$

$$C_i \geq y_i^t C_{DL}^t \quad \text{if user } i \text{ is a CU} \tag{7}$$

$$C_i \geq y_i^t C_{UL}^t \quad \text{if user } i \text{ is a DU} \tag{8}$$

$$C_{DL}^t \geq C_{UL}^t \tag{9}$$

Both the objectives are subject to the constraints as given in (3)- (9). Equation (3) constraints that a user is satisfied if and only if the data received is at least the data requested. Equations (4) and (5) compute the data received by user i at a certain slot, under the standard assumption that transmission rates are proportional to the CQI value used in the transmission. These equations can be replaced when we apply a different model of transmission. Equation (6) constraints that the total number of resource blocks cannot exceed T. Equations (7) and (8) constrain that a user receives data at a certain time-slot if and only if its CQI value is at least the CQI used in the transmission. Finally, (9) requires that CQI in a certain UL slot cannot exceed the DL slot immediately preceding it.

B. MAXIMIZING THE SATISFIED PROFIT: NP-HARDNESS PROOF

We now prove that our problem of maximizing satisfied profit in NP-hard; hence, justifying the need to consider approximation algorithms.

Our proof is by a polynomial-time reduction from the well-known NP-hard Subset Sum problem.

To prove our problem is NP-hard, our proof process follows two standard steps:

1. First, given an instance of the Subset Sum problem, we show that we can reduce the instance to an instance of our problem.

2. Second, show this reduction takes polynomial time. As a result, for the outcome of our problem, we can correctly determine the answer to the original Subset Sum problem.

In definition 1, we have defined our multicast problem.

Definition 1: Given a two-hop TDD based D2D multicast network, with users having unique CQIs, data request rates, and profits, we aim to maximize the profit obtained by the satisfied users assuming to have limited RBs. Further, in this model, at the first time slot (DL) the e-NB transmits to the CUs and in the subsequent time slot (UL) CUs forward the data to the child DUs in the second hop. If more than one CU wishes to forward the same data received from the base station simultaneously to their corresponding child, then all of them will be compelled to use the most common CQI.

In definition 2, we describe the subset sum problem, a well-known NP-hard problem that we use in our reduction.

Definition 2: The Subset sum is a decision problem: given a set S of k positive integers, $S = \{s_1, s_2, \ldots, s_k\}$, and a target value t. Determine a subset S' of S, such that its sum exactly equals to t.

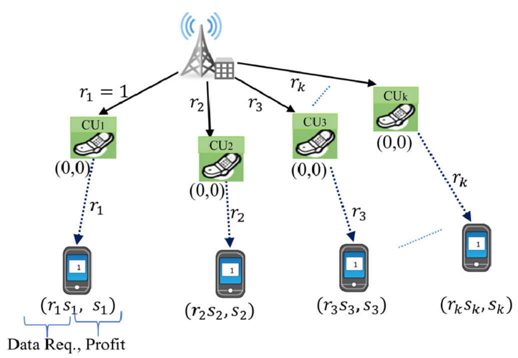


FIGURE 2. The instance transformed from the Subset Sum problem.

In particular, the network topology, after the transformation, will be a tree with each CU having a single child DU as shown in Fig. 2.

There are k CUs connected from the base station with CQI value of CU_i equal to:

$$r_i = 1 \quad \text{if } i = 1$$

$$r_i = (1 + \sum_{j=1 \text{ to } i-1} s_j) r_{i-1} \quad \text{if } i \neq 1$$

The data request and the profit of every CU are both zero. Let DU_i denote the child of CU_i. The CQI of each child DU from its respective parent CU is set to r_i, while the data request and profit of DU_i are $r_i s_i$ and s_i, respectively.

In general, any element s_i of set S represents the number of RBs needed and profit offered by DU_i when data is transmitted at CQI r_i.

The following is a key lemma about the property of the transformed instance.

Lemma 1: Let D* denote a set of DUs that offers the maximum satisfied profit with t resource blocks. Then, the maximum satisfied profit is at least t if and only there exists a transmission schedule that sends exactly s_i resource blocks (using CQI r_i) to each DU_i in D* and $\sum_{DU_i \in D^*} s_i = t$.

Proof: ($\leftarrow$) From the resource blocks sent, all the DUs in D* will be satisfied, so that total profit is at least that amount offered by these DUs, which is $\sum_{DU_i \in D^*} s_i = t$. Thus, the ''if'' part follows.

($\rightarrow$) Consider a transmission schedule that satisfies all DUs in D*. Let R_i denote the number of resource blocks sent to a satisfied DU_i in D*. If $R_i \geq s_i$, one can retain exactly s_i resource blocks to send to DU_i, and upgrade the remaining to use any higher CQI values so that all the satisfied users in D* are still satisfied. On the other hand, we claim that if $R_i < s_i$, then one must be using at least

$$Y = s_1 + s_2 + \ldots + s_i - R_i$$

resource blocks with lower CQIs than r_i. In such a case, simply retain exactly $s_1 + s_2 + \ldots + s_{i-1}$ resource blocks with s_j of them using $r_j, j < i$ (thus satisfying all DU_1 through DU_{i-1}), and convert the remaining resource blocks with CQI r_i (thus keeping all those already satisfied users still satisfied, and arriving the first case of $R_i \geq s_i$).

To see why the claim is true, suppose to the contrary, send less than Y blocks with CQIs lower than r_i. Then, the data received by DU_i is less than

$$\begin{aligned} &Yr_{i-1} + R_i r_i \\ &\quad \leq (s_i - R_i)(s_1 + s_2 + \ldots + s_{i-1} + 1) r_{i-1} + R_i r_i \\ &\quad = (s_i - R_i) r_i + R_i r_i = s_i R_i, \end{aligned}$$

which is impossible. Thus, the claim follows.

From the above discussion, an optimal schedule can be converted into one that sends exactly s_i resource blocks (using CQI r_i) to each DU_i in D*. These resource blocks are sufficient to satisfy all DUs in D*, so that no further resource blocks will be needed. Consequently, we have $\sum_{DU_i \in D^*} s_i \geq t$ (since the total profit is at least t) and $\sum_{DU_i \in D^*} s_i \leq t$ (since the total transmitted resource blocks is at most t). This implies $\sum_{DU_i \in D^*} s_i = t$, and thus the ''only if'' part follows as required. ■

With the help of Lemma 1, the direct connection between the original instance of Subset Sum, and the transformed instance of our problem will be established by the following lemma.

Lemma 2: Suppose that the number of resource blocks in the transformed instance is t. Then, the maximum satisfied profit is t if and only if the original instance of Subset Sum contains a subset whose sum is exactly t.

Proof: It is easy to check that the following two statements are equivalent:

1. There exists a transmission schedule that sends exactly s_i resource blocks (using CQI r_i) to each DU_i in D* and $\sum_{DU_i \in D^*} s_i = t$.
2. The original instance of Subset Sum contains a subset whose sum is exactly t.

So, Lemma 2 follows based on the above and Lemma 1. Hence proved. ■

For instance, let us consider an example of a subset sum problem with input multi-set $S = \{2, 3, 5, 6, 12, 18, \ldots\}$ and target is $t = 120$. Using the above transformation, one can construct a network topology as in Fig. 2, and determine the CQI, data rate, and profit values of each user. For instance, the CQI of CU_1 to CU_3 (and from these CUs to their corresponding child DUs DU_1 to DU_3) are 1, 3, 18, respectively. The data rates and profits of DU_1 to DU_3 be (2, 2), (9, 3), (90, 5), respectively.

