Network selection and resource allocation for multicast in HetNets

Ma, Maode; Ma, Dong

2014

Ma, D., & Ma, M. (2014). Network selection and resource allocation for multicast in HetNets. *Journal of network and computer applications*, 43, 17-26

<https://hdl.handle.net/10356/105165>

<https://doi.org/10.1016/j.jnca.2014.03.015>

© 2014 Elsevier Ltd. This is the author created version of a work that has been peer reviewed and accepted for publication by *Journal of Network and Computer Applications*, Elsevier Ltd. It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at:
[DOI:<http://dx.doi.org/10.1016/j.jnca.2014.03.015>].

Downloaded on 26 Apr 2023 21:59:35 SGT

Network Selection and Resource Allocation for Multicast in HetNets

Dong Ma^{1,2,*} and Maode Ma²

¹*School of Urban Rail Transportation, Soochow University, Suzhou 215131, China*

²*School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798*

ABSTRACT

In a heterogeneous wireless system, the member stations of a multicast stream may be distributed in many networks within the system. Meanwhile, each member station may have multiple network interfaces available simultaneously. Therefore, the selection of networks for serving the stream and the allocation of radio resources in the selected networks play significant roles in improving the utilization of system resource, especially for the stream with a large group size. In this paper, a novel joint network selection and resource allocation scheme is proposed aiming to minimize the overall resources consumed by the multicast stream in the heterogeneous wireless system. The fundamental idea is to gather the member stations locating in the same network region and to transmit the stream with possibly high modulation and coding rates. We first formulate the scheme as a binary integer programming problem, and then derive a computationally efficient Lagrangian-based heuristic algorithm to solve the problem, where the architecture of interworking system is taken into account. Simulation results show that the proposed scheme will cost the minimum system bandwidth compared to other reference schemes.

Keywords: HetNets, multicast, network selection, resource allocation.

1. Introduction

In recent years, an explosive growth in mobile data demands has been driven by media-hungry devices (e.g., smart phones). Although the data rates provided by advanced wireless access technologies have also been greatly improved, the link efficiencies are approaching their fundamental limits. It's believed that the only scalable way to meet the current "capacity crunch" is the interworking of heterogeneous networks (HetNets) (Andrews, 2013). Typically, these networks may have different access technologies including Long Term Evolution (LTE), Wireless Local Area Network (WLAN) and Worldwide Interoperability for Microwave Access (WiMAX). It is also possible for them to employ the same technology but with different transmission power (i.e., coverage), network density and data rate. The Heterogeneous Cellular Network (HCN) is a fast growing approach of the second type. In this paper, we consider the efficient multicast in HetNets since it has attracted much research attention due to the highly increased popularity of mobile video services like IPTV.

Within a HetNet, the member stations of a multicast stream may be distributed in multiple networks, meanwhile each member station may be covered by more than one network as shown in Fig. 1. The issue of how to select suitable networks to serve the multicast stream then arises naturally. If the member stations locating at the same region connect to different networks while only one can serve all of them, there is a waste of system radio resources. The problem becomes even worse when more stations join the group. It is thus expected that the member stations of a stream could cluster together such that fewer networks are involved in transmission.

However, this could not be the unique factor considered in the network selection for the multicast in HetNets. It is known that adaptive modulation and coding has been widely supported by current wireless access technologies as a resource allocation technique. Usually, a station close to the Base Station (BS) presenting good channel conditions can deploy a high modulation and coding rates to reduce the resource consumption. If the station moves away from BS later, a more robust Modulation and Coding Scheme (MCS) may be required to maintain the link stability, resulting in a lower data rate and more consumed radio resources. In traditional homogeneous networks, since the member stations of a multicast stream observe different channel conditions, the most robust MCS corresponding to the member station with the worst channel conditions has to be deployed to guarantee a successful transmission to all the member stations. In a HetNet, however, the situation becomes more challenging and interesting because the member station under worse channel conditions may experience better conditions in other available networks. If it connects to another network, fewer radio resources may be consumed from the entire system point of view. Therefore, the channel conditions of the member stations and the resource allocation for the stream should be taken into account by the network selection strategy.

For a clearer explanation, three potential network selection strategies are compared briefly in Fig. 1. The member stations under strategy ① randomly connect to their available networks leading all the three networks to be involved in the transmission (i.e., $A \rightarrow N2(30)$, $B \rightarrow N1(10)$, $C \rightarrow N3(5)$). Since the rule of strategy ② is only to gather the member stations, the tier-1 network N1 is selected to serve all these member stations with its most robust MCS. Nevertheless, with the resource allocation taken into account by strategy ③, A and B connect to N1 with the least robust MCS deployed while C which is further away connect to N3. The mini-

* Corresponding author.

Email-address: mado0001@e.ntu.edu.sg, maode_ma@pmail.ntu.edu.sg

Fig. 1. Comparison of different network selection strategies.

total bandwidth cost is therefore introduced in this case, which reveals that the resource allocation should be tightly integrated into the network selection process.

Based on the above considerations, we propose a joint Network Selection and Resource Allocation (NSRA) scheme for the multicast service in HetNets. The objective is to minimize the overall consumption of the bandwidth cost for the multicast service within the entire system. We first formulate the scheme as a binary integer programming problem, and then derive a Lagrangian-based heuristic algorithm to find solutions. The specific architecture of the interworking system is taken into account to make the algorithm computationally efficient. Moreover, we address some important issues for the implementation of the scheme.

The remainder of the paper is organized as follows. In Section 2, some related works are briefly presented. In Section 3, we first describe the system model and notations, and then formulate the NSRA scheme to an optimization problem. Furthermore, a Lagrangian-based heuristic algorithm for solving the problem is introduced and the implementation issues of the algorithm are also discussed. In Section 4, the performance of our scheme is evaluated by comparing with other schemes. Finally, the paper is concluded and summarized in Section 5.

2. Related works

So far, there have been a number of research works published for the interworking of heterogeneous networks, especially in the recent two years with the rapid development of HCN. These works cover various aspects including network selection (see review provided by Wang and Kuo, 2013), vertical handoff (see review provided by Zekri et al., 2012), and resource allocation (Singh et al., 2013; Ramaboli et al., 2012). The latest work (Andrews, 2013) strongly confirmed the trend of HetNet and provided valuable recommendations on seven key factors in the system design. We have also proposed solutions of proactive vertical handoff between WiMAX and WLAN networks to improve the quality of service for users (Ma and Ma, 2009; Ma and Ma, 2012). In addition, we have introduced a resource management scheme in (Ma and Ma, 2013) to dynamically balance the network traffic load from the network operators' point of view. However, most of existing works for the interworking of heterogeneous networks are designed for the unicast service.

