A Distributed Multi-service Resource Allocation Algorithm in Heterogeneous Wireless Access Medium

Muhammad Ismail, *Student Member, IEEE,* and Weihua Zhuang, *Fellow, IEEE*

Abstract**—In this paper, radio resource allocation in a heterogeneous wireless access medium is studied. Mobile terminals (MTs) are assumed to have multi-homing capabilities. Both constant bit rate and variable bit rate services are considered. A novel algorithm is developed for the resource allocation. Unlike existing solutions in literature, the proposed algorithm is distributed in nature, such that each network base station / access point can perform its own resource allocation to support the MTs according to their service classes. The coordination among different available wireless access networks' base stations is established via the MT multiple radio interfaces in order to provide the required bandwidth to each MT. A priority mechanism is employed, so that each network gives a higher priority on its resources to its own subscribers as compared to other users. Numerical results demonstrate the validity of the proposed algorithm.**

Index Terms**—Heterogeneous wireless networks, resource allocation, multi-homing, network utility, distributed solutions.**

I. INTRODUCTION

CURRENTLY there exist different wireless access networks with different capabilities in terms of bandwidth, latency, coverage area, or cost [2]. These wireless access networks include wireless metropolitan area networks (WMANs), cellular networks, wireless local area networks (WLANs), and so on. The integration of these different networks can help to support user roaming and provide various classes of services with different network resource demands. However, to be able to satisfy the required bandwidth by the mobile terminals (MTs) via different available wireless networks and make efficient utilization of the available resources from these networks, new mechanisms for bandwidth allocation and call admission control are required.

In literature, there exist various works that study the problem of resource allocation in a heterogeneous wireless access medium. These works can be classified in two categories. The first category includes the solutions that utilize the single radio interface of an MT, in which the MT obtains its required bandwidth from a single access network. The second category includes the solutions where multiple radio interfaces of an MT are used simultaneously to satisfy the user's requirement. The resource allocation solutions from this category are known as multi-homing solutions, in which the MT obtains its required bandwidth from all available wireless access networks.

In this paper, the resource allocation problem in a heterogeneous wireless access medium is studied. MTs are assumed to have multi-homing capabilities. Both constant bit rate (CBR) and variable bit rate (VBR) services are considered. A novel algorithm is developed for such a problem. While existing solutions in literature call for a central resource manager to perform the resource allocation, this newly developed algorithm allows each network base station (BS) / access point (AP) to solve its own utility maximization problem and performs its own resource allocation to satisfy the MT requirement, according to its service class. Hence, no need for a central resource manager. The MT plays an active role in the resource allocation operation by performing coordination among different available wireless access networks' BSs/APs to satisfy its required bandwidth. Each network employs a priority mechanism in order to give a higher priority on its resources to its own subscribers than to the other users.

The rest of this paper is organized as follows: Section II reviews the related work. Section III describes the system model. In Section IV, the problem formulation is developed. Section V discusses the proposed algorithm. Section VI presents numerical results and discussions. Finally, conclusions are drawn in Section VI.

II. RELATED WORK

The problem of resource allocation in heterogeneous wireless access networks is studied in [3] - [7]. The existing solutions can be classified in two categories based on whether a single radio interface or multiple radio interfaces of an MT are used simultaneously for the same application. Each category can then be further divided into two groups based on whether the proposed solution can support single class or multiple classes of service.

The resource allocation solutions that belong to the first category are studied in [3] - [5]. In [3], a utility function based resource allocation scheme is introduced for a single service class code division multiple access (CDMA) cellular network and WLAN. In [4], two resource management schemes are proposed for bandwidth allocation and admission control in a heterogeneous wireless access environment with different classes of service. The mechanisms provided in [3] and [4] needs a central resource manager to find the optimum bandwidth allocation. In [5], a distributed resource allocation mechanism is developed to find the optimum bandwidth allocation for a given set of voice users and best effort users in a heterogeneous wireless access environment. While a distributed mechanism is developed in [5], only a single network is considered in obtaining the required bandwidth. The

---

Manuscript received January 17, 2011; revised June 20, 2011; accepted August 30, 2011. The editors coordinating the review of this paper and approving it for publication were X. Shen, Q. Zhang, P.R. Kumar, and Z. Han.