Finally, the following shows that the given transformation takes polynomial time. Note that the transformation generates exponentially large numbers (such as r_i and $r_i s_i$), so one must be careful while arguing that the transformation indeed takes polynomial time. First, let I denote the original instance of Subset Sum problem, and I' denote the transformed instance of our problem. Furthermore, let $||I||$ and $||I'||$ be the number of bits to encode I and I', respectively. Thus, we have $||I||$ is at least $\Omega(k + \log t + \log \max\{s_i\})$ bits. Following steps show the generation of I' in time polynomial to $||I||$. Recall that I' contains a two-hop tree, with size linear to k, so that the topology can be constructed in O(k) time. Next, it has a CQI value, a data request value, and a profit value associated with an edge or a node. The largest of these values is bounded by $O(kmax\{s_i\})^{O(k)}$ whose encoding thus takes $\log O(kmax\{s_i\})^{O(k)} = O(k) \times O(\log k + \log \max\{s_i\})$ bits, which is polynomial in $||I||$. Moreover, each value can be generated in time polynomial to its encoding length. In summary, the transformation takes polynomial time. This completes the proof. ■

The above discussion leads directly to the following theorem.

Theorem 1: Maximizing the satisfied profit in a D2D network with finite resource blocks is NP-hard. ■

IV. ALGORITHMS

This section is devoted to our proposed algorithms. We first describe a greedy algorithm and two approximation algorithms to maximize the satisfied profit, in Sections IV. A, IV. B and IV. C, respectively. A brief description of these algorithms is as follows.

The greedy algorithm chooses the users based on their ratio of offered profit to the required number of RBs. Due to the limited availability of RBs, not all users might be served.

The main intuition here is to serve the user with the maximum average profit to get satisfied at the earliest to grab its profit. It mainly has 3 phases. In the admission phase, we assign RBs to a user if its RB requirement can be met. In the reclaiming phase, excess RBs with a CU will be taken back and in the unifying step, all DUs will try to use a common CQI while retaining already satisfied users unaltered.

In the first approximation algorithm, given a range of CQIs and a set of RBs, we enumerate all possible combinations of CQI on the set of RBs to multicast the data in the DL and UL slots. Following this, we determine the profit obtained for each CQI combination. Later, the CQI combination that offers the maximum profit either in DL or UL will be our solution. The main reason behind this method is that at least we can obtain half of the total profit of an optimal solution, where the optimal solution is the result achieved by satisfying users in the DL and UL considered together.

In the second approximation algorithm, given a set of RBs, and a range of CQIs, we divide the entire range of CQIs into multiple smaller groups based on the desired approximation ratio. Then, we enumerate all possible CQI combinations from each of these groups on the set of RBs to multicast in DL and UL to determine the profit of each group. Finally, the combination of CQIs from the group that offers the maximum profit will be chosen as our solution.

Then, to address the objective of maximizing the number of satisfied users, we propose a greedy algorithm in IV. D to select the users who can maximize satisfied user count. In our multicast model, when a user is served, its peer may also receive the data and get satisfied at the same time. Our algorithm targets at finding such users that have the potential to increase the satisfied users count. The algorithm arranges the users based on the ratio of the number of satisfiable users to the required number of RBs and serves those with higher ratios first. Thus, the user with the highest potential to satisfy multiple users simultaneously will be served first to maximize the user count.

In our optimal multicast solution, we exhaustively try all CQI combinations with our available RBs to the users in DL and UL; and determine the profit obtained by all those satisfied users by considering the DL and UL together.

A. MAXIMIZING THE SATISFIED PROFIT: GREEDY HEURISTIC

As the total RBs are limited, it is intuitive to select a user that has high potential to maximize the profit. Our greedy-based heuristic in this subsection follows such an intuition. The heuristic consists of two main steps:

Step 1:

i) Setting up the priority among the users. Here, we first determine the minimum number of RBs needed by each user, assuming that it does not receive any data shared by the other users. This is computed by dividing its data request by its CQI value as in (10), where R_i is the minimum number of RBs needed by user i.

ii) Next, we determine the maximum profit per RB that each user can offer, which is computed by dividing its profit by the minimum number of RBs that it needs. Such a value is called the *weight* of the user. Finally, we arrange all users in descending order of their weights, which sets up the *priority* among all users. (In our algorithm, when two users have the same weight, we break ties by first selecting one whose CQI value is lower (in case of CU) or whose parent CQI value (in case of DU) is lower. If the CQI values are still the same, we break ties randomly.)

$$R_i = \lceil dR_i/C_i \rceil \tag{10}$$

Step 2:

Assigning RB to satisfy users, iteratively, based on the priority in Step 1. Here, we consider the users one by one, starting from the one with the highest priority, and do the following:

1) Use an *admission* procedure to decide if we want to satisfy this user. The invariant is that when we decide to satisfy a user, this user will be assigned enough RBs to satisfy its data request, while *all the already satisfied users remain satisfied*.
2) If the user is a CU, we assign sufficient RBs to the user to satisfy its data request, using its CQI for transmission. After this, we perform a *reclaim* procedure, which examines if any RBs that are already assigned to the other CUs can be returned to the system, without changing the satisfiability status.
3) If the user is a DU, we consider all the satisfied siblings (i.e., those DUs sharing the same parent as the user) and get the minimum CQI value among the user and the siblings (known by the *CQI feedback*). Then, we compute the number of RBs needed to satisfy the user and all the satisfied siblings, where the minimum CQI is used for transmission among them. We call this a *unify* procedure. After this, we will find out the extra number of RBs that should be sent to the user's parent, assign these extra RBs to the parent, and perform a *reclaim* procedure on the CUs as in Step 2.

Let us now present the details of three procedures used in Step 2 as follows.

1) ADMISSION

1. If the user is a CU: We check if the remaining number of RBs with e-NB is at least the minimum RB needed by the user. If so, the user is admitted. Otherwise, the user is skipped.
2. If the user is a DU: We compute the number of RBs needed to serve the user, and all the satisfied siblings, after the *unify* process (see below). Consequently, the parent CU of the user may require extra RBs to serve the user and the siblings, if the user is admitted. Next, we check if the remaining number of RBs is more than or equal to the number of extra RBs needed by the parent CU. If so, the user is admitted. Otherwise, the user is skipped.

2) UNIFY

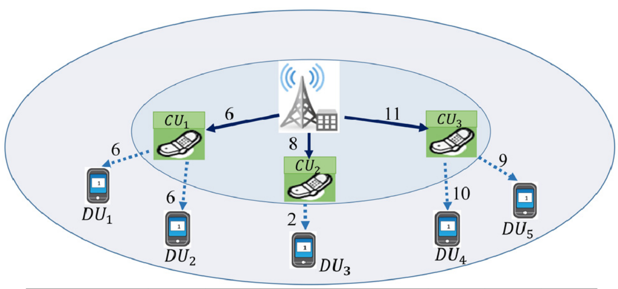
1. If the user does not have an already satisfied sibling: We assign sufficient RBs to the user, using its CQI value.
2. Else, all transmissions from the parent to the user, and the satisfied siblings are assumed to use a common CQI, which is set to the minimum CQI among the user and these siblings. We re-compute the number of RBs needed to satisfy the user and these siblings, and assign sufficient RBs (in total) to the parent to allow this to happen.