Although the multicast in wireless networks has attracted much more research attention, the previous works have mainly focused on the homogeneous systems. Many approaches have addressed the issue of determining suitable MCSs (i.e., allocating radio resources) for the multicast applications in order to effectively utilize the radio resource in a network (Lin et al., 2013; Kuo et al., 2011; Meddour et al., 2012).

Among the very few works for the multicast in HetNets, the deployment issues and network architectures for multicasting in cooperative relay-based HetNets have been addressed in (Peng et al., 2011; Xie et al., 2011). The handoff procedure has been defined in (Ying et al., 2011; Xu et al., 2014). The solution presented in (Yang and Chen, 2008:7(4)) is supposed to be the first piece of work addressing the network selection for multicast in HetNets, which aims to minimize the bandwidth cost consumed by the stream. The authors have further extended the scheme to be used in the special case of multicasting scalable videos with layered structure in (2008:7(2); 2009). However, by these solutions, the channel conditions of individual member stations have not been differentiated and the same bandwidth cost has been configured to all the member stations in a network. That is, the resource allocation has not been taken into account in the network selection process, which makes the scheme unsuitable for advanced wireless access technologies supporting adaptive modulation and coding. Moreover, the schemes require the member stations to perform network selection operations and iteratively exchange bandwidth cost information with BSs. The signaling overhead over wireless links and the implementation complexity of mobile devices further reduce the feasibility of these schemes. To the best of our knowledge, our solution is the first piece of work jointly considering the network selection and resource allocation for minimizing the bandwidth consumption of multicast streams in HetNets.

3. Design of the NSRA scheme

3.1. System model and notation

Here, we mainly focus on the network selection and resource allocation over the wireless local loop as suggested in (Lin et al., 2013), because the bandwidth of wireless links is much scarcer than that of the wired links and thus becomes a dominant factor in the overall system optimization. We assume that adaptive modulation and coding is supported by all the networks within the system since this is the trend of advanced wireless technologies. Letting M denote the group of member stations requesting the stream, for each member station $m \in M$, there are a set of networks N_m available. The available networks of all the member stations compose the network set N . For the network n ($n \in N$), the MCSs supported by the covered member stations are included in the set $B_n = \{b_1, b_2, \dots, b_p\}$, where $b_1 < b_2 < \dots < b_p$ in terms of robustness. (n, b) denotes the MCS b ($b \in B_n$) in the network n , under which the number of radio resource units (e.g., slots) consumed by the stream is $r_{n,b}$. Obviously, we have $r_{n,b1} < r_{n,b2} < \dots < r_{n,bp}$ because a more robust MCS leads to a lower data rate and hence consumes more bandwidth resources. If the bandwidth cost per unit resource in the network n is c_n , the bandwidth cost of MCS (n, b) can be calculated by $c_{n,b} = c_n \times r_{n,b}$. It is worthy to be noted here that the flexibility of our bandwidth cost-based solution is provided by c_n , which makes our solution very suitable for HetNets. For instance, if the network n tends to be congested, the network operator can increase c_n such that the stream tends to flow into other networks with lower costs. Moreover, if the stream users have a preference on a certain network, this can be realized by decreasing corresponding c_n of that network.

In order to reduce the radio resources consumed by the stream, clearly, the MCSs with lower robustness are preferred for transmission. In other words, the finally selected MCS must be one of the least robust MCS supported by the member stations covered by that network. We thus only include the least robust MCS supported by each member station in the network n into B_n to reduce the problem size. For the member station m covered by the network n , usually only a subset of MCSs can be used limited by channel conditions, which are listed in $B_{n,m}$. If the least robust MCS supported by m is (n, b_{pp}) , normally we have $B_{n,m} = \{b_{pp}, b_{pp} + 1, \dots, b_p\}$.

Within the system, usually some of the member stations locate in the same region, especially when the stream has a large group size. A region covers a set of member stations, which have the same set of available networks and support the same set of MCSs. If one of such member stations can successfully receive the stream, others can also at the same time. Then we include only one of the member stations within the same network region into the member station set M to further reduce the problem size. Any m in M hence represents not only a member station but also logically a specific region. In other words, the size of the member station set $|M|$ will not always increase with the joining of group member stations as it is upper bounded by the number of regions within the system. The notations used in the NSRA scheme are summarized in Table 1.

Table 1
Notations used in the NSRA scheme.

N	the set of networks covering the stream members.
B_n	the set of MCSs supported by the member stations covered by the network n .
c_n	the bandwidth cost per unit resource in the network n .
$c_{n,b}$	the bandwidth cost introduced with the MCS (n, b) .
$r_{n,b}$	the number of resource units consumed with the MCS (n, b) .
$M_{n,b}$	the set of member stations that can be served by the MCS (n, b) .
M	the set of member stations in the system.
N_m	the set of networks available to the member station m .
$B_{n,m}$	the set of MCSs in the network n available to the member station m .

3.2. Mathematical formulation

The function of our NSRA scheme is to select networks for delivering the stream to all its member stations within the system, meanwhile determine the amount of bandwidth resource allocated to the stream in each selected network (i.e., determine the MCS deployed). The objective is to minimize the overall bandwidth cost taken by the stream within the whole heterogeneous overlay system.

Here, we define a binary variable $x_{n,b}$ for the MCS (n, b) , that is

$$x_{n,b} = \begin{cases} 1, & (n,b) \text{ is selected for transmission} \\ 0, & \text{otherwise} \end{cases}.$$

The state "1" for $x_{n,b}$ indicates that the network n is involved in the service meanwhile it needs to allocate proper resource required by the MCS (n, b) for serving the stream. Therefore, once the states of all the elements in the vector $\mathbf{x} = \{x_{n,b} \mid n \in N, b \in B_n\}$ are determined, both the network selection and resource allocation operations will be completed. In addition, " $x_{n,b} = 1$ " also indicates

that $|M_{n,b}|$ member stations can successfully receive the stream with a bandwidth cost of $c_{n,b}$ paid. It is obvious that to minimize the overall system bandwidth cost consumed by the stream, we prefer to select the MCS (n, b) with larger $|M_{n,b}|$ (i.e., gathering member stations) and lower $c_{n,b}$ (lower robustness).

