Relay-Aided Downlink Data Broadcast in LTE-Advanced or 802.16m WiMAX-Based Wireless Networks

Cheng-Hsien Lin, Ren-Hung Hwang, Jeng-Farn Lee, Chuan-Kang Ting
Department of Computer Science and Information Engineering, National Chung Cheng University, Taiwan
vexonelit@gmail.com, {rhhwang, jflee, ckting}@cs.ccu.edu.tw

Abstract

In LTE-Advanced or 802.16m WiMAX-based relay networks, the downlink broadcast service requires the network to decide not only the data path but the amount of resource that should be reserved for the transmissions once a broadcast session is triggered. This is done by allocating the scarce resource to the Base Station (BS) and a subset of Relay Stations (RSs) such that all subscribed Mobile Stations (MSs) can receive data. Resolving this allocation problem is challenging since selecting a different set of RSs and allocating different amounts of the resource to them would yield different network topologies and thus diverse data paths from the BS to the MSs. Existing works either do not specialize in downlink broadcast scenario or provide solutions that are relied on some specific prerequisites. This motivates us to formulate the studied problem as an integer linear programming (ILP) model. Considering the computational complexity of ILP, we proposed two polynomial-time algorithms, the Resource Diminishing Principle (RDP) and Enhanced-RDP. Their performance was evaluated through a serious of simulations. The results showed that the E-RDP algorithm boosts the resource utilization by 5 ~ 59%; and outperforms existing methods in terms of the resource utilization by 31 ~ 66%. This paper further conducted a performance analysis on the worst-case performance of the E-RDP algorithm.

Keywords: Broadcast, Path construction, Relay networks, Resource allocation.

1 Introduction

Internet access is now ubiquitous. People make phone calls or access social networks (e.g., Facebook) through smart-phones or tablets anytime, anywhere. Broadband wireless technology is being improved constantly. Two leading standards, 3GPP LTE-Advanced and 802.16m WiMAX, provide high data rates for real-time applications (e.g., Internet Protocol television (IPTV) or video conferencing) that need to transmit massive amounts of data. However, it is less likely that a single base station (BS) could cover an enormous area and with an acceptable data rate due to the deployment cost of wireless infrastructures, obstacles, the fading channel, the mobility pattern, etc. Thus, both standards adopt relay stations (RSs) [1], which can collaborate with the BS to improve the overall throughput and coverage [2-3].

In this paper, we focus on the downlink broadcast of massive amounts of data over an LTE-Advanced or IEEE 802.16m WiMAX based wireless relay network, operating in non-transparent, in-band mode. Although we take real-time multimedia transmission as the target application, the proposed method is applicable to other applications that transmit large amounts of data. The network in a cell is typically comprised of one BS, multiple RSs, and multiple mobile stations (MSs), which subscribe to a real-time multimedia application. An MS can receive data either from the BS directly or via one or more RSs relaying data to it. Before delivering data to MSs, an RS must receive a copy of the data from the BS. In this paper, we stipulate that an RS can only assist the BS or MSs to forward traffic; in other words, it cannot relay traffic to or from other RSs.

In general, the network needs to decide not only the data path but the amount of resource that should be reserved for the transmissions once a broadcast session is triggered. The amount of the resource required for a successful transmission between a sender and a receiver depends on the modulation and coding scheme (MCS) chosen in response to the channel condition (e.g., suppose that the data rate of a real-time multimedia is M, and that the corresponding transmission rate of chosen MCS is b. Then the amount of resource required to the sender is M/b). In addition, the chosen MCS also decides the coverage range of a sender. When there are multiple receivers that wish to get the same data, a substantial amount of the resource can be saved by the sender broadcasting the data only once, rather than unicasting the same data multiple times. However, the sender must select a suitable MCS to ensure that all designated receivers receive the content successfully. More resource is needed when a more robust MCS is chosen. It is crucial that the resource is utilized efficiently to deliver the real-time multimedia to all the designated MSs since the radio resource such as the bandwidth, resource blocks, and time-slots is shared by all network components in the same cell. We observe that resolving the resource allocation problem can also determine the data path from the BS to all the designated

*Corresponding author: Ren-Hung Hwang; E-mail: rhhwang@cs.ccu.edu.tw
DOI: 10.6138/JIT.2016.17.1.20130330

MSs. That is, the problem to find a subset of RSs and allocate each of them different amount of resource (or selecting different MCS for the BS and RSs) such that all designated MSs can be served and the total required resource is minimized. The resource allocation problem is challenging since selection of RSs and allocation of resource affect network topologies and thus routing paths from the BS to the MSs.

Several approaches have been proposed to resolve the resource allocation problem in different environments and to satisfy diverse requirements. First, the works [4-7] focus on the downlink unicast resource allocation issue. Their objective is to find the optimum resource allocation for the BS and a set of RSs that will maximize the system throughput and service range. On the other hand, the studies [8-9] address the uplink many-to-one unicast resource allocation issue, where multiple MSs send data to the BS. Since these works do not address the path construction and resource allocation tasks simultaneously, their approaches can only be used in unicast applications but not suitable for downlink broadcast purposes. Second, the studies [10-11] investigated the recipient maximization problem under the limited resource constraint. The objective is to allocate the limited resource to the BS and a subset of RSs efficiently so that the largest number of subscribers can receive the requested multimedia content. However, their approaches may either not be applied on directly or allocate resource in a less efficient way. For example, Dynamic Station Selection (DSS) proposed in [10] only runs on the network whose topology has already been known, whereas the inflexible allocation strategy of Multicast Subscriber Selection (MSS) proposed in [11] might cause the fraction of the allocated resource to be wasted. Moreover, MSS assigns the entire resource budget to the BS in the initial step. Nevertheless, we do not know how many resource allocated to the BS is sufficient when employing MSS in the initial step. Third, the minimum resource broadcast issue in wireless ad-hoc networks has been well-studied [12-21], where any node can act as a relay in wireless ad-hoc networks. However, unlike nodes in wireless ad-hoc networks, only RSs can help forward traffic in cellular-based relay networks. This difference prevents us adopting the proposed solutions. Moreover, in practice, the performance of existing solutions may not be good enough. For example, approximation algorithms guarantee that their performance is always within a certain bound; however, it is not necessary mean they do perform well in most cases. More importantly, some of their results are based on specific assumptions or properties (e.g., the algorithm proposed in [22] needs to employ the approximation algorithm devised in [23] in order to obtain a Steiner tree. Nevertheless, the algorithm requires a great amount of computing time to yield a Steiner tree according to [24]).