3) RECLAIM

When a CU, say $CU_i\mathrm{CU_i}$, is assigned with additional RBs, all the other CUs with equal or higher CQI values can receive the same data sent to CU_i. Consequently, the number of RBs needed by the other CUs to fulfill their data requests (or their children's data requests) may be reduced.

For example, when we transmit the data to the priority CU (say user *A*) by a specific CQI in the DL slot, the other CU in the DL (say user *B*) with a higher channel quality than the priority user may also receive as the transmission is multicast. Now, suppose user *B* had already received data due to some other DL transmission previously, then the data received in the current DL will be an excess. In such a case, this user *B* may replace its RBs of higher CQI received in the previous DL slot by the RBs of lower CQI received in the current DL slot. In this way, both the users (*A* and *B*) can share the same RBs, and the RBs of higher CQI will be given back to the system.

To perform the procedure, we shall examine each of these CUs in turn, and take back as many RBs as possible such that *all the satisfied users remain satisfied*. This step is performed by brute force checking and can be sped up slightly when we use binary search to find out the maximum RBs to be taken when we examine a particular CU.



User	CU_1	CU_2	CU_3	DU_1	DU_2	DU_3	DU_4	DU_5
Profit	10	30	40	15	22	15	20	30
Data Req.	100	300	400	150	220	150	200	300

FIGURE 3. An example of a D2D multicast scenario.

Let us consider the example in Fig. 3 to get a better understanding of our heuristic. The CQI value of each link is shown beside the link; while the profit and data request of each user is shown in the lower part of the Fig. 3 as a table. We assume that $T = 10$ (the total available RBs) and CQI with value x offers 10x bits/RB.

TABLE 1. Priority computation steps.

User	Profit	Min RB	Weight	Priority
CU_3	40	4	10	1
DU_4	20	2	10	2
CU_2	30	4	7.5	3
DU_5	30	4	7.5	4
DU_2	22	4	5.5	5
CU_1	10	2	5	6
DU_1	15	3	5	7
DU_3	15	8	1.875	8

In *Step 1*, we compute the minimum number of RBs required by each user (assuming no data is received from other users), then obtain the *weights* (= *profit/minimum RB*) of all users to assign priority to them. The resulting priority is given in Table 1.

In *Step 2*, we examine the users one by one and perform resource allocation. Recall that in our algorithm, we assume all RBs sent to a CU will be done via its CQI; while all these RBs will be reused by its children via a common CQI, (which is the minimum CQI among all its satisfied children).

A *triplet* (#RB, CQIown, CQIchild) can summarize the resource allocation, specifying the number of RBs sent to each CU, its own CQI, and the common CQI that it relays the message to its children, respectively.

First allocation step (an example of admitting a CU): We first choose the first priority user CU_3 who needs 4RBs. From the admission rule, the RB requirement of CU_3 is smaller than total remaining RBs (i.e., $T = 10$). So, e-NB will admit and serve CU_3 by assigning 4 RBs of CQI 11 to satisfy it. At this moment, there are no satisfied children, so we denote such an allocation by writing (4, 11, 11) in allocation step 1 for CU_3 in Table 2. The total remaining RBs is updated to $10 - 4 = 6$.

Second allocation step (an example of admitting a DU): Next, we choose the second priority user DU_4 who needs 2RBs. Since it is a DU, we first simulate the unify procedure to see if sufficient RBs are available to serve it. Here, DU_4 does not have any satisfied siblings, we will assume its parent CU communicates with it using its own CQI value. As its parent (CU_3) has received enough RBs where all can be reused to serve it, no extra RBs are needed, so that DU_4 is admitted, and after that, served. The CQI value of DU_4 is 10, so in Table 2, we write (4, 11, 10) in allocation step 2 for its parent CU_3 to indicate such an assignment. The total remaining RBs is still 6.

Third allocation step (an example of reclaim after admitting a CU): After that, we look at CU_2 who is next in the priority list. CU_2 needs 4 RBs so that e-NB will admit CU_2 since it has enough remaining RBs (6). Thus, it will assign 4 RBs at CQI 8 in DL slot to satisfy it. Note that CU_3, an already satisfied user, will also receive data from this transmission, and the reclaim procedure is invoked, which finds that 3 RBs that are sent to CU_3 earlier can be taken back, such that all the satisfied users are still satisfied.

To indicate the above assignment, and the changes, we write (4, 8, 8) in the entry for CU_2, and (1, 11, 10) in the entry for CU_3, in allocation step 3 of Table 2. The total remaining RBs is updated to 5.

Fourth allocation step (example of unify after admitting a DU): Next, we examine DU_5, and find that it can be admitted following a simulation run of unify. As DU_5 has an already satisfied sibling (DU_4), our algorithm requires that all of them will receive RBs with a common CQI, which is equal to min $\{9, 10\} = 9$. Then, by checking, we see that no RBs is actually needed from its parent CU_3, since they both receive 320 bits from CU_2 by data sharing. Nevertheless, we would still write (1, 11, 9) in the entry for CU_3 in the allocation step 4 of Table 2, to indicate that CU_3 itself needs one RB from e-NB, and the common CQI of its satisfied children is 9. As no extra RBs are needed to send to CU_3, the total remaining RBs is still 5. (Note that in this step, if the CQI of DU_5 is smaller, say 3, then we will need extra RBs to satisfy all satisfied children using the common CQI value 3; consequently, extra RBs will be sent to the parent CU_3, and then the reclaim procedure will be invoked.)

TABLE 2. Resource allocation steps (first 4 steps).

Step	1	2	3	4
CU_1				
CU_2			(4, 8, 8)	(4, 8, 8)
CU_3	(4, 11, 11)	(4, 11, 10)	(1, 11, 10)	(1, 11, 9)

The above allocation steps have demonstrated how all the different procedures are performed. We will skip the remaining allocation steps for brevity. Finally, the last column of Table 2 (after all allocation steps are done) shows for each CU the number of RBs sent from e-NB, the CQI used by e-NB to the CU, and the common CQI used for it to send to its children. All the information of the previous allocation steps are discarded.

Finally, we sketch the time complexity of our heuristic algorithm. Step 1 can be performed in $O(N \log N)$ time by computing the weights in $O(N)$ time and then perform sorting. In Step 2, we consider the N users, one by one, where for each user, it takes $O(N^2)$ time to check (by brute force) if it can be admitted. In case it is a CU, the reclaim step takes $O(TN^3)$ time by a brute force checking, but it can be improved to $O(N^3 \log T)$ time with binary search [or even to $O(N^2 \log T)$ time by using some standard data structure techniques]. In case it is a DU, the unify step takes $O(N^2)$ time, which is possibly followed by a reclaim step for its parent CU. In summary, Step 2 takes $O(N^3 \log T)$ time for each user. Thus, the time complexity of the heuristic is $O(N^4 \log T)$.

B. MAXIMIZING SATISFIED PROFIT: APPROXIMATION ALGORITHM 1

This subsection describes our first approximation algorithm. Before going into the details, we introduce the following lemma.

Lemma 3: Consider the special case of our problem where the multicast network does not have DUs. That is, the topology of connections forms a one-hop tree, and the transmission from e-NB to the CUs is essentially a broadcast. Then, when e-NB transmits a sequence of RBs by different CQIs at different time instances, the total data received by a CU remains the same irrespective of the order of the CQIs used.