Based on the above considerations, we formulate the NSRA scheme as the following Binary Integer Programming problem (BIP)

$$\text{Minimize } C = \sum_{n \in N} \sum_{b \in B_n} c_{n,b} x_{n,b} \quad (1)$$

$$\text{Subject to } \sum_{n \in N_m} \sum_{b \in B_{n,m}} x_{n,b} \geq 1, \quad \forall m \in M, \quad (2)$$

$$\sum_{b \in B_n} x_{n,b} \leq 1, \quad \forall n \in N, \quad (3)$$

$$x_{n,b} \in (0, 1), \quad \forall n \in N \text{ and } \forall b \in B_n. \quad (4)$$

Without loss of generality, we assume an integer for the bandwidth cost $c_{n,b}$ of each MCS. The constraint (2) ensures that any member station m ($m \in M$) can receive the stream from at least one available network with a supported MCS. The involvement of constraint (3) guarantees that in each network $n \in N$, the stream is served with at most one MCS, i.e., the stream can only be modulated and encoded once in each selected network. For the final optimal solution of the minimization problem, this constraint may be unnecessary. It is because if two MCSs (n, b_1) and (n, b_2) ($b_1 < b_2$) are selected by a solution simultaneously, (n, b_1) will be redundant as all the member stations served by (n, b_1) can be served by the more robust MCS (n, b_2) . Such cases can be removed during the minimization process. However, this constraint well reflects the overlay structure of HetNets. It can then be used to expedite the minimization process and make the problem computationally cheaper as shown in the next subsection.

From the objective function (1) and the constraint (2), we can see that the BIP problem is a special case of set covering problem. In a set covering problem, each set covers some elements with an assigned cost and the problem calls for the minimum-cost group of sets with each element covered by at least one set. Each MCS in our scheme can be regarded as a set in the set covering problem. Specially, our BIP problem presents an overlay structure where the network with a larger coverage overlay the network with a smaller coverage and the MCSs with more robustness overlay those with lower robustness from the viewpoint of the member stations covered by them. It is known that the set covering problem is NP-hard in strong sense (Garey and Johnson, 1979). An effective way to solve the problem is deploying heuristic methods to find the optimal or near-optimal solutions. With the development of mathematical programming, a number of heuristic algorithms have been derived for solving the real-world set covering instances utilizing the structural properties of these instances to improve the performance of algorithms (Caprara et al., 2000). In the next subsection, we derive a Lagrangian-based heuristic algorithm to solve the BIP problem, which takes into account the overlay structure of our NSRA scheme. Lagrangian approach is especially suitable for the integer programming problem with a relatively easy objective function complicated by a set of side constraints (Fisher, 1985), just like our BIP problem. It usually can find good solutions at reasonable computational cost, and thereby has been increasingly used in large-scale mathematical programming including the set covering problem (Beasley, 1993; Ceria et al., 1998).

3.3. Algorithm for solving the problem

In this subsection, we first relax the BIP problem with a Lagrangian approach, by which the complicated constraints can be replaced with a penalty term in the objective function. The relaxed problem then becomes easier to be solved, which can provide a lower bound to the optimum of the original problem. Based on this Lagrangian relaxation, we then derive a heuristic algorithm to find the optimal or near-optimal solution of the original BIP problem.

3.3.1. Lagrangian relaxation

By the algorithm, the constraint (2) is chosen to be relaxed. We define a Lagrangian multiplier vector $\mathbf{u} = \{u_m \geq 0 \mid m \in M\}$ for all the member stations in M and attach it to the constraint (2). Penalty functions are then constructed and brought into the objective function in (1). Subsequently, we can obtain the following Lagrangian Relaxation Problem (LRP)

$$\begin{aligned} \text{Minimize } C(\mathbf{u}) &= \sum_{n \in N} \sum_{b \in B_n} c_{n,b} x_{n,b} + \sum_{m \in M} u_m (1 - \sum_{n \in N_m} \sum_{b \in B_{n,m}} x_{n,b}) \\ &= \sum_{n \in N} \sum_{b \in B_n} (c_{n,b} - \sum_{m: (n,b) \in B_{n,m}} u_m) x_{n,b} + \sum_{m \in M} u_m \\ &= \sum_{n \in N} \sum_{b \in B_n} \varepsilon_{n,b} x_{n,b} + \sum_{m \in M} u_m \end{aligned} \quad (5)$$

$$\text{Subject to } \sum_{b \in B_n} x_{n,b} \leq 1, \quad \forall n \in N, \quad (3)$$

$$x_{n,b} \in (0, 1), \quad \forall n \in N \text{ and } \forall b \in B_n, \quad (4)$$

where $\varepsilon_{n,b} = c_{n,b} - \sum_{m:(n,b) \in B_{n,m}} u_m$ is the relative cost of the MCS (n, b) . Here, u_m can be seen as the utility obtained when the member station m receives the stream, and $\sum_{m:(n,b) \in B_{n,m}} u_m$ is the total utility obtained from all the member stations in M that can be served by the MCS (n, b) . That is, if the MCS (n, b) is selected for transmission, the system obtains a utility of $\sum_{m:(n,b) \in B_{n,m}} u_m$ with a bandwidth cost of $c_{n,b}$ paid. Obviously, the MCS with smaller relative cost $\varepsilon_{n,b}$ is preferred.

Since the Lagrangian multiplier u_m of each member station $m \in M$ is nonnegative and $(1 - \sum_{n \in N} \sum_{b \in B_{n,m}} x_{n,b}) \leq 0$ is required by the constraint (2), the optimum of the LRP problem is less than that of the original BIP problem. And thus, the optimal solution of LRP problem under a given multiplier vector $\mathbf{u}$ provides a lower bound on the optimum of BIP. Because the complicated constraint (2) has been relaxed, it is easier to solve LRP problem and then find the lower bound of the optimum of BIP.

3.3.2. Lagrangian heuristic algorithm

Based on the LRP described above, we derive a heuristic algorithm to find the optimal or near-optimal solution of the original BIP problem. The fundamental idea is to iteratively update the Lagrangian multiplier vector $\mathbf{u}$ to generate a series of the optimums of LRP problem and feasible solutions of the original BIP problem. As discussed above, an optimal solution of LRP can provide a lower bound to the optimum of BIP. Moreover, any feasible solution of BIP must match an objective value greater than its final optimum and then constitutes an upper bound. The algorithm is to iteratively increase the lower bound and decrease the upper bound in order to make them as close to the optimum of BIP as possible.