This paper builds an integer linear programming (ILP) model for the downlink broadcast resource allocation problem in that no existing models and methods are suitable for our purpose. Since the ILP model belongs to the NP class, it is difficult to compute the optimal solution in a short time. To resolve the downlink broadcast path construction problem, we propose an approach called the Resource Diminishing Principle (RDP) algorithm which can output a solution rapidly and does not require specific assumptions or properties. Furthermore, this paper presents the Enhanced-RDP (E-RDP) to boost the overall resource utilization by preventing the selected RSs from being allocated more resource than actually needed. The E-RDP incorporates an auxiliary rule that is used to regulate the serving range of the selected RSs. The results on a series of simulations demonstrate that E-RDP can achieve satisfactory performance with various parameter settings, which validate its robustness. In addition, our analysis shows that the worst performance of E-RDP is bounded by H_N of the optimal solution under a given assumption, where H_N stands for the Harmonic series, and N is the number of designated MSs.

The remainder of this paper is organized as follows. Section 2 describes the system model and the notations used in the paper. The section also discusses the integer linear programming model used to represent the downlink broadcast path construction problem. Section 3 describes the components of E-RDP in detail and provides the complexity analysis and the performance analysis. In Section 4, we present the simulation results and assess the performance of E-RDP. Section 5 contains some concluding remarks.

2 System Model and the ILP Modeling

This section presents the network model and notions. An integer linear programming formulation is proposed to model the downlink broadcast resource allocation problem.

2.1 System Model and Notions

This paper considers a wireless relay network comprised of one BS, M RSs, and N MSs. To distinguish between the network nodes, each one is assigned a unique integer as its identification (ID). Let the ID of the BS be 0, the IDs of the RSs range from 1 to M, and the IDs of the RSs range from $M + 1$ to $M + N$. In addition, assume that a real-time multimedia streaming must be broadcast from the BS to all MSs via the relay network. An MS receives the streaming data from the BS or an RS, depending on the result of the path construction mechanism. To form a topology for the multimedia broadcast, the system must find a suitable set of RSs and allocate the resource to the BS and RSs appropriately. In this paper, we do not consider the spatial reuse, and assume that the BS and RSs would be scheduled on disjoint radio resources to avoid the intra interference. The radio resource is viewed as a general form that is representative of any type of resource used in existing radio technologies. For example, both LTE-A or 802.16m WiMAX adopt OFDMA as the multiple access mechanism. The resource in this context refers to the total number of resource blocks used to broadcast the streaming data. For a downlink transmission between a sender i (i.e., the BS or an RS) and a receiver j (i.e., an RS or MS), the resource required by the sender i is denoted by $R_{i,j}$, which is determined by the MCS chosen by i in response to the channel condition feedback from the receiver j. Let $\varepsilon = [\varepsilon_0, \varepsilon_1, ..., \varepsilon_M, \varepsilon_{M+1}, ..., \varepsilon_{M+N}]^T$ be a column vector, where $\varepsilon_i = [R_{i,0}, R_{i,1}, ..., R_{i,M}, R_{i,M+1}, ..., R_{i,M+N}]$ is a row vector that records the amount of resource required for a successful transmission from i to j. Note that some $R_{i,j}$ values are zero since their transmission links are disallowed in the problem setting (i.e., an RS cannot send traffic to the BS or other RSs, and MSs can only receive traffic). Table 1 lists the notations used throughout the paper.

2.2 Integer Linear Programming Model

The proposed integer linear programming formulation is based on the network flow model. The goal of the path construction problem is to yield a network topology with the minimum resource consumption on the BS and the set of selected RSs such that all designated MSs can receive the broadcast data. In the problem model, there are three basic variables for $0 \le i, j \le M + N$: (1) $X_{i,j} \in \{0, 1\}$ are binary variables: $X_{i,j} = 1$ if the transmission from node i to node j is used in the final solution, and 0 otherwise; (2) $F_{i,j} \in Z$ are flow variables that represent the amount of flow passing through link (i, j); (3) $Y_i \in R$ stands for the amount of the resource the system allocates to node i. In addition, we let $V = \{0, 1, ..., M, M + 1, ..., M + N\}$ denotes the set of all network nodes, let s denote the source node, and let D denotes the set of intended destination nodes (in our context, s is the BS and $D = \{M + 1, ..., M + N\}$ stands for all MSs).

The ILP is defined as follows.