Example: If e-NB transmits 3 RBs with CQIs [1, 2, 2], respectively at different time instants, then for any CU, the data received would be the same if e-NB transmits 3 RBs with the CQIs in a different order, say [2, 1, 2] or [2, 2, 1].

Proof: When e-NB transmits (i.e., broadcasts) an RB with a certain CQI c, the amount of data received by user i depends only on c and C_i (all bits if $c \leq C_i$, and 0 bits otherwise), which is independent of the other RBs. Thus, the ordering of the CQIs used does not affect the total data received by any user. Hence proved. ■

Our first approximation algorithm proceeds as follows.

1) Compute a solution that maximizes the total satisfied profit from the CUs. That is, we ignore all the DUs (as if their offered profits are all 0s).
2) Compute a solution that maximizes the total satisfied profit from the DUs. That is, we ignore all the CUs (as if their offered profits are all 0s).
3) Compare the solutions in the above steps. Return the one with a higher total satisfied profit.

Let $\mathbb{C}$ be the set of CQIs used by the users in the network and C be the size of $\mathbb{C}$. As there are T RBs, by Lemma 3, we can compute the optimal solution of Step 1 by first enumerating all size-T combinations of (possibly repeating) elements in $\mathbb{C}$, and for each combination, finding out the satisfied users and their total profit. The number of combinations is

$$\mathcal{L} = \binom{T + C - 1}{C - 1}$$

so that the total time required is $O(CN\mathcal{L}) = O\left(CNT^C\right)$.

For Step 2, since the CUs are ignored, the scenario is equivalent to e-NB sending directly to some DU D in each round (using its parent CU as a dummy relay and using D's CQI in both hops). Consequently, the network topology is equivalent to a one-hop tree, with e-NB connecting directly to all the DUs. So, we can compute the optimal solution in this step using the same method as in Step 1, taking $O(CNT^C)$ time in total. Finally, in Step 3, the total profits of both steps are compared in $O(1)$ time, and the desired output is reported. The total time of all steps is thus bounded by $O(CNT^C)$.

Next, we will give the approximation ratio analysis. Let OPT_1 and OPT_2 be the optimal profits in Step 1 and Step 2, respectively. Let OPT be the optimal profit in this problem. Note that OPT is no larger than OPT_1 if we only consider profits generated by satisfied CUs, and similarly OPT is no larger than OPT_2 if we only consider profits generated by satisfied DUs. This implies the following relation:

$$OPT \leq OPT_1 + OPT_2 \leq 2 \max (OPT_1, OPT_2) \quad (11)$$

Since Step 3 chooses a solution with total profit $\max(OPT_1, OPT_2)$, performing transmission in the same way as specified by Step 3 must guarantee a profit at least that value (since we may get extra profits from the ignored users), thus our algorithm achieves an approximation ratio of 2.

C. MAXIMIZING SATISFIED PROFIT: APPROXIMATION ALGORITHM 2

Our second approximation algorithm is a simple adaptation of our first approximation algorithm, where the target is to provide a tradeoff between running time and approximation ratio. The algorithm runs as follows:

1) Partition the set $\mathbb{C}$ of CQI (evenly) into g groups $\mathbb{C}_1, \mathbb{C}_2, \ldots, \mathbb{C}_g$, each of size at most $C' = \lceil C/g \rceil$.
2) For each group $\mathbb{C}_i$ of CQIs,
   a) Assume that we can only use CQIs from $\mathbb{C}_i$.
   b) Run Approximation Algorithm 1.
3) Return the best profit among all g solutions obtained in Step 2.

The above algorithm runs Approximation Algorithm 1 for g times, each time taking $O\left(C'NT^{C'}\right)$ time. In total, the running time is $O\left(CNT^{C'}\right)$ time.

As for the approximation ratio, let $OPT_{i,1}$ and $OPT_{i,2}$ denote the optimal profit attained by satisfied CUs and by satisfied DUs, respectively, when transmissions are restricted to use only those CQIs from $\mathbb{C}_i$. Also, we let OPT denote the optimal profit of this problem. Then, we have:

$$\begin{aligned} OPT &\leq \sum_{i=1 \text{ to } g} (OPT_{i,1} + OPT_{i,2}) \\ &\leq 2g \max_i \left(OPT_{i,1}, OPT_{i,2}\right) \end{aligned} \tag{12}$$

On the other hand, performing transmission in the way specified by Step 3 guarantees a total profit of at least $\max_i \left(OPT_{i,1}, OPT_{i,2}\right)$, so that our algorithm achieves an approximation ratio of $2g$.

D. MAXIMIZING SATISFIED USERS: GREEDY HEURISTIC

All algorithms discussed for maximizing satisfied profit can readily be used for maximizing the number of satisfied users, by setting the profit of each user to be the same. Here, we provide a slightly different greedy heuristic for the latter problem. The algorithm is exactly the same as that presented in Section IV. A, except that we use a different way to determine the weights of users in Step 1; once the priority is set, Step 2 will run in the same way as before.

First, we determine the minimum number of RBs as in (10). Later, we find the weight of CUs and DUs as follows:

1. If the user is a CU, we assume that a minimum number of RBs is transmitted to the CU to satisfy it, using its own CQI. We then obtain the maximum satisfied user count, which is the number of CUs that are satisfied with this transmission. For instance, when we multicast to user A, peer CUs B, C, and D also may get satisfied. Then, the satisfied user count for user A becomes four.
2. If the user is a DU, we assume that a minimum number of RBs is transmitted to the DU, relayed via its parent CU, using the DU's own CQI in both hops. We then obtain the maximum satisfied user count, which is the number of CUs and DUs that are satisfied with this transmission.
3. The weight of a user is equal to its maximum satisfied user count divided by the minimum number of RBs to satisfy the user. Users are prioritized by descending order of their weights.

Intuitively speaking, the priority rule in Section VI.A pessimistically computes the minimum potential of profit increase (per RB) for each user, while the priority rule in this section optimistically computes the maximum potential of profit (i.e., user count here) increase for each user. The latter rule can easily be generalized for the maximizing satisfied profit problem (using maximum satisfied profit instead of maximum satisfied user count to calculate the weight) though it might not be suitable for that. Since a bias is easily created when there is a group of users with high combined profits who can satisfy one and the other (but with mediocre profit per user). In contrast, it seems that such a bias cannot be easily created for the satisfied user's problem, because all users carry the same profit (or, if so, then the selected users may as well give a good enough solution). Thus, this priority rule is used primarily in the maximizing satisfied users heuristic, but not in the maximizing satisfied profit heuristic.

Finally, the time complexity of this algorithm remains the same as the heuristic algorithm in Section VI. A.

V. PERFORMANCE EVALUATION

A. SIMULATION SETTINGS

In this section, we present our experimental setup and simulation results. We simulated our experiments with Matlab, by making suitable modifications to the LTE system level simulator available online [36]. Here, the experimental scenario consists of a single base station that controls the CUs and DUs. We have used the TDD model, where we divide the communication into DL and UL slots. The e-NB communicates with CUs in the DL and data transmission from CUs to DUs takes place in the UL slots. We have considered that the CQI value of any user depends on the distance, power, and path loss. The main simulation parameters are listed in Table 3. In the experimental setup, the users will be distributed randomly in a single cell around the e-NB.