The details of the algorithm are illustrated in Fig. 2. $\underline{C}$ and $\bar{C}$ record the best known lower and upper bounds obtained by the algorithm, respectively. $\bar{x}$ is the feasible solution of BIP corresponding to $\bar{C}$. Initially, the multiplier of each member station u_m is set to be 1. In step 1, the LRP problem in (5) is solved under the multiplier vector $\mathbf{u}$ of current iteration. The relative cost $\varepsilon_{n,b}$ of each MCS (n, b) ($n \in N, b \in B_n$) is calculated first. As above mentioned, the MCS with smaller relative cost is preferred by the system for the more utility obtained by it and lower bandwidth cost paid. Meanwhile, it is required by the constraint (3) that at most one MCS can be selected for transmission in each network. Therefore, for each network n , we find out its MCS, b^* , with the minimum relative cost because it presents the highest probability of being selected. If its relative cost $\varepsilon_{n,b^*} \leq 0$, we set $x_{n,b^*}(\mathbf{u}) \leftarrow 1$; otherwise, $x_{n,b^*}(\mathbf{u}) \leftarrow 0$. Once the binary variables of all the MCSs are determined, the optimal solution of the LRP problem under $\mathbf{u}$ is obtained as $\mathbf{x}(\mathbf{u}) = \{x_{n,b}(\mathbf{u}) \mid n \in N, b \in B_n\}$. At the same time, the optimum of LRP problem $C(\mathbf{u})$ can also be calculated, which provides a new lower bound to the optimum of BIP. The best known lower bound $\underline{C}$ will be updated if $C(\mathbf{u})$ is greater.

By step 1, the optimal solution of LRP problem $\mathbf{x}(\mathbf{u})$ is obtained; however, this solution may not be feasible to the original BIP problem, which is caused by the fact that the constraint (2) is relaxed in LRP and thus it cannot be guaranteed that all the member stations can be served by the solution $\mathbf{x}(\mathbf{u})$. Even though, $\mathbf{x}(\mathbf{u})$ can also give us useful information on the structure of the optimal solution of BIP. We then construct a feasible solution of BIP in step 2 based on $\mathbf{x}(\mathbf{u})$, which is referred to as $\bar{x}(\mathbf{u})$. Firstly, we check each member station in M . If the member station m cannot be served by $\mathbf{x}(\mathbf{u})$ (i.e., $\sum_{n \in N} \sum_{b \in B_{n,m}} x_{n,b}(\mathbf{u}) = 0$), we seek its available network and MCS with the minimum relative cost (n'', b'') , and set $\bar{x}_{n'',b''}(\mathbf{u}) \leftarrow 1$ to add this MCS to the solution $\bar{x}(\mathbf{u})$. If another MCS b (definitely $b < b''$) has already been selected in the network n'' , we set $\bar{x}_{n'',b}(\mathbf{u}) \leftarrow 0$ because it is overlaid by b'' and thus unnecessary. Once all the member stations can be served by $\bar{x}(\mathbf{u})$, it becomes a feasible solution of BIP. Then the corresponding objective value can be computed by (1), which provides a new upper bound of the optimum and may update $\bar{C}$. So far, both the upper bound $\bar{C}$ and the lower bound $\underline{C}$ of current iteration have been obtained. If these two bounds are identical, they are just the optimum of BIP problem, and then the algorithm will be stopped here.

Otherwise, if the optimal solution is not found at this iteration, we need to update the multiplier vector $\mathbf{u}$ in order to improve the solution of the next iteration to make both the upper and lower bounds as close to the optimum of BIP as possible. This is done in step 3 by a subgradient approach, where the subgradient associated with the current utility of each member station is defined as

$$s_m(\mathbf{u}) = \frac{\partial C(\mathbf{u})}{\partial u_m} = (1 - \sum_{n \in N} \sum_{b \in B_{n,m}} x_{n,b}(\mathbf{u})), \quad \forall m \in M. \quad (6)$$

If $s_m(\mathbf{u}) < 0$, more than one available network of the member station m is selected for service, then we reduce its utility u_m for the next iteration. While if $s_m(\mathbf{u}) = 0$, just one network serves m which is the case desired by the algorithm; and thus, u_m will be kept unchanged in this case. If $s_m(\mathbf{u}) > 0$ (i.e., $s_m = 1$), none of the available networks of m is selected for service under current utility vector $\mathbf{u}$. The utility of such member station needs to be increased in order to increase the probability for its available networks to be selected at the next iteration. The nonnegative parameter λ controls the step size along the subgradient direction, which is assigned a starting value and halved when $\bar{C}$ has not been decreased for a specified number of iterations.

At the end of each iteration, the necessity of each MCS will be checked in order to continually reduce the problem size during the execution of the algorithm. If a MCS (n, b) is not selected by the current solution of LRP (i.e., $x_{n,b}(\mathbf{u}) = 0$) but with $C(\mathbf{u}) + \varepsilon_{n,b} > \bar{C}$, it will be removed from the MCS set of the network n because a MCS leading to the lower bound greater than the upper bound cannot be in the optimal solution (Beasley, 1993). The algorithm stops when λ is less than a threshold λ_{thr} , or the number of iterations t exceeds a threshold T . In such cases, a near-optimal solution ($\bar{x}$ and $\bar{C}$) is achieved by the algorithm.

3.3.3. Implementation details and complexity analysis

3.3.3.1 Implementation details

In a heterogeneous wireless system, the radio resources of networks can be regarded as being in a resource pool. To efficiently exploit the pooled radio resources, the joint radio resource management or Common Radio Resource Management (CRRM) is strongly suggested by current research works and standards (Skehill et al., 2007; 3GPP TR 25.881, 2002; Hasib and Fapojuwo,

Fig. 2. Lagrangian heuristic algorithm.

2008). With CRRM, a central controller is usually introduced such as the CRRM server (CRMS) (3GPP TR 25.881, 2002) defined by the Third Generation Partnership Project (3GPP). The central controller can collect measurements and exchange management information with local RRM entities (e.g., BSs), and apply a common algorithm from a complete network view. Local RRM entities report the events (including the joining, leaving, and change of supported MCSs) happening at its covered member stations to the central controller, which will then decide the initiation of the algorithm. As mentioned in Section 3.1, any m in M indeed represents a network region which includes a set of member stations with the same set of available networks and supported MCSs. Therefore, if a member station leaves while other member stations locating in the same region still exist, the set M is not changed and thus our algorithm will not be invoked. Only when an event changes the network regions, our algorithm is invoked by the central controller. If a solution is found after executing the algorithm, the central controller will inform the selected networks to serve the stream with the MCS designated by the solution.