$$\min \textstyle\sum_{\forall i} Y_i. \tag{1}$$

subject to

$$Y_i - R_{i,j}X_{i,j} \ge 0, \forall i, j \in V, i \ne j, \tag{2}$$

$$L_D X_{i,j} - F_{i,j} \ge 0, \forall i, j \in V, i \ne j, \tag{3}$$

Table 1 The Notations Used Throughout the Paper

M	The number of RSs (each of them is denoted by i, $1 \le i \le M$)
N	The number of MSs (each of them is denoted by i, $M + 1 \le i \le M + N$)
$R_{i,j}$	The resource required by the sender i which chooses the MCS in response to the channel condition feedback from the receiver j
ε	The column vector ($e = [\varepsilon_0, \varepsilon_1, ..., e_M, e_{M+1}, ..., e_{M+N}]^T$)
ε_i	The row vector ($\varepsilon_i = [R_{i,0}, R_{i,1}, ..., R_{iM}, R_{i,M+1}, ..., R_{i,M+N}]$)
$X_{i,j}$	The binary variable ($X_{i,j} = 1$ if the transmission form node i to node j is used in the final solution, and 0 otherwise)
$F_{i,j}$	The flow variable (integer numbers) that represents the amount of flow passing through link (i, j)
$\mathbf{Y}$	The allocated resource of the system ($\mathbf{Y} = [Y_0, Y_1, ..., Y_M, Y_{M+1}, ..., Y_{M+N}]$)
Y_i	The amount of the resource the system allocated to node i
V	The set of all network nodes
D	The set of intended destination nodes (in our context, $D = \{M + 1 \le i \le M + N\}$
L_D	The cardinality of set D
Λ	The set of intended destination nodes that have yet to receive the multimedia streaming data
s	The source node (in our context, s is the BS)
Γ	The set of RSs ($\Gamma = V - (D \cup \{s\})$)
$a_{i,j}$	The additional resource required to the sender i such that the receive j can receive the data form i successfully
P_H	A set of RSs with a higher priority (P_L stands for the set of low priority RSs)
λ	A specific threshold that is used to classify RSs into P_H or P_L
θ	A parameter that is used to control the specific threshold λ

$$L_D = \sum_{j=0}^{M+N} F_{i,j};\ i = \text{source},\ j \ne i, \tag{4}$$

$$0 = \sum_{j=0}^{M+N} F_{j,i};\ i = \text{source},\ j \ne i, \tag{5}$$

$$\sum_{j=0}^{M+N} F_{j,i} - \sum_{j=0}^{M+N} F_{i,j} = 0;\ \forall i \in V - D - \{\text{source}\},\ j \ne i, \tag{6}$$

$$0 = \sum_{j=0}^{M} F_{i,j};\ \forall i \in V - D - \{\text{source}\},\ j \ne i, \tag{7}$$

$$\sum_{j=0}^{M+N} F_{j,i} = 1;\ \forall i \in D,\ j \neq i, \tag{8}$$

$$0 = \sum_{j=0}^{M+N} F_{i,j};\ \forall i \in D,\ j \neq i, \tag{9}$$

$$X_{i,j} \in \{0, 1\}; \forall i, j \in V, i \neq j, \tag{10}$$

$$F_{i,j} \geq 0; \forall i, j \in V, i \neq j. \tag{11}$$

The objective function in Equation (1) is to minimize the sum of resource allocated to the set of selected RSs. The constraints stated in Equation (2) is defines the relations between Y_i and $X_{i,j}$, where $R_{i,j}$ is the (i, j)-th element in the vector ε. In a solution where a sender i broadcasts data to a set of receivers, this constraint ensures that the system will allocate enough of the resource to node i such that the set of receivers can receive the data. Equation (3) represents the constraint that connects the relations between $X_{i,j}$ and $F_{i,j}$, and ensures that $X_{i,j} = 1$ if $F_{i,j} > 0$, where LD is the cardinality of set D.

Moreover, we define the flow control constraints for s, the transit nodes (the RSs), and the destination nodes. For the source node s, Equation (4) requires that s should generate exactly LD units of flow and direct them through a set of out-going links. Meanwhile, Equation (5) states that no flow should pass through or gather at s. Since the RSs are only responsible for relaying flow, Equation (6) is used to model that the RSs do not produce or store any flow. Furthermore, Equation (7) states that it is impossible for flow to pass through more than one RS. Regarding each destination node, Equation (8) ensures that it keeps exactly one unit of flow while Equation (9) ensures that it does not forward that flow to other nodes.

Finally, Equations (10) and (11) express the integrality of the binary variables and the non-negativity of the flow variables, respectively. Since the proposed ILP model belongs to the NP class, it is difficult to achieve the optimal solution in a short time. Thus, we propose an E-RDP algorithm to resolve the downlink broadcast resource allocation problem.

3 Enhanced-Resource Diminishing Principle (E-RDP) Algorithm

In this section, we discuss the proposed E-RDP algorithm, which provides a simple yet efficient way to construct a data path towards all the designated MSs for a cellular-based relay network. The first sub-section describes the ordinary RDP algorithm in detail; and the second sub-section introduces a helpful rule to boost the algorithm's performance. Finally, we present a complexity analysis and performance analysis of the E-RDP algorithm.

3.1 The RDP Algorithm

Before discussing the rationale and operations of the RDP algorithm, we define the notations used in the remainder of the paper. Let Λ denote the set of MSs that have not received the multimedia streaming data yet. By specifying the BS s as the source node, the set of RSs Γ can be determined by $V - (D \cup \{s\})$.

The rationale behind the RDP algorithm is that the total resource consumption can be reduced by utilizing specific RSs to send the multimedia streaming data to some MSs, instead of requiring the BS to broadcast the data to all MSs. Moreover, since we need to serve all designated MSs in the network, the resource should be first allocated to the BS and an RS in order to serve the MS that requires the BS the most additional resource to serve it. This step of allocation can help to serve multiple MSs simultaneously to boost resource utilization. In each round (iteration), RDP first selects a pair of nodes (u, v) where $v \in \Lambda$ is an MS that costs the BS the most additional resource to serve currently; and u is the RS that requires the lowest resource cooperates with the BS to serve MSv. Then, RDP performs two key operations. First, it allocates the appropriate amount of resource to the BS and the RS_u (by updating Ys and Yu). Next, it determines which MSs in Λ can be served by the BS, the RS_u, or both. Once an MS has been served, RDP removes it from the set Λ. The algorithm repeats the above operations until Λ is empty, i.e., the paths from the source to all destination nodes have been assigned.