B. SIMULATION RESULTS FOR MAXIMIZING SATISFIED PROFIT

We have evaluated our simulation performance by creating two scenarios. In the first scenario, we maintained the number of RBs as a constant and varied the UEs from $10 \sim 50$. In the second scenario, we kept the number of users a constant and observed the performance by varying the number of RBs from $10 \sim 50$. In both the scenarios, the CQI, data request, and profit values of each CU and DU are assigned randomly.

TABLE 3. Main simulation parameters.

Parameter	Value
Cell radius	500 m [26]
Frame Structure	Type 2 (TDD)
TTI	1 ms
TDD configuration	1
Carrier Frequency	2.5 GHz
eNodeB Tx power	46 dBm
D2D node Tx power	23 dBm
Noise power	-174 dBm/Hz
Path loss (cell link)	128.1 + 37.6 log (d), d [km]
Path loss (D2D link)	46.8 + 20 log(f/5) + 16.9 log(d), f [GHz], d [m]
Shadowing standard deviation	10 dB (cell mode); 12 dB (D2D mode)
RB size	12 sub-carriers, 0.5 ms
Sub-carrier spacing	15 kHz

In this two-hop network, we allow any user to assume the role of CU or DU. In addition, the child DUs of a CU are assigned randomly.

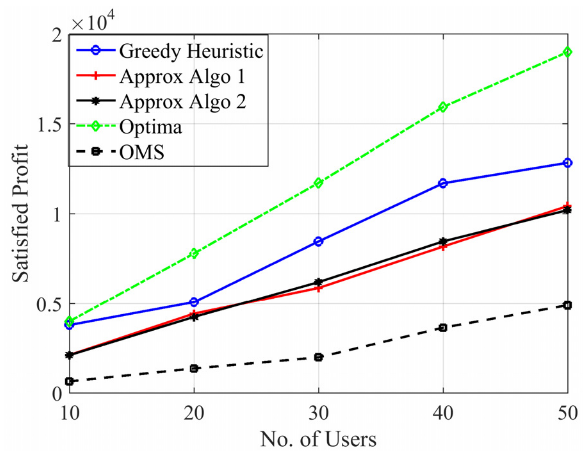


FIGURE 4. Satisfied profit for fixed number of RBs.

In Fig. 4, we determine the satisfied profit by varying the number of UEs from 10 to 50 for a fixed set of 25 RBs. In general, the overall profit increases as we increase the number of users. In the case of our Approximation Algorithm 1, we assign the RBs in the DL and UL slots by all possible combination of CQIs. Later, we choose the maximum profit between the two hops. However, we have avoided the repeated combinations to reduce the running time as mentioned in the description of our algorithm in Section IV. B. In the case of Approximation Algorithm 2, we have made multiple groups of CQIs and chosen the group that offers the maximum profit between the two hops. The practical performance of profit value in case of both approximation algorithms remains almost similar as many users had their CQI values concentrated more at one region of entire CQI's range. As a result, both algorithms may end up choosing almost the same users. However, the profit value of both the algorithms is at least 50% the value of the optimal solution at every instant. We observed that, this performance is due to user distribution, where in most of the simulation rounds we had almost an equal number of users distributed in DL and UL. In the optimal solution, all the satisfied users from both the hops contribute to the total profit. We used exhaustive search to determine the optimal solution. In this case, we assigned an RB by different possible CQI combinations to both the hops, and we try until all the RBs are tested for every CQI combination. Our greedy heuristic algorithm uses a greedy based priority rule to choose the users having the best profit; as a result, the total profit is higher than the other two approximation algorithms. We compared our algorithms with OMS [12] by modifying it for a two-hop network. The OMS makes user groups based on the best CQIs at a TTI and serves the subgroups by different CQIs at each TTI. By this, it compromises with the number of users it can serve. As a result, its performance is lower than our proposed algorithms. The worst-case performance of our approximation algorithms can be drawn by the theoretical discussions as shown in section IV. B and C.

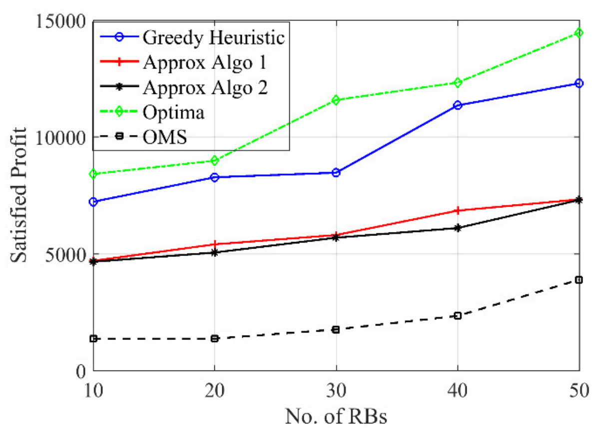


FIGURE 5. Satisfied profit for fixed number of UEs.

In Fig. 5, we have measured the profit for a constant user set of 30 and by varying the RBs. As we vary the RBs, initially the available RBs are not sufficient for all the 30 users to get satisfied of their data request. Therefore, the profit collected will be low, and it increases gradually in all the algorithms as the number of RBs increases. In the case of Approximation Algorithm 1, we varied the range of users CQI randomly from 1 to 15. For assigning RBs, we could use only up to 5 CQIs to maintain low running time. In Approximation Algorithm 2, we observe that the profit will be maximized when the number of groups is two, especially in the first group, where the CQIs were half the range. This is due to the dispersion of more users on the lower CQI range with satisfiable data request rates. The performance of Approximation Algorithm 1 is slightly better than Approximation Algorithm 2, as it scans the entire range of CQI as a single group and the user distribution is more dispersive in this case. Nevertheless, both algorithms exhibit at least 50% of the optimal performance. For our greedy heuristic, it has a better performance than the other two approximation algorithms, despite there is no approximation ratio guarantee. In the case of OMS, it groups users with maximum CQI; however, the number of users that can get satisfied will be lower and so will be the profit.

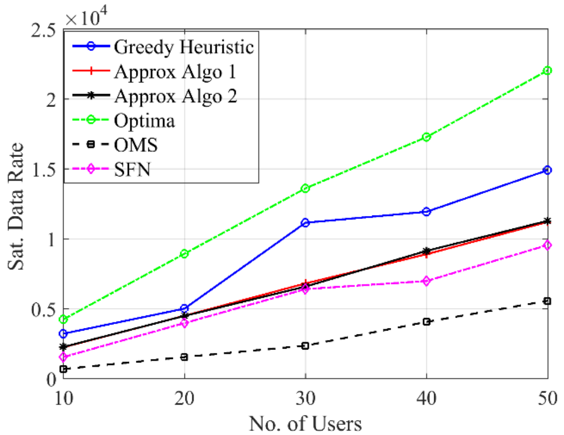


FIGURE 6. Overall satisfied data rate for fixed RBs.