On the other hand, if there is no central controller in the system, our algorithm can still work in a distributed manner enabled by the Lagrangian-based approach. The networks exchange the event information of their covered member stations. Upon the variation of the region set, the algorithm will be initiated. At the beginning of an iteration, each network n ($n \in N$) calculates the relative cost $\varepsilon_{n,b}$ ($b \in B_n$), the binary variable $x_{n,b}$, and the network relative cost $\varepsilon_n = \sum_{b \in B_n} \varepsilon_{n,b} x_{n,b}(\mathbf{u})$ in step 1. If a network n finds a member station m covered by it with $\sum_{b \in B_{n,m}} x_{n,b}(\mathbf{u}) = 0$, the (n, b'') with the minimum $\varepsilon_{n,b}$ ($b \in B_{n,m}$) will be chosen first in step 2. Further, if this member station was located in an overlapped area, the network n will transmit $\varepsilon_{n,b''}$ to its overlapped networks to find (n'', b'') . This way, $\bar{x}(\mathbf{u})$ can be determined by each network and then $c_n = \sum_{b \in B_n} c_{n,b} \bar{x}_{n,b}(\mathbf{u})$ can be calculated. Furthermore, $x^{m,n} = \sum_{b \in B_{n,m}} x_{n,b}(\mathbf{u})$ will be calculated by the network n for its covered member station m in step 3. By now, each network n can compress ε_n, c_n , and $x^{m,n}$ into one message and broadcast it to other networks in N . Based on these parameters, each network can determine the lower bound $\underline{C}$, the upper bound $\bar{C}$, $s_m(\mathbf{u}) = \sum_{n \in N_m} x^{m,n}$ and further updates the utility u_m for the next iteration.

3.3.3.2 Complexity analysis

By our NSRA scheme, an event occurs at the moment of a member station's joining, leaving, or change of its supported MCSs (e.g., for movement). As discussed above, however, the proposed algorithm is invoked only when the event changes the region set M , i.e., a new region appears or an existing region disappears. Especially for a stream with a large group size, the joining, leaving or movement of many member stations may not trigger the algorithm since their network regions are kept by other member stations.

This is a nice feature because the computational complexity of our algorithm is upper-bounded by the number of network regions and will not always rise with the number of member stations.

Besides the computational consideration, the signaling complexity is an inherent cost of the coordination between networks in HetNets although it is undesirable. In general, the networks with relatively large coverages (e.g., macro and pico BSs) can exchange information very fast via high speed wired dedicated backhauls directly or via a BS controller (Son et al., 2011). The signaling overhead is thereby a nontrivial issue mainly for small-coverage networks (e.g., femto BSs). However, it should be noted that such networks are usually deployed at home or office for residential use; and thus the events reported by these networks mainly include the joining/leaving of member stations.

Now let us investigate the signaling overhead of our scheme mathematically. With the centrally-controlled approach of our scheme, the signaling overhead is introduced by the local RRM entities reporting events to the central controller. We assume that the event occurs on a joining/leaving of a member station and is reported by only one network. Almeroth and Ammar (1997) found that the exponential distribution works well for the group joining inter-arrival time and the membership duration of short sessions lasting less than a day or two. The overhead arrival is thus modeled as Poisson process. We then borrow the conclusions of (Xia et al., 2012) where the signaling overhead in HetNets is evaluated by an overhead quality contour. For the Poisson overhead arrival, the overhead quality contour is established as (7) and (8) for the backhaul and wireless signaling cases, respectively.

$$Q = \left\{ (T, A, d, p_e) : p_e = 1 - \left(\prod_{i=1}^X \frac{\mu_i}{\mu'_i} \right) F(d, A, \{\mu'_i\}_{i=1}^X) \right\} \quad (7)$$

$$Q = \left\{ (T, A, d, p_e) : p_e = e^{-\eta d} q_{k,r} \{\beta(d)\} + \int_0^d q_{k,r} \{\beta(x)\} \eta e^{-\eta x} dx \right\} \quad (8)$$

These equations determine the feasible set of overhead packet parameters (T, A, d, p_e) , where T is the overhead inter-arrival time, A is the packet size, d is the required deadline and p_e is the outage probability. Other parameters including X , which is the number of backhaul servers between BSs and μ_i which is the service rate for the overhead packet at the backhaul server i , reflect the backhaul configurations. Taking the backhaul signaling case as an example, the acceptable overhead inter-arrival time T could be found by the equation (7) under specific backhaul configurations and pre-defined overhead requirements (i.e., given d and p_e). If current overhead exceeds this threshold, the BS could delay invoking the proposed algorithm.

By the distributed approach of our scheme, the signaling overhead consists of two parts: the signaling of reporting an event to other networks and the signaling of exchanging parameters between networks when the algorithm is invoked. The signaling of the first part is similar to the central-control case, and the corresponding overhead quality contour can also be established based on (7) and (8). If the algorithm is invoked, the message containing parameters (e.g. ε_n) will be exchanged between the networks within each iteration. Thus, we could admit that the inter-arrival time of this overhead follows deterministic distribution. Then the corresponding overhead quality contour can be established based on the equations (23) and (34) in (Xia et al., 2012) for backhaul and wireless signaling cases, respectively.

To provide the reader an intuitive view on the signaling complexity, we compare our scheme with the interference management (IM) schemes (Son et al., 2011) which are also essential for the implementation of HetNets. By IM schemes, the networks usually need to exchange user information slot-by-slot (in milliseconds). By our scheme, however, the signaling is mainly triggered by the joining/leaving of member stations (in minutes or even hours) and can thus be almost neglected. In real applications, the messages reporting events of our scheme could be compressed into IM messages.

4. Performance evaluation

To demonstrate the effectiveness of the proposed scheme, we compared the proposed NSRA scheme as well as the Lagrangian heuristic algorithm with other reference schemes. In this section, the system configurations are described first, and then the simulation results for the comparison among different schemes and algorithms are presented and discussed.

4.1. System configurations

Three types of wireless networks named as N1, N2 and N3 networks have been considered in the simulation as illustrated in Fig.1. The key parameters are listed in Table 2. We set the parameters of N1 based on WiMAX networks where 7 MCSs are available (Kuo et al., 2011). Taking the most robust MCS (BPSK 1/2) and the least robust MCS (QAM64 3/4) as examples, they correspond to throughputs of 3.8768 Mbps and 34.904 Mbps, respectively. We thereby roughly assume the numbers of resource units

Table 2
Key parameters used for simulation.