3.2 Enhancing the RDP Algorithm

In each round the chosen (u, v) pair is a local optimum, which may allocate more resource than that of the optimal solution to the chosen RS_u. If the service range of the BS extends to a certain level such that each RS is able to cover a relatively small area of the rest network, the resource utilization can be much efficient than that RSs have a relatively larger service range. Taking the wireless relay network in Figure 1 as example, Figure 1(a) shows the design discipline mentioned above, where RDP would choose RS_1 and RS_2 to serve MS_1 and MS_2, MS_3 respectively, for the local optimum in the steps of resource allocation. However, as shown in Figure 1(b), if RDP chooses RS_3 to serve MS_1, it would save a substantial amount of resource since the BS essentially needs to cover RS_2. To resolve this problem, we use an auxiliary rule which gives a higher priority to the RSs whose required resource to serve MS_v is within a specified threshold λ. Consequently, in each round the resultant E-RDP will first consider the RSs that have higher priority when choosing the suitable RS_u. In our performance evaluation, we introduce a parameter θ defined by the ratio of the radius to the BS's radio range to reflect the value of λ.

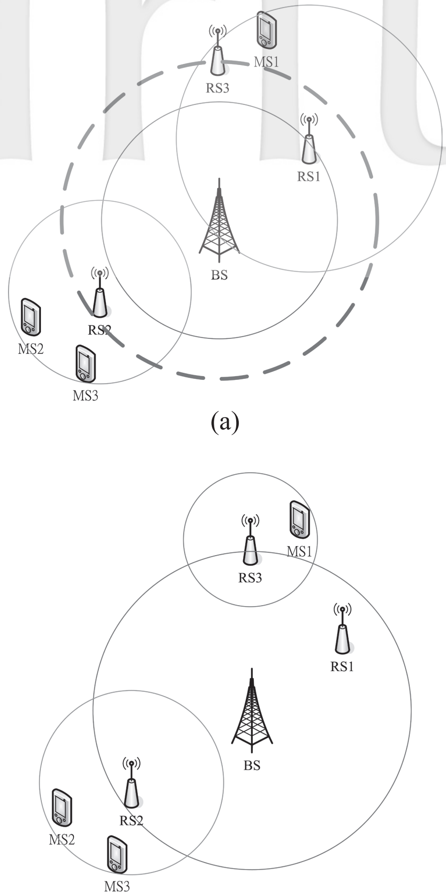


Figure 1 The Influence of the Proposed Enhancement Rule: (a) RDP without the Enhancement Rule; (b) RDP with the Enhancement Rule (E-RDP)

3.3 Complexity of the E-RDP Algorithm

The E-RDP algorithm is summarized in Figure 2. The input to E-RDP is ε, V, D, λ, and the source node. In the initialization process, E-RDP initializes the crucial parameters, and takes $O(M + N)$ to complete the process. Next, E-RDP executes steps 1 ~ 4 repeatedly until all designated MSs have been served. In the beginning of each round, E-RDP takes $O(M)$ to divide the set of RSs into two groups based on the given threshold λ. In step 2, E-RDP seeks for a pair of nodes (RS_u, MS_v), where it takes $O(N)$ time on searching for the $MS_v \in \Lambda$ and $O(M)$ time on identifying the RS_u that cooperates with the BS to serve MSv such that the additional resource requirement is the lowest among all RSs. Next, it takes $O(1)$ to allocate the amount of additionally required resource to RS_u and the BS in step 3. In step 4, E-RDP changes the state of MS_j ($j \in \Lambda$) which can be served in this round, by removing MS_j from the set Λ. The time complexity of step4 is $O(N)$. Since E-RDP repeats the above operations at most N times, the time complexity of E-RDP is $O(MN + N^2)$, which is polynomial.

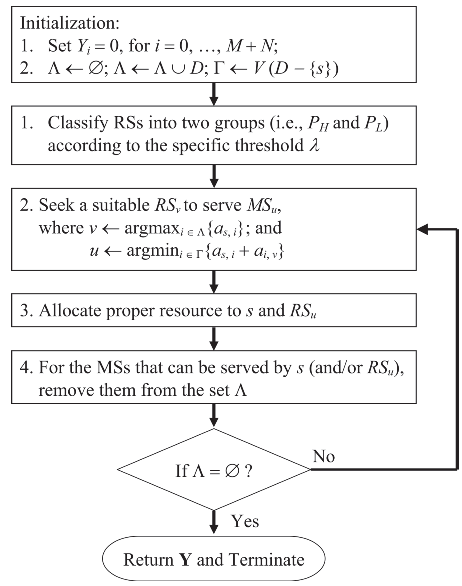


Figure 2 The Flow Chart of E-RDP

3.4 Performance Analysis of the E-RDP Algorithm

The performance analysis of E-RDP considers the following situation. Let A denote the set of MSs which have been served by the relay network. In each round of resource allocation process, E-RDP yields a sub-tree which connects at least one MS. Assume that in a single round (namely the i-th round) the set of MSs $U_i \subseteq D$ is served by E-RDP. Further, we define the resource utilization associated with the MSs which belong to U_i as $\frac{c(U_i)}{n_i}$, where $c(U_i)$ denotes the sum of additional resource required by the selected RS_u and the BS; and $n_i = |U_i - A|$ is the number of unserved MSs which belong to U_i. Thus, in the i-th round, the j-th to $(j + n_i)$-th MS would be served by the network, each of them costs the relay network $\frac{c(U_i)}{n_i}$ units of resource. Let $c(E\text{-}RDP)$ be the total resource allocated to the set of selected RSs (including the BS) in the solution obtained from E-RDP. Assume that E-RDP executes σ rounds and outputs the solution, then we have $c(E\text{-}RDP) = \sum_{i=1}^{\sigma} c(U_i) = \sum_{k=1}^{N} c(k)$, where $c(k)$ is the cost to pick up the k-th MS in U_i. Moreover, we assume that the amount of resource allocated to the RSs selected by E-RDP never exceeds λ. Under this assumption, $c(E\text{-}RDP)$ reaches its maximum when each selected RS is allocated λ units of resource. To obtain the performance gap, we need to find the relation between the resource utilization of the k-th MS added in A and the optimal solution.

Lemma 1: Let k be the index that represents the k-th MS being served by E-RDP, and $c(OPT)$ be the total resource allocated to the set of selected RSs (including the BS) in the optimal solution. For $1 \le k \le N$, it holds that $c(k) \le \frac{c(OPT)}{N - k + 1}$.