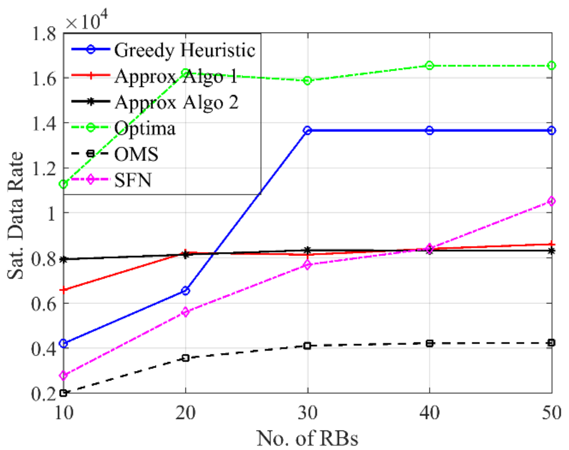


FIGURE 7. Overall satisfied data rate for fixed UEs.

In Figs. 6 and 7, we measure the total data rate of the satisfied users for the profit maximization algorithm.

In Fig. 6, by fixing the number of RBs to 25, we vary the number of users and measure the overall data rate. As more users are added into the system, the total data rate of the satisfied user will increase gradually. However, the variation of data rate from one user set to the other is not the same, as each user set may possess different CQIs and data rates. The data rate and profit of each user are independent, where we set their values as random. Our greedy heuristic algorithm performs better than the other two approximation algorithms. The overall data rates of both approximation algorithms are almost similar due to similar CQI values occupied by most of the user set in both the cases. We compared the overall data rate with SFN [15] which is a relay based two-hop scheme. SFN initially confirms the CQI of DU to be better than CU and enables only those users with best CQI values. As a result, its performance is better than OMS and is very close to our algorithms. However, in our scenario the channel qualities are random, and no such mandatory requirement on child DUs CQI are guaranteed.

In Fig. 7, we measure the overall data request of satisfied users for the fixed UEs scenario. As we vary the number of RBs, the accumulated data rate increases in all the algorithms. However, the performance of our greedy heuristic algorithm becomes close to the optimal solution when the number of RBs is increased to 30 due to sufficient availability of RBs for the users to get satisfied. Later, it remains constant because almost all the satisfiable users have already been satisfied.

We observe that the performances of our approximation algorithms are almost similar because the number of users with CQI values concentrated in the lower range is more which makes Approximation Algorithm 2 to perform similarly to Approximation Algorithm 1.

On the other hand, until the RBs reaches to 25, the performance of greedy heuristic is worse than these two approximation algorithms. In the cases of SFN and OMS, the data rate gradually increases, but their performances are both worse than our algorithms.

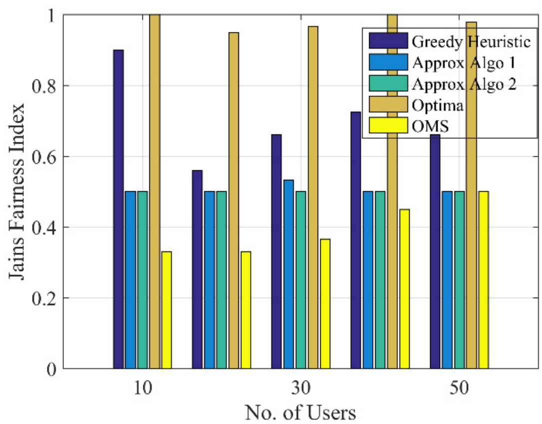


FIGURE 8. Jain's fairness index for fixed RBs.

In Fig. 8, we have shown the Jain's fairness index (FI) [29] for varying number of UEs. It is a measure of how well the users are served with the data. An FI value of 1 represents the maximum fairness.

We can observe that both approximation algorithms have almost similar fairness index as they perform similarly when we consider throughput and users that get served. Our greedy heuristic algorithm has higher fairness index as it satisfies more users at a higher data rate than the rest. In the case of OMS, the fairness index is lowest as it cannot satisfy as many users as the proposed algorithms.

In Fig. 9, we show the Jain's fairness index for varying number of RBs. Again, both approximation algorithms have almost similar fairness index as they satisfy almost same users and collect nearly equal data throughput. Our greedy heuristic has higher fairness index, which is near to optimal as it satisfies more users comparatively.

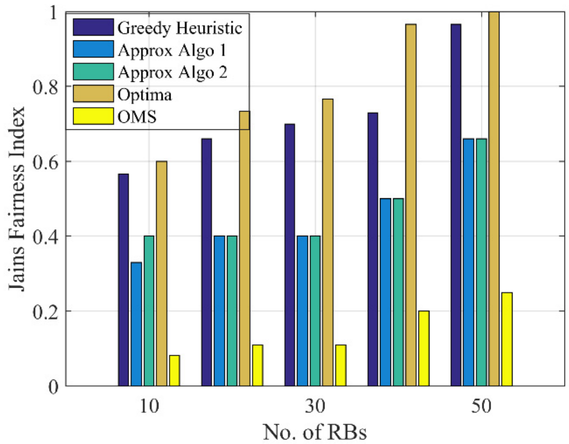


FIGURE 9. Jain's fairness index for fixed UEs.

However, the FI value in Fig. 9 is lower than the value in Fig. 8 until the number of RBs reaches to 30 as the collected data rate due to satisfied users of all the algorithms is comparatively lower. The main reason is the poor channel quality of the UEs. As the number of RBs increases, the users get sufficient RBs to satisfy their data requests. As a result, the fairness index also increases when the number of RBs increases to 40 and later. The same effect has shown by the OMS also in this scenario.

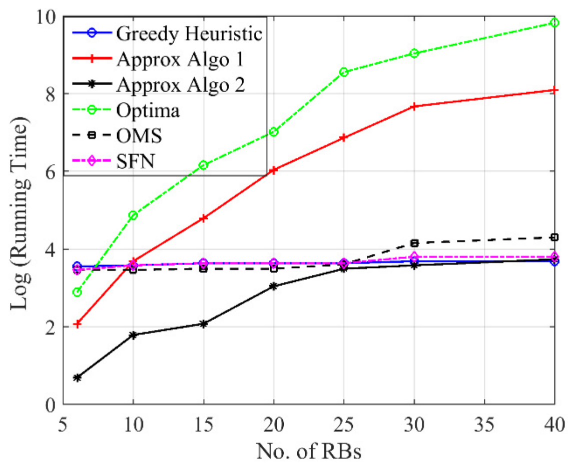


FIGURE 10. Running time of algorithms (in seconds).

In Fig. 10, we show the running time (in log scale in the y-axis) of all the algorithms. When we increase the number of RBs, the running time will grow exponentially in the case of the optimal algorithms than the approximation algorithms, as we need to check all possible CQI combinations for the chosen set of RBs. However, in the case of our greedy heuristic, the running time remains almost the same, as its dependency on the number of RBs is low. Similarly, in the cases of SFN (which overlaps with greedy in the figure) or OMS, the time required remains almost the same. In general, the running time is significant and mainly depends on the various CQI combinations that we need to apply when varying the number of RBs as shown in Fig. 10.

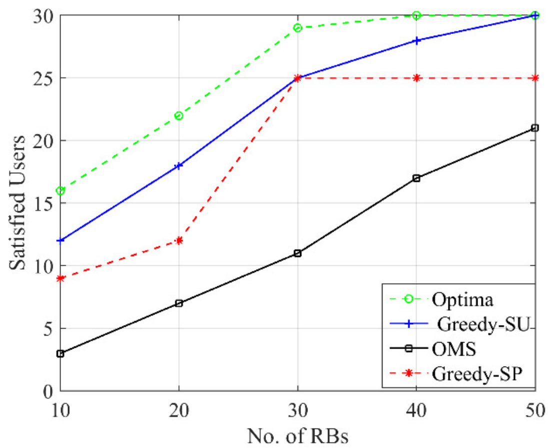


FIGURE 11. Satisfied Use's Count for fixed UEs.