	N1	N2	N3
Network radius (m)	3000	2000	100
Number of MCSs	7	5	3
Resource units needed with each MCS ($r_{n,b}$)	{1, 2, 3, 4, 5, 6, 10}	{1, 2, 3, 4, 6}	{1, 2, 5}
Velocity of stations (m/min)	100	100	15

consumed by the stream under these two MCSs (i.e., $r_{n,b}$) are 10 and 1. Similarly, 5 MCSs are considered in the N2 network according to the cellular standard - High Speed Downlink Packet Access (HSDPA) as in (Dottling et al., 2002). For the N3 network with small-coverage, we assume the three data rates commonly deployed by IEEE 802.11b WLANs (i.e., 2 Mbps, 5.5 Mbps and 11 Mbps) are used for the multicast stream. Corresponding $r_{n,b}$ are roughly set to be 5, 2 and 1, respectively.

In the simulation, each station moves along a random route and randomly decides whether to move and the direction of movement at the beginning of each minute. Within N3 networks that are usually deployed at indoor areas such as home and office, the stations move with a relatively low speed. While in N1 and N2 networks at outdoor areas, they move faster. Based on the observations of (Almeroth and Ammar, 1997), the group member arrival is modeled as Poisson process with a rate of 0.1 per minute, and the membership duration is exponentially distributed with a rate of 0.1 per minute. Each simulation runs with a duration of 1000 minutes.

4.2. Evaluation of NSRA scheme

Since our NSRA scheme is the first piece of work considering both the network selection and resource allocation for minimizing the bandwidth consumption of multicast streams in HetNets to the best of our knowledge, for evaluating its performance, we compare it with other three reasonable reference schemes which work based on the ideas of unicasting schemes and possibly deployed in practical multicasting in HetNets. The first reference scheme (i.e., **RC**) is a typical one of the schemes without control in the network selection, by which the member stations Randomly Connect to their available networks. The second reference scheme (i.e., **MBC**) is a typical one of the WLAN-first (Song et al., 2007) or service charge related (Choi and Choi, 2007) unicasting schemes, where the stations always connect to the available network with the Minimum Bandwidth Cost in order to reduce monetary cost for users. The third reference scheme (i.e., **MOBC**) works based on a general idea of load balancing schemes (e.g. (Ma and Ma, 2013)), which accepts an incoming station into the network leading to the Minimum increase on the Overall system Bandwidth Cost. By all these approaches, the stream is transmitted in the selected networks with the least robust MCS available to their served member stations to minimize the bandwidth cost consumed. In the simulation, our scheme is executed in a centrally controlled manner.

To thoroughly investigate the performance of different schemes, we compare the schemes in multiple scenarios by changing various factors including:

- System size and structure

We use one N1 network with 30 N3 networks to compose a two-tier overlay system; and one N1 network with four N2 networks and 40 N3 networks to compose a three-tier overlay system. The N3 networks are evenly distributed in each MCS region of the N1 network. There are 100 stations in the two-tier overlay system and 200 stations in the three-tier overlay system.

- Mobility of stream

For some multicast services, such as IPTV and video telephony, users usually stay in indoor areas (e.g., home or office). We compare the schemes in such resident scenario and general mobile scenario, separately. In a resident scenario, the stations are uniformly distributed in N3 networks and move only within their covering N3 networks. In a general mobile scenario, the stations are randomly distributed within the N1 network at the beginning and then move in random routes.

- Bandwidth cost

We adjust the bandwidth cost $c_{n,b}$ of MCS (n, b) by changing the bandwidth cost per unit resource c_n based on the equation $c_{n,b} = c_n \times r_{n,b}$. The networks at the same tier are configured with the same bandwidth cost for one unit resource in simulation; and then, c_n can also be written as c_{N1} , c_{N2} and c_{N3} .

Fig. 3 to Fig. 6 compare the schemes in eight scenarios. From these simulation results, we can see that RC always introduces the highest system bandwidth cost. Due to the loss of cooperation, the member stations within the same network region cannot be gathered by RC; consequently, the overlay networks covering the same network region are selected by different member stations at the same time. Especially when a considerable number of stations request the stream, almost all the networks within the system are involved in transmission with the most robust MCSs deployed. Then the system bandwidth is seriously wasted as a result.

Compared to RC, the schemes MBC and MOBC are more intelligent but they seem working relatively well in limited cases. MBC achieves similar performance with the proposed NSRA scheme when there are few member stations and there is a relatively large difference between the bandwidth costs of different tier networks (e.g., comparing (b) to (a) in Fig. 3), because the member stations tend to cluster to the network with a lower bandwidth cost in such cases. Also, MBC works better in mobile scenarios where the mobile stations can obtain more opportunities to switch connections between networks to reduce the cost. However, by MBC, the network selection is performed by individual member stations without an entire system view. There exists the case that multiple member stations connect to several N3 networks separately due to their lower bandwidth costs, but the total bandwidth cost consumed in these N3 networks is higher than the cost of one N2 or N1 network that can serve all these member stations. With the increase of the number of member stations, this problem becomes more serious resulting in the performance of MBC even approaching RC.

By MOBC, the radio resources within the system are centrally controlled and the possibility of serving the stream over the same region via multiple networks can be decreased. However, the network selection for a newly joined member station is influenced by the previous state of networks. For instance, the N1 network is selected at some time for the member station $M1$ with a bandwidth cost 10 consumed. Later, the member station $M2$ joins the group and also connects to N1 because no additional bandwidth cost is introduced by doing so. If $M1$ leaves after a while, N1 is still maintained by $M2$ although it may only consume a bandwidth cost of 1 in its available N3 network. N1 may be continually kept by subsequent member stations, which is the reason that the system

Fig. 3. Two-tier model with resident application: (a) $c_{N1} = c_{N3} = 1$; (b) $c_{N1} = 5, c_{N3} = 1$.