Proof: We prove the lemma by induction with the induction parameter $k \in [1, N]$.

Basis step: $k = 1$. In each round, the target MS_v requires the BS the most amount of additional resource to serve; and the selected RS_u cooperating with the BS to serve the target MS_v costs the relay network the minimal resource consumption. Hence, $c(k = 1)$ would be the minimal in the first round of E-RDP, and is smaller than $\frac{c(OPT)}{N}$. In other words, if the condition $c(k = 1) > \frac{c(OPT)}{N}$ is true, then we can conclude that either E-RDP is not deterministic or the resource consumption $\frac{c(OPT)}{N}$ is not sufficient for the optimal solution to serve the target MS_v. However, both cases would lead to a contradiction.

Inductive step: $2 \le k \le N$. Suppose that the inequality $c(l) \le \frac{c(OPT)}{N-l+1}$ holds for $2 \le l \le k-1$. Now we prove that the inequality is also valid for $l = k$. If both the l-th and the $(l-1)$-th MSs belong to U_{i-1} to, then $c(l) = c(l-1) = \frac{c(U_{i-1})}{n_{i-1}} \le \frac{c(OPT)}{N-l+2} < \frac{c(OPT)}{N-l+1}$ holds since $f(k) = \frac{c(OPT)}{N-k+1}$ is an increasing function. Now we consider the case in which the l-th MS belongs to U_i. Based on the assumption that the amount of resource allocated to the RSs which are selected by E-RDP never exceeds λ, $c(U_{i-1}) \ge c(U_i)$ when $c(E\text{-}RDP)$ reaches its maximum. Hence we have $c(l) = \frac{c(U_i)}{n_i} \le \frac{c(OPT)}{N-l+1}$.

From the above derivation, we can conclude that for $1 \le k \le N$, $c(k) \le \frac{c(OPT)}{N-k+1}$ holds.

Theorem 2: The upper bound of $c(E\text{-}RDP)$ is never large than $H_N \cdot c(OPT)$ only if the amount of resource allocated to each selected RS never exceeds λ, where H_N stands for the Harmonic series.

Proof: According to the assumption, $c(E\text{-}RDP)$ has the poorest performance when the amount of resource allocated to each selected RS is exactly equal to λ. By applying the lemma 1, we have $\sum_{k=1}^{N} c(k) \le \sum_{k=1}^{N} \frac{c(OPT)}{N-k+1} = c(OPT) \cdot \sum_{k=1}^{N} \frac{1}{k} = c(OPT) \cdot H_N$.

4 Performance Evaluation

This study evaluated the performance of the proposed approach via a series of simulations. To reflect the channel condition between a sender i and a receiver j, we use a two-ray ground propagation model without fading [25]. In the model, the required resource $R_{i,j}$ is inversely proportional to $(cd_{i,j})^{\alpha}$, where $d_{i,j}$ is the distance between the sender i and the receiver j, α is the channel attenuation factor, which is usually set to $\alpha \in [2, 4]$, and c is a constant. Different radio propagation models can be used as long as the resource required by $R_{i,j}$ can be estimated. The simulation results of three α settings yield the same performance patterns and have the same space limitations; therefore, we only consider $\alpha = 3$ in the following discussion. To measure $d_{i,j}$, we deploy network nodes in a circular area on a two-dimensional plane (in Figure 3). The radius of the circular area and the range of the RSs are denoted by r_{SA} and r_{RS} respectively. The BS is placed in center (0, 0) of the circular area, and the MSs are placed randomly in the area. The RSs are deployed randomly in a specific region of $[\Delta, r_{RS}]$, where the parameters Δ and r_{RS} are used to generate two classifications of the RS deployment. [0, 100] is set to represent as a general case that tests the robustness of the proposed approach, while [40, 60] is set to represent as a practical case where the RSs are located in the middle of the BS and the MSs. In addition, we fix the value of c at 1/100 to represent a special case where a set of MSs within 100 units of the BS's radio range can receive the media content broadcast by the BS. Thus, the performance result of a specified approach in the simulations can be seen as a multiple of performance of the special case. The threshold λ for the auxiliary rule is computed as $(\theta c r_{SA})^{\alpha}$ where the value of θ depends on the value of r_{SA}. We choose two sets of values of r_{SA}: $\theta = \frac{1}{2}$ for $r_{SA} = 100$, and $\theta = \frac{2}{3}$ for $r_{SA} = 150$. Table 2 summarizes the parameter settings used in the simulations.

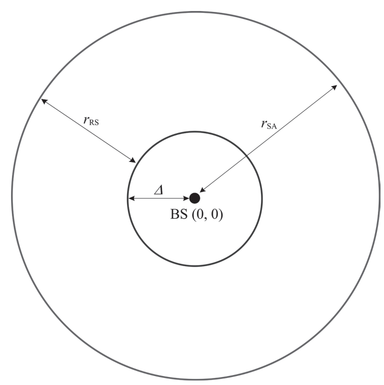


Figure 3 Node Placement in the Simulations

Table 2 The Parameter Settings Used in the Simulations

		(r_{SA}, λ)		
		(100, 0.125)		(150, 1.0)
(Δ, r_{RS})	(0, 100)	# of RS: 5 ~ 9	# of RS: 10 ~ 40	# of RS: 20 ~ 40
		# of MSs: 100	# of MSs: 200	# of MSs: 200
	(40, 60)	# of RS: 5 ~ 9	# of RS: 10 ~ 40	# of RS: 20 ~ 40
		# of MSs: 100	# of MSs: 200	# of MSs: 200

The intension of the simulations is twofold: We are interested in the performance gap among E-RDP and the other approaches, and in the extent that the auxiliary rule can improve. In the simulations, we compare the performance of the E-RDP, RDP, BIP [12] algorithms, the utility-based method (referred to as Utility) and the BF method. BF computes the optimal solution in a brute-force manner, while the Utility method uses a performance-to-cost-ratio criterion (i.e., the number of additional MSs that can be served, divided by the amount of extra resource allocated to RS_i and the BS) to choose a pair of RS_i and MS_j in each round of the allocation process.