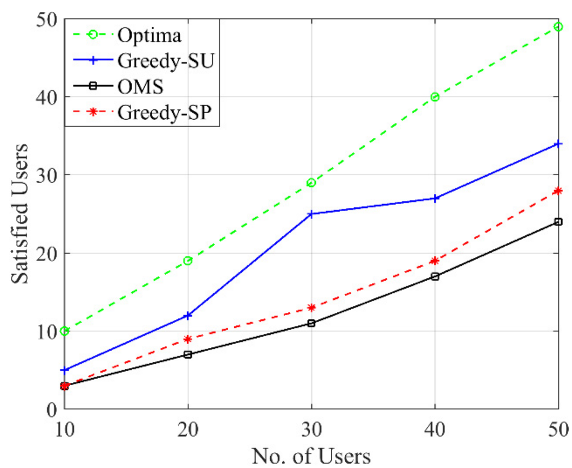


FIGURE 12. Satisfied Use's Count for fixed RB.

C. SIMULATION RESULTS FOR MAXIMIZING SATISFIED USERS

The simulation results of the problem of maximizing the number of satisfied users are shown in Figs. 11 and 12. In Fig. 11, we fix the number of users to be 30 and vary the number of RBs to determine the number of users whose data request has been met entirely. Here, Greedy-SU represents our proposed greedy algorithm in Section IV. D, while Greedy-SP represents the greedy algorithm in Section IV. A when all user profits are set to 1. The number of satisfied users increases linearly at the beginning, and the increment gradually slows down when the number of RBs reaches a certain value (30). This is because the high priority users will be chosen at the beginning, and the supplied RBs will be used by them efficiently to satisfy more users. Later, as we increase the number of RBs further, almost every user by then have already got satisfied. Nevertheless, our proposed

Greedy-SU has better performance than Greedy-SP due to its priority rule. In the case of OMS, the curve increases almost linearly; however, the performance is worse than both our greedy algorithms.

In Fig. 12, we measure the number of satisfied users for the case of fixed RBs = 25, and a varying number of users. Again, our proposed Greedy-SU algorithm gives the best performance. We observe that, after input user set exceeds 30 users, the number of satisfiable users increases slowly of Greedy-SU. This is because later on, those user's RB requirement exceeds the available RBs; hence they cannot be satisfied by our greedy algorithm. Greedy-SP and OMS give similar performance, where the satisfied users count increases linearly, though Greedy-SP is still (slightly) better than OMS.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have considered the problem of revenue maximization of the telecom operator when users have different data request rates, profit, and CQI. We proved this problem is NP-hard and proposed three algorithms; one is a greedy heuristic while the other two has approximation ratio guarantees. All the three algorithms do better while comparing to candidate algorithms regarding profit, fairness index, and satisfied data rate in all the scenarios. We have addressed the special case for maximizing the number of satisfied users and provided another greedy heuristic. It has a higher number of satisfied users when compared to the other candidate algorithms for the case of fixed UEs and fixed RBs.

In our future work, we intend to address the problem of maximizing the profit of maximally satisfied users, which is an NP-hard problem, too. The said objective has significance when the operator has limited resources and to decide whether to satisfy more users while compromising the overall revenue or to maximize the revenue while satisfying fewer users. This will be important for the telecom operator to achieve a balance between these two parameters to survive in the competitive market.

REFERENCES

[1] (2016). *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update: 2016-2021*. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html

[2] 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. 2016.

[3] X. Chen, B. Proulx, X. Gong, and J. Zhang, "Exploiting social ties for cooperative D2D communications: A mobile social networking case," *IEEE/ACM Trans. Netw.*, vol. 23, no. 5, pp. 1471–1484, Oct. 2015.

[4] Y. Xiao, K.-C. Chen, C. Yuen, and L. A. DaSilva, "Spectrum sharing for device-to-device communications in cellular networks: A game theoretic approach," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DYSPAN)*, McLean, VA, USA, Apr. 2014, pp. 60–71.

[5] L. Lei, Z. Zhong, C. Lin, and X. Shen, "Operator controlled device-to-device communications in LTE-advanced networks," *IEEE Wireless Commun.*, vol. 19, no. 3, pp. 96–104, Jun. 2012.

[6] B. Zhou, H. Hu, S.-Q. Huang, and H.-H. Chen, "Intracluster device-to-device relay algorithm with optimal resource utilization," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2315–2326, Jun. 2013.

[7] C. W. Yeh, G. Y. Lin, M. J. Shih, and H. Y. Wei, "Centralized interference-aware resource allocation for device-to-device broadcast communications," in *Proc. IEEE Int. Conf. Internet Things (iThings)*, Sep. 2014, pp. 304–307.

[8] C. P. Lau, A. Alabbasi, and B. Shihada, "An efficient live TV scheduling system for 4G LTE broadcast," *IEEE Syst. J.*, vol. 11, no. 4, pp. 2737–2748, Dec. 2017.

[9] N. Cassiau and D. Ktenas, "Satellite multicast for relieving terrestrial eMBMS: System-level study," in *Proc. IEEE 82nd Veh. Technol. Conf. (VTC Fall)*, Boston, MA, USA, Sep. 2015, pp. 1–5.

[10] I. F. Akyildiz, S. Nie, S.-C. Lin, and M. Chandrasekaran, "5G roadmap: 10 key enabling technologies," *Comput. Netw.*, vol. 106, pp. 17–48, Sep. 2016.

[11] R. O. Afolabi, A. Dadlani, and K. Kim, "Multicast scheduling and resource allocation algorithms for OFDMA-based systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 240–254, 1st Quart., 2013.

[12] T.-P. Low, M.-O. Pun, Y.-W. P. Hong, and C.-C. J. Kuo, "Optimized opportunistic multicast scheduling (OMS) over wireless cellular networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 791–801, Feb. 2010.

[13] X. Lin, R. Ratasuk, A. Ghosh, and J. G. Andrews, "Modelling, "Modeling, analysis, and optimization of multicast device-to-device transmissions," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4346–4359, Aug. 2013.

[14] S. Mitrofanov, A. Anisimov, and A. Turlikov, "EMBMS LTE usage to deliver mobile data," in *Proc. 6th Int. Congr. Ultra Modern Telecommun. Control Syst. Workshops (ICUMT)*, Oct. 2014, pp. 60–65.

[15] L. Militano, M. Condoluci, G. Araniti, A. Molinaro, A. Iera, and G. M. Muntean, "Single frequency-based device-to-device-enhanced video delivery for evolved multimedia broadcast and multicast Services," *IEEE Trans. Broadcast.*, vol. 61, no. 2, pp. 263–278, Jun. 2015.

[16] J.-P. Sheu, C.-C. Kao, S.-R. Yang, and L.-F. Chang, "A resource allocation scheme for scalable video multicast in WiMAX relay networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 1, pp. 90–104, Jan. 2013.

[17] U. K. Sarkar, P. P. Chakrabarti, S. Ghose, and S. C. De Sarkar, "A simple 0.5-bounded greedy algorithm for the 0/1 knapsack problem," *Inf. Process. Lett.*, vol. 42, pp. 173–177, May 1992.