Fig. 4. Two-tier model with general mobile application: (a) $c_{N1} = 10, c_{N3} = 1$; (b) $c_{N1} = 20, c_{N3} = 1$.

bandwidth cost of MOBC usually tends to be remained at a level with the joining of member stations (e.g., when the member station number is over 60 in Fig. 6(b)). Another example is that 5 member stations join the group one by one. Each member station is connected to a N3 network with a bandwidth cost of 1 introduced at each time leading to a total bandwidth cost of 5. However, if all these member stations connect to one N2 network, the overall cost is only 2. That is why MOBC works relatively bad when there are few stations in the system. Comparatively, the performance of MOBC in resident scenarios is better than in mobile scenarios, because the scheme can adapt to the joining of member stations but not to their movements.

It is clearly shown in the figures that our NSRA scheme always introduces the minimum bandwidth cost into the system compared to the reference schemes, which is caused by the fact that our scheme can select networks for service from the entire system perspective meanwhile excellently adapt to the change of network status and link conditions. That is, when the joining, leaving or movement of a member station causes the change of network region set, our scheme is invoked to determine the networks currently suitable for the stream which can keep minimizing its consumed bandwidth cost. It is indicated in the figures that the member stations within the system can be intelligently gathered by our scheme. Especially in the case that some networks are configured with higher bandwidth costs for per unit resource (e.g., becoming congested), the minimal bandwidth cost obtained by our scheme means it can distribute the traffic from a hot area to cold areas most effectively. In this context, our scheme works well as a load balancing tool.

Another important observation obtained from the simulation results is that, although the overall system bandwidth cost also progressively increases by our NSRA approach with the joining of member stations, the maximum cost is known which indeed is the bandwidth cost corresponding to the most robust MCS of the tier-1 N1 network. That is, the worst case by our approach is to select the tier-1 network to serve all the member stations; hence, the overall system bandwidth cost can be guaranteed to be below a threshold. In Fig. 3(a), the N1 network has a comparable bandwidth cost with the N3 network, and then it is selected for transmission by our NSRA scheme and by MOBC at most of time. But in the similar case of three-tier system as shown in Fig. 5(a), the overall system bandwidth cost is still 10 by our NSRA approach while is increased to be 20 by MOBC because the N2 network is also selected by MOBC during some periods and thus the performance cannot be guaranteed.

4.3. Evaluation of Lagrangian heuristic algorithm

In this subsection, we evaluate the performance of our LRP algorithm from several aspects. Firstly, we compare the algorithm with a greedy-based algorithm, which is another representative heuristic algorithm for solving the set covering problem. A greedy

Fig. 5. Three-tier model with resident application: (a) $c_{N1} = c_{N2} = c_{N3} = 1$; (b) $c_{N1} = 5, c_{N2} = 2, c_{N3} = 1$.

Fig. 6. Three-tier model with general mobile application: (a) $c_{N1} = 10, c_{N2} = 5, c_{N3} = 1$; (b) $c_{N1} = 20, c_{N2} = 10, c_{N3} = 1$.

algorithm for solving our NSRA scheme is developed in Fig. 7, which iteratively selects the most cost effective MCS until all the member stations can be served. A two-tier overlay network is simulated. Meanwhile, the software CPLEX (CPLEX, 2010) has also been used to solve the NSRA scheme. The results obtained by the three approaches are compared in Fig. 8, which shows that our scheme can achieve equivalent performance with CPLEX and better performance than the greedy algorithm, especially, when there are more member stations in the system.

Furthermore, we compare the LRP algorithm with the Lagrangian algorithm derived in [28] (Yang and Chen, 2008:7(4)), where a network selection scheme is designed for the multicast in HetNets while the resource allocation is not taken into account and all the stations in a network are assumed to be with the same bandwidth cost. For comparison, we configure each network with only one MCS. A three-tier overlay network is simulated and all the Lagrangian related parameters (e.g., λ) used in these two algorithms are identical. In the simulation, we found that if the network bandwidth costs are relatively low, the performances of the two algorithms are similar. Otherwise, more iterations are usually required by the algorithm in [28] to obtain the optimal solution, because it reduces the upper bound slower from a high value to the optimum. Fig. 9 compares the performance of the two algorithms with c_{N1} set to be 20, c_{N2} uniformly selected from 1 to 10 and c_{N3} uniformly selected from 1 to 3. Each point in the curve is an average from 1000 runs with the algorithms compared directly in each run. It is shown in Fig. 9(b) that much less iterations are required by our algorithm, especially when there are not too many member stations in the system. The possible reason is that the algorithm in [28] depends on the stream members to select networks; hence, decreasing the number of stream member stations will slow down the speed of the convergence of upper bound and lower bound. However, our algorithm is performed by network components without the involvement of member stations. It can find the optimum with much less iterations in such scenarios. As the number of iterations needed by the algorithm in [28] reaches the threshold $T = 2000$ with a much higher probability with only near-optimal solutions found at that case, more bandwidth cost is consumed as shown in Fig. 9(a).

5. Conclusion

1. let $M' = \emptyset, x_{n,b} \leftarrow 0, \forall n \in N, \forall b \in B_n$.
2. find $(n', b') \leftarrow \arg \min_{n \in N, b \in B_n} \{c_{n,b} / |M_{n,b} \setminus M'|\}$; let $x_{n',b'} \leftarrow 1$.
3. let $M' \leftarrow M' \cup M_{n',b'}$;
 if $|M'| \neq |M|$, then return to step 2, endif.
4. for any (n, b) with $x_{n,b} = 1$, (in descending order of $c_{n,b}$)
 if all $m \in M_{n,b}$ with $\sum_{(n',b') \in B_{n,m} \setminus (n,b)} x_{n',b'} = 1$,
 then let $x_{n,b} \leftarrow 0$,
 endif;
 endfor.

Fig. 7. Greedy algorithm for solving the NSRA scheme.

Fig. 8. Comparison with greedy algorithm and CPLEX.

(a)

(b)

Fig. 9. Comparison with the algorithm in [28].

In this paper, we design a joint network selection and resource allocation scheme for the multicast service in heterogeneous wireless networks, which aims to minimize the overall system bandwidth cost consumed by the multicast stream. We formulate the scheme as a binary integer programming problem, and then derive a Lagrangian-based heuristic algorithm to find solutions of the problem. By the scheme, the member stations at the same network region can be effectively gathered and the stream can be served with possibly highest modulation and coding rates to reduce the resource consumption. The simulation results demonstrate that our scheme can introduce the minimum bandwidth cost into the system; meanwhile, the derived algorithm can find good solutions for the scheme with lower computational cost compared with other reference schemes and algorithms.