Firstly, we examine the performance gap between the proposed solutions and the BF method. We set $r_{SA} = 100$, $\Delta = 0$, $r_{RS} = 100$, $N = 80$, and range the number of RSs from 3 to 7. Here we do not show the results of BF for the number of RSs large than 7 due to its forbidden computation cost. Figure 4 shows that the performance gap between E-RDP and BF is small (less than 3 percent), but the difference between RDP and BF is relatively large. For example, the largest performance gap between RDP and BF is up to 48 percent (between BIP and BF is up to 56%; between Utility and BF is up to 29%), which indicates that the auxiliary rule can effectively assist E-RDP save a substantial amount of resource.

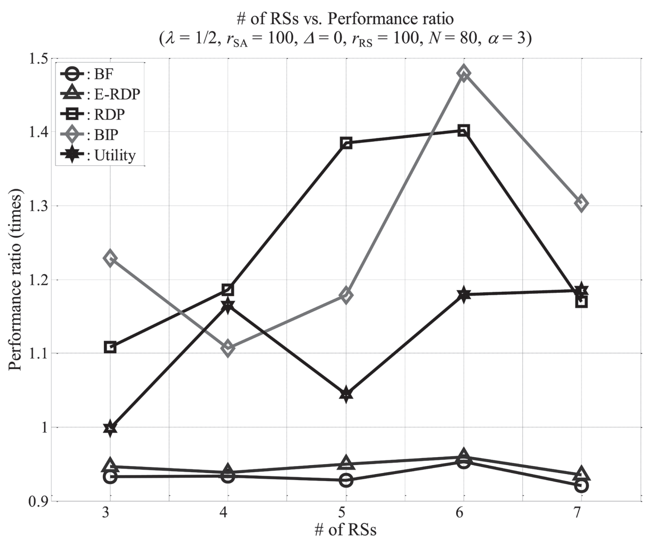


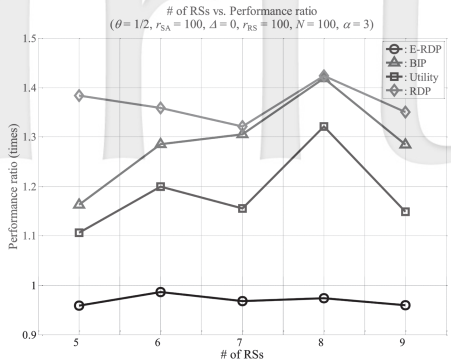
Figure 4 Performance Comparison of BF, E-RDP, RDP, BIP, and Utility Algorithms with Different Numbers of RSs

Next, we examine the performance for different numbers of RSs. We first set $N = 100$ $r_{SA} = 100$, and manipulate the number and distribution of the RSs. Figures 5(a) and 5(b) show the cases for $\Delta = 0$, $r_{RS} = 100$. E-RDP outperforms the other approaches as $5 \leq M \leq 40$. Interestingly, RDP does not gain any advantage over BIP and Utility when $M \leq 10$; however, its performance improves when the number of RSs exceeds 20. The reason is that RDP is less likely to allocate an excessive amount of the resource to an RS because there are more good candidate RSs that can be chosen to serve MS_v. On the other hand, Figures 5(c) and 5(d) show the cases for $\Delta = 40$, $r_{RS} = 60$. We observe that the performance pattern is similar to that in Figures 5(a) and 5(b). In Figure 5(c), the performance of Utility is quite close to that of E-RDP when $M = 6$ and 8. This is because the allocation decisions made by Utility are similar to that of E-RDP; thus the performance gap is rather limited. In Figure 5(d), the increasing number of RSs with their position distribution provides Utility good chance to find a more efficient way to allocate the resource when the number of RSs exceeds 20. From this set of the results, we observed that the auxiliary rule contributes five to fifty-nine percent of performance improvement. E-RDP is able to save two to forty-four percent of the allocated resource compared to the Utility method; and save sixteen to sixty-six percent of the allocated resource compared to the BIP method.

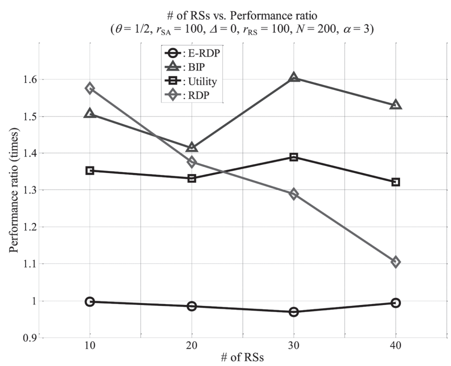
This study also simulates the scenario where a subset of MSs outside the BS's radio range has to receive the media content via the RSs. In the simulations, we set $r_{SA} = 150$, and stipulate that the radius of the BS's radio range is at most 100 units. Figure 6(a) shows the results when $\Delta = 0$ and $r_{RS} = 100$. We observe that the overall resource consumption of E-RDP is the lowest among the compared approaches. Moreover, the consumption decreases as M increases. Figure 6(b) shows a similar performance pattern when $\Delta = 40$ and $r_{RS} = 60$. The performance of BIP and Utility improve when the RSs are placed in favorable positions. Interestingly, the performance of RDP is very close to that of E-RDP. The reason is because under the scenario setting, each selected RS can serve the MSs outside the BS's radio range without requiring the amount of resource that exceeds the specified threshold λ. From this set of the results, the proposed methods can improve the resource utilization by six to twenty-three percent compared to the Utility method; and improve the resource utilization by sixteen to forty-one percent compared to the BIP method. The above results demonstrate that E-RDP works well and it is not affected by the distribution of the RSs. The proposed auxiliary rule makes an important contribution in that it helps E-RDP select suitable RSs without excessive resource consumption.