[18] N. K. Pratas and P. Popovski, "Network assisted device to device (D2D) direct proximity discovery with underlay communication," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Diego, CA, USA, Dec. 2015, pp. 1–6.

[19] Z. Zhou, M. Dong, K. Ota, B. Gu, and T. Sato, "Stackelberg-game based distributed energy-aware resource allocation in device-to-device communications," in *Proc. IEEE Int. Conf. Commun. Syst. (ICCS)*, Nov. 2014, pp. 11–15.

[20] L. song, D. Niyato, Z. Han, and E. Hossain, "Game-theoretic resource allocation methods for device-to-device (D2D) communication," *IEEE Wireless Commun. Mag.*, vol. 21, no. 3, pp. 136–144, Jun. 2014.

[21] C. Xia, S. Xu, and K. S. Kwak, "Resource allocation for device-to-device communication in LTE—A network: A stackelberg game approach," in *Proc. IEEE 80th Veh. Technol. Conf. (VTC-Fall)*, Vancouver, BC, Canada, Sep. 2014, pp. 1–5.

[22] C. Xu, L. Song, D. Zhu, and M. Lei, "Subcarrier and power optimization for device-to-device underlay communication using auction games," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, NSW, Australia, Jun. 2014, pp. 5526–5531.

[23] Z. Zhou, M. Dong, K. Ota, R. Shi, Z. Liu, and T. Sato, "Game-theoretic approach to energy-efficient resource allocation in device-to-device underlay communications," *IET Commun.*, vol. 9, no. 3, pp. 375–385, Dec. 2014.

[24] L. Zhang, Z. He, K. Niu, B. Zhang, and P. Skov, "Optimization of coverage and throughput in single-cell E-MBMS," in *Proc. IEEE 70th Veh. Technol. Conf. (VTC-Fall)*, Barcelona, Spain, Sep. 2009, pp. 1–5.

[25] A. Alexious, C. Bouras, V. Kokkinos, A. Papazois, and G. Tsichritzis, "Spectral efficiency performance of MBSFN-enabled LTE networks," in *Proc. IEEE Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Oct. 2010, pp. 361–367.

[26] L. Militano, M. Condoluci, G. Araniti, and A. Iera, "Multicast service delivery solutions in LTE-Advanced systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2013, pp. 5954–5958.

[27] L. Militano, D. Niyato, M. Condoluci, G. Araniti, A. Iera, and G. M. Bisci, "Radio resource management for group-oriented services in LTE-A," *IEEE Trans. Veh. Technol.*, vol. 64, no. 8, pp. 3725–3739, Aug. 2015.

[28] G. Araniti, M. Condoluci, A. Iera, A. Molinaro, J. Cosmas, and M. Behjati, "A low-complexity resource allocation algorithm for multicast service delivery in OFDMA networks," *IEEE Trans. Broadcast.*, vol. 60, no. 2, pp. 358–369, Jun. 2014.

[29] G. Araniti, M. Condoluci, L. Militano, and A. Iera, "Adaptive resource allocation to multicast services in LTE systems," *IEEE Trans. Broadcast.*, vol. 59, no. 4, pp. 658–664, Dec. 2013.

[30] J. Chen, M. Chiang, J. Erman, G. Li, K. K. Ramakrishnan, and R. K. Sinha, "Fair and optimal resource allocation for LTE multicast (eMBMS): Group partitioning and dynamics," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, Kowloon, Hong Kong, May 2015, pp. 1266–1274.

[31] G. Zhang, K. Yang, P. Liu, and J. Wei, "Power allocation for full-duplex relaying based D2D communication underlaying cellular networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4911–4916, Oct. 2015.

[32] R. Cohen and G. Grebla, "Multi-dimensional OFDMA scheduling in a wireless network with relay nodes," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, Toronto, ON, Canada, May 2014, pp. 2427–2435.

[33] R. Cohen and G. Grebla, "Multidimensional OFDMA scheduling in a wireless network with relay nodes," *IEEE/ACM Trans. Netw.*, vol. 23, no. 6, pp. 1765–1776, Dec. 2015.

[34] *Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN)*, document TS 36.300, 3GPP, Sep. 2012.

[35] H.-L. Chao, C.-K. Chang, and C.-L. Liu, "A novel channel-aware frequency-domain scheduling in LTE uplink," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2013, pp. 917–922.

[36] M. Rupp, "Vienna LTE-A simulators," LTE System Level Simulator Documentation, Vienna Univ. Technol., Vienna, Austria, Tech. Rep. v1.2r310. [Online]. Available: https://www.nt.tuwien.ac.at/research/mobile-communications/vccs/vienna-lte-a-simulators/

[37] R. B. Jagadeesha, J.-P. Sheu, and W.-K. Hon, "User satisfaction based resource allocation schemes for multicast in D2D networks," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2017, pp. 1–5.

JAGADEESHA R. BHAT (S'17) received the M.Tech. degree in digital electronics and communication from Visvesvaraya Technological University, India. He is currently pursuing the Ph.D. degree with the Department of Computer Science, National Tsing Hua University, Taiwan. His research interest includes wireless networks, mobile computing, and next generation networks.

JANG-PING SHEU (S'85–M'86–SM'98–F'09) received the B.S. degree in computer science from Tamkang University, Taiwan, in 1981, and the M.S. and Ph.D. degrees in computer science from National Tsing Hua University, Taiwan, in 1983 and 1987, respectively. He is currently a Chair Professor with the Department of Computer Science, National Tsing Hua University.

He was the Chair of the Department of Computer Science and Information Engineering, National Central University, from 1997 to 1999, and the Director of the Computer Center, from 2003 to 2006. He was the Director of the Computer and Communication Research Center, National Tsing Hua University from 2009 to 2015. He was an Associate Dean of the College of Electrical and Computer Science, National Tsing Hua University from 2016 to 2017. His current research interests include wireless communications, mobile computing, and software-defined networks. He was an Associate Editor of the IEEE Transactions on Parallel and Distributed Systems and the *International Journal of Sensor Networks*. He is an Advisory Board Member of the *International Journal of Ad Hoc and Ubiquitous Computing* and the *International Journal of Vehicle Information and Communication Systems*.

He is a member of the Phi Tau Phi Society. He was a recipient of the Distinguished Research Awards of the National Science Council, China, from 1993 to 1994, from 1995 to 1996, and from 1997 to 1998, the Distinguished Engineering Professor Award of the Chinese Institute of Engineers in 2003, the K.-T. Li Research Breakthrough Award of the Institute of Information and Computing Machinery in 2007, the Y. Z. Hsu Scientific Chair Professor Award in 2009, the Pan Wen Yuan Outstanding Research Award in 2014, the Academic Award in Engineering from the Ministry of Education, and the Medal of Honor in Information Sciences from the Institute of Information and Computing Machinery, in 2016 and 2017, respectively.

WING-KAI HON received the bachelor's degree (Hons.) in computer science and the master's and Ph.D. degrees under the supervision of T.-W. Lam, from The University of Hong Kong, Hong Kong, in 1997, 2000, and 2005, respectively. During his Ph.D. studies, he visited the National University of Singapore, under the guidance of W.-K. Sung, for one year. He was a Post-Doctorate Researcher with Purdue University, West Lafayette, IN, USA, under the supervision of J. Vitter.

He is currently an Associate Professor with the Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan. His current research interests include indexing, data compression, external memory data structures, and combinatorial optimization.