Newly defined abbreviations

NSRA: the proposed Network Selection and Resource Allocation scheme; **BIP**: the Binary Integer Programming problem; **LRP**: the Lagrangian Relaxation Problem; **N1, N2, N3**: three types of networks listed in the descending order of coverage; **RC**: the first reference scheme where the stations Randomly Connect to their available networks; **MBC**: the second reference scheme where the stations always connect to their available networks with the Minimum Bandwidth Cost; **MOBC**: the third reference scheme where an incoming station is accepted into the network leading to the Minimum increase on the Overall system Bandwidth Cost.

References

- 3GPP TR 25.881, 2002. Improvements of RRM across RNS and RN2/BSS V5.0.0. <<http://www.3gpp.org/ftp/Specs/html-info/25881.htm>>.
- Almeroth KC, Ammar MH. Multicast group behavior in the Internet's Multicast Backbone (MBone). *IEEE Communication Magazine* 1997;35(6):124-129.
- Andrews JG. Seven ways that HetNets are a cellular paradigm shift. *IEEE Communications Magazine* 2013;51(3):136-144.
- Beasley JE. Lagrangean relaxation. In C. R. Reeves (ed.): *Modern Heuristics Techniques for Combinatorial Problems*, Blackwell Scientific Publications; 1993, p. 243-303.
- Caprara A, Toth P, Fischetti M. Algorithms for the set covering problem. *Annals of Operations Research* 2000;98(1-4):353-371.
- Ceria S, Nobili P, Sassano A. A Lagrangian-based heuristic for large-scale set covering problems. *Mathematical Programming* 1998;81(2): 215-228.
- Choi Y, Choi S. Service charge and energy-aware vertical handoff in integrated IEEE 802.16e/802.11 networks. In: 26th IEEE international conference on computer communications INFOCOM; 2007.p. 589-597.
- CPLEX: mathematical programming optimizer; 2010. <<http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/>>.
- Dotling M, Michel J, Raaf B. Hybrid ARQ and adaptive modulation and coding schemes for high speed downlink packet access. In: The 13th IEEE international symposium PIMRC; 2002. p. 1073-7077.
- Fisher ML. An applications oriented guide to Lagrangian relaxation. *Interfaces* 1985;15(2):10-21.
- Hasib A, Fapojuwo AO. Analysis of common radio resource management scheme for end-to-end QoS support in multiservice heterogeneous wireless networks. *IEEE Transactions on Vehicular Technology* 2008;57(4):2426-2439.
- Garey MR, Johnson DS. *Computers and Intractability – A Guide to the Theory of NP-Completeness*. W.H. Freeman and Company; 1979.
- Kuo WH, Liao W, Liu T. Adaptive resource allocation for layer-encoded IPTV multicasting in IEEE 802.16 WiMAX wireless networks. *IEEE Transactions on Multimedia* 2011;13(1):116-124.
- Lin KC, Shen WL, Hsu CC, Chou CF. Quality-differentiated video multicast in multirate wireless networks. *IEEE Transactions on Mobile Computing* 2013;12(1):21-34.
- Ma D, Ma M. A QoS-based vertical handoff scheme for interworking of WLAN and WiMAX. In: *IEEE Global Telecommunications Conference (GLOBECOM'09)*; 2009. p. 2260-2265.
- Ma D, Ma M. A QoS oriented vertical handoff scheme for WiMAX/WLAN overlay networks. *IEEE Transactions on Parallel and Distributed Systems* 2012;23(4):598-606.
- Ma D, Ma M. Proactive load balancing with admission control for heterogeneous overlay networks. *Wireless Communications and Mobile Computing* 2013; 13(18):1671-1680.
- Meddour DE, Abdallah A, Ahmed Toufik, Boutaba R. A cross layer architecture for multicast and unicast video transmission in mobile broadband networks. *Journal of Network and Computer Applications* 2012;35(5):1377-1391.
- Peng M, Liu Y, Wei D, Wang W, Chen HH. Hierarchical cooperative relay based heterogeneous networks. *IEEE Wireless Communications* 2011;18(3):48-56.
- Singh S, Dhillon HS, Andrews JG. Offloading in heterogeneous networks: modeling, analysis, and design insights. *IEEE Transactions on Wireless Communications* 2013;12(5):2487-2497.
- Skehill R, Barry M, Kent W, O'Callaghan M, Gawley N, McGrath S. The common RRM approach to admission control for converged heterogeneous wireless networks. *IEEE Wireless Communications* 2007;14(2):48-56.
- Son K, Lee S, Yi Y, Chong S. REFIM: a practical interference management in heterogeneous wireless access networks. *IEEE Journal on Selected Areas in Communications* 2011; 29(6):1260-1272.
- Song W, Jiang H, Zhuang W. Performance analysis of the WLAN-first scheme in cellular/WLAN interworking. *IEEE Transactions on Wireless Communications* 2007;6(5):1932-1943.
- Wang L, Kuo GS. Mathematical modeling for network selection in heterogeneous wireless networks – a tutorial. *IEEE Communications Surveys & Tutorials* 2013;15(1):271-292.
- Xia P, Jo HS, Andrews JG. Fundamentals of inter-cell overhead signalling in heterogeneous cellular networks. *IEEE Journal of Selected Topics in Signal Processing* 2012;6(3):257-269.

- Xie X, Rong B, Zhang T, Lei W. Improving physical layer multicast by cooperative communications in heterogeneous networks. *IEEE Wireless Communications* 2011;18(3):58-63.
- Xu YH, Chow CO, Tham ML, ISHII H. An enhanced framework for providing multimedia broadcast/multicast service over heterogeneous networks. *Journal of Zhejiang University-SCIENCE C* 2014;15(1):63-80.
- Yang DN, Chen MS. Efficient resource allocation for wireless multicast. *IEEE Transactions on Mobile Computing* 2008;7(4):387-400.
- Yang DN, Chen MS. Bandwidth efficient video multicasting in multiradio multiCellular wireless networks. *IEEE Transactions on Mobile Computing* 2008;7(2):275-288.
- Yang DN, Chen MS. Data broadcast with adaptive network coding in heterogeneous wireless networks. *IEEE Transactions on Mobile Computing* 2009;8(1):109-125.
- Ying K, Yu H, Wang X, Luo H. Multicast/broadcast service over heterogeneous networks. In: *IEEE Global Telecommunications Conference (GLOBECOM)*;2011. p. 1-5.
- Zekri M, Jouaber B, Zeghlache D. A review on mobility management and vertical handover solutions over heterogeneous wireless networks. *Computer Communications* 2012;35(17): 2055-2068.