5 Conclusion

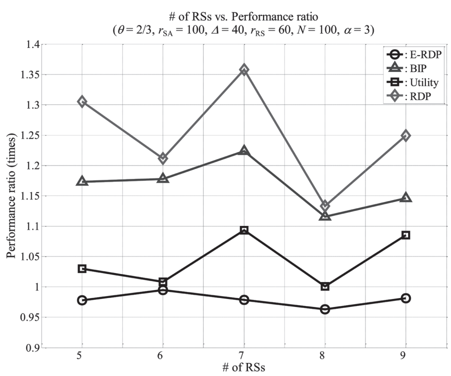
In this paper, we consider the resource allocation problem in the downlink broadcast of massive amounts of data over wireless relay networks. The objective is to find a subset of RSs and to allocate an appropriate amount of resource to each of them such that the BS and all the designated MSs can be connected via the chosen subset of RSs. Meanwhile, the total allocated resource for the BS



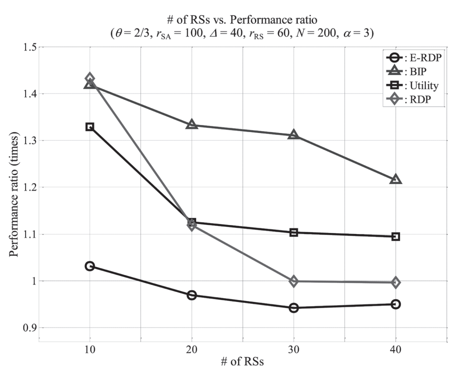
(a)



(b)

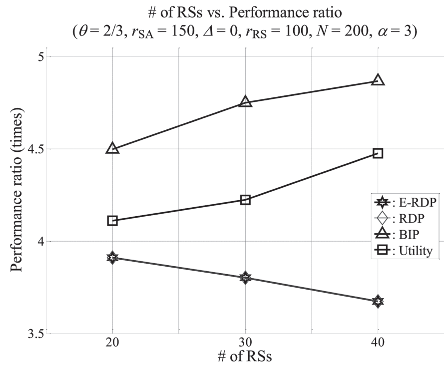


(c)

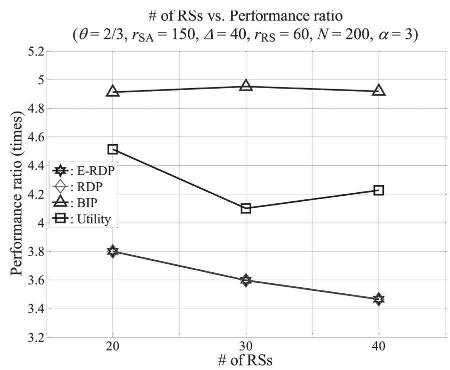


(d)

Figure 5 Performance Comparison of E-RDP, BIP, Utility, and RDP Algorithms, under Different Parameter Settings: (a) $\Delta = 0$, $r_{RS} = 100$, and $5 \leq M \leq 9$; (b) $\Delta = 0$, $r_{RS} = 100$, and $10 \leq M \leq 40$; (c) $\Delta = 40$, $r_{RS} = 60$, and $5 \leq M \leq 9$; (d) $\Delta = 0$, $r_{RS} = 60$, and $10 \leq M \leq 40$



(a)



(b)

Figure 6 Performance Comparison of E-RDP, BIP, and Utility Algorithms under Different Distributions of RSs' as well as M: (a) $\Delta = 0$, $r_{RS} = 100$; (b) $\Delta = 40$, $r_{RS} = 60$

and the selected set of RSs are to be minimized. To resolve this problem, we present an ILP to formulate the studied problem. In view of the computational complexity of ILP, we propose the RDP and the E-RDP algorithms to address the problem. The E-RDP improves the resource utilization of the RDP by introducing an auxiliary enhancement rule that tries to choose the RSs with high priority to cooperate with the BS. The simulation results give evidence to back up the claim that five to fifty-nine percent of the allocated resource can be saved. They also show that the E-RDP performs well in various scenarios and outperforms existing methods in terms of the resource utilization by at most thirty-one to sixty-six percent. Apart from the simulation results, we also conduct a theoretical analysis to illustrate that the performance of E-RDP is bound to the optimal solution.

Some directions remain for future study. First, the spatial reuse issue and the intra-interference issue will be taken into account. A new ILP model for the joint resource allocation and assignment problem will be presented, and an integrated solution will also be proposed. Second, other factors such as carrier aggregation and call-admission control are worthy of further investigation for an integrated solution.

References

[1] K. Loa, C.-C. Wu, S.-T. Sheu, Y. Yuan, M. Chion, D. Huo and L. Xu, IMT-Advanced Relay Standards, *IEEE Communications Magazine*, Vol. 48, No. 8, pp. 40-48, August, 2010.

[2] B. Li, Y. Qin, C. P. Low and C. L. Gwee, A Survey on Mobile WiMAX, *IEEE Communication Magazine*, Vol. 45, No. 12, pp. 70-75, December, 2007.

[3] K. Park, C. G. Kang, D. Chang, S. Song, J. Ahn and J. Ihm, Relay-Enhanced Cellular Performance of OFDMA TDD System for Mobile Wireless Broadband Services, *IEEE 16th International Conference on Computer Communications and Networks*, Honolulu, HI, 2007, pp. 430-435.

[4] J. Cai, X. Shen, J. W. Mark and A. S. Alfa, Resource Allocation in Wireless Relay Networks, *IEEE Global Telecommunications Conference*, San Francisco, CA, 2006, pp. 1-5.

[5] O. Oyman, OFDM2A: A Centralized Resource Allocation Policy for Cellular Multi-hop Networks, *40th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, 2006, pp. 656-660.

[6] M. Kaneko and P. Popovski, Adaptive Resource Allocation in Cellular OFDMA System with Multiple Relay Stations, *IEEE Vehicular Technology Conference Spring*, Dublin, Éire, 2007, pp. 3026-3030.

[7] T. Girici, C. Zhu, J. Agre and A. Ephremides, Optimal Radio Resource Management in Multihop Relay Networks, *IEEE Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks and Workshops*, Berlin, Germany, 2008, pp. 443-451.

[8] G. Li and H. Liu, Resource Allocation for OFDMA Relay Networks with Fairness Constraints, *IEEE Journal on Selected Areas in Communications*, Vol. 24, No. 11, pp. 2061-2069, November, 2010.

[9] M. K. Awad and X. Shen, OFDMA Based Two-Hop Cooperative Relay Networks Resource Allocation, *IEEE International Conference on Communications*, Beijing, China, 2008, pp. 4414-4418.

[10] W.-H. Kuo and J.-F. Lee, Multicast Recipient Maximization in IEEE 802.16j WiMAX Relay Networks, *IEEE Transactions on Vehicular Technology*, Vol. 59, No. 1, pp. 335-343, September, 2010.

[11] W.-H. Kuo and J.-F. Lee, Multicast Routing Scheme for Recipient Maximization in Wireless Relay Networks, *IEEE Transactions on Vehicular Technology*, Vol. 59, No. 8, pp. 4002-4011, July, 2010.

[12] J. E. Wieselthier, G. D. Nguyen and A. Ephremides, On the Construction of Energy-Efficient Broadcast and Multicast Trees in Wireless Networks, *IEEE International Conference on Computer Communications*, Tel Aviv, Israel, 2000, pp. 585-594.

[13] P.-J. Wan, G. Calinescu, X.-Y. Li and O. Frieder, Minimum-Energy Broadcast Routing in Static Ad Hoc Wireless Networks, *IEEE International Conference on Computer Communications*, Anchorage, AK, 2001, pp. 1162-1171.

[14] A. K. Das, R. J. Marks, M. El-Sharkawi, P. Arabshahi and A. Gray, Minimum Power Broadcast Trees for Wireless Networks: Integer Programming Formulations, *IEEE International Conference on Computer Communications*, San Francisco, CA, 2003, pp. 1001-1010.

[15] F. Li and I. Nikolaidis, On Minimum-Energy Broadcasting in All-Wireless Networks, *26th Annual IEEE Conference on Local Computer Networks*, Tampa, FL, 2001, pp. 193-202.

[16] O. Egecioglu and T. F. Gonzalez, Minimum-Energy Broadcast in Simple Graphs with Limited Node Power, *IASTED International Conference on Parallel and Distributed Computing and Systems*, Richardson, TX, 2001, pp. 334-338.

[17] M. Cagalj, J. P. Hubaux and C. Enz, Minimum-Energy Broadcast in All-Wireless Networks: NP-Completeness and Distribution Issues, *ACM Annual International Conference on Mobile Computing and Networking*, Atlanta, GA, 2002, pp. 172-182.

[18] A. K. Das, R. J. Marks, M. El-Sharkawi, P. Arabshahi and A. Gray, r-shrink: A Heuristic for Improving Minimum Power Broadcast Tree in Wireless Networks, *IEEE Global Telecommunications Conference*, San Francisco, CA, 2003, pp. 523-527.

[19] S. Guo and O. Yang, A Dynamic Multicast Tree Reconstruction Algorithm for Minimum-Energy Multicasting in Wireless Ad Hoc Networks, *IEEE International Conference on Performance, Computing, and Communications*, Phoenix, AZ, 2004, pp. 637-642.

[20] J. Cartigny, D. Simplot and I. Stojmenovic, Localized Minimum-Energy Broadcasting in Ad-hoc Networks, *IEEE International Conference on Computer Communications*, San Francisco, CA, 2003, pp. 2210-2217.

[21] R. J. Marks, A. K. Das, M. El-Sharkawi, P. Arabshahi and A. Gray, Minimum Power Broadcast Trees for Wireless Networks: Optimizing Using the Viability Lemma, *IEEE International Symposium on Circuits and Systems*, Phoenix-Scottsdale, AZ, 2002. pp. I-273-I-276.

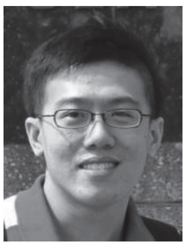
[22] W. Liang, Constructing Minimum-Energy Broadcast Trees in Wireless Ad Hoc Networks, *ACM International Symposium on Mobile Ad Hoc Networking and Computing*, Lausanne, Switzerland, 2002, pp. 112-122.

[23] M. Charikar, C. Chekuri, T.-Y. Cheung, Z. Dai, A. Goel, S. Guha and M. Li, Approximation Algorithms for Directed Steiner Problems, *Journal of Algorithms*, Vol. 33, No. 1, pp. 73-91, October, 1999.

[24] M.-I. Hsieh, E. H. Wu and M. F. Tsai, FasterDSP: A Faster Approximation Algorithm for Directed Steiner Tree Problem, *Journal of Information Science and Engineering*, Vol. 22, No. 6, pp. 1409-1435, November, 2006.

[25] T. S. Rappaport, *Wireless Communications: Principles and Practices*, Prentice-Hall, 1996.

Biographies



Cheng-Hsien Lin received the PhD degrees in computer science and information engineering from National Chung Cheng University, Chiayi, Taiwan in 2013. He is an Android Apps engineer in the RealTime Corporation. His research interests include the resource management and optimization in the 4G networks.

Ren-Hung Hwang received his PhD from University of Massachusetts. He is distinguished professor of the department of Computer Science and Information Engineering and the Dean of the Engineering College at National Chung Cheng University, Taiwan. His research interests include wireless networks, Internet of Things, and cloud computing.

Jeng-Farn Lee received his PhD degree from National Taiwan University, Taiwan, in 2007. He was a postdoctoral fellow with the Institute of Information Science, Academia Sinica, until July 2007. In August 2007, he joined the Department of Computer Science and Information Engineering, National Chung Cheng University as an assistant professor.

Chuan-Kang Ting received the Dr. rer. nat. degree from the University of Paderborn, Germany, in 2005. He is currently an associate professor at Department of Computer Science and Information Engineering, National Chung Cheng University, Taiwan. His research interests include evolutionary computation and its applications in music, networks, and bioinformatics.