The authors are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada, e-mail:{m6ismail, wzhuang}@bbcr.uwaterloo.ca.

This work is presented in part at IEEE ICC 2011 [1].

resource allocation mechanisms that belong to this category based on a single interface of an MT suffer from the following shortcomings: 1) The incoming call is blocked if no network in the service area can individually satisfy the bandwidth requirement of the MT, as a result these mechanisms do not fully exploit the available resources from different networks; 2) These mechanisms do not improve the system capacity of the individual networks.

The resource allocation solutions that belong to the second category are studied in [6] and [7]. In [6], the concept of utility fairness is applied to allocate bandwidth to different types of traffic. In [7], the problem of bandwidth allocation in a heterogeneous wireless access medium is formulated as a non-cooperative game. The mechanisms of [6] and [7] support MTs with multi-homing capabilities. Thus, each MT can obtain its required bandwidth for a specific application from all available wireless access networks. This has the following advantages [8]: Firstly, with multi-homing capabilities, the available resources from different wireless access networks can be aggregated to support applications with high required bandwidth using multiple threads at the application layer; Secondly, these mechanisms allow for mobility support since at least one of the used interfaces will remain active during the call duration; Finally, the multi-homing concept can reduce the call blocking rate and improve the system capacity.

However, these exiting resource allocation mechanisms for a heterogeneous wireless access environment that support MTs with multi-homing capabilities need a central resource manager to perform the resource allocation and admission control. The need for the central resource manager arises from the fact that the allocated bandwidth from each network BS/AP to a given connection should sum up to the bandwidth required by that connection. Hence, a global view of the BS/AP capacity of every network is needed to coordinate the allocations from different networks to satisfy the required bandwidth for that connection. This global view is provided by the central resource manager. This is not practical in a case that these networks are operated by different service providers. A central resource manager that controls the operation of different networks' BSs/APs in such a case raises some issues related to: 1) the question of which network will be in charge of the operation and maintenance of the central resource manager, considering the fact that such network will control the resources of other networks; 2) changes required in different network structures and operations in order to account for such a central manager; 3) the fact that, if the central resource manager breaks down, the whole multi-homing service fails and this may extend to the operation of the different networks. Hence, in such a networking environment it is desirable to have a distributed solution that enables each network BS/AP to solve its own utility maximization problem and to perform its own resource allocation and admission control, while at the same time cooperates with other available networks BSs/APs to support MTs with multi-homing capabilities.

In this paper, a distributed algorithm for resource allocation in heterogeneous wireless access medium for MTs with multi-homing capabilities is proposed. Each wireless access network BS/AP, in this algorithm, solves its own utility maximization problem to allocate its resources so that the MTs requirements can be satisfied. Two classes of service are considered, namely, CBR and VBR services. When sufficient resources are available from different networks BSs/APs, VBR services are allocated the maximum required bandwidth. On the other hand, when all available networks BSs/APs with overlapped coverage areas reach their capacity limitation, VBR services are degraded towards the minimum required bandwidth using the resources from different overlapped networks BSs/APs. The work of [6] employs a utility fairness concept to ensure that all MTs with VBR service are degraded simultaneously with the same amount of resources within the same wireless access network and according to the same utility change among different wireless access networks. This, however, does not take into consideration the fact that different MTs are the subscribers of different networks and, as a result, they should not be treated equally by each network. It is more practical that each network supports first its own subscribers and ensures that they are satisfied with the maximum possible required bandwidth, while at the same time it supports the subscribers of other networks. To accomplish this, our proposed algorithm employs a priority mechanism so that each network can give a higher priority in allocating its resources to its subscribers as compared to the other users.

III. SYSTEM MODEL

Consider a geographical region where a set $\mathcal{N}$ of wireless access networks with different access technologies is available, $\mathcal{N} = \{1, 2, \ldots, N\}$. Each network is operated by a unique service provider. Each network, $n \in \mathcal{N}$, has a set $\mathcal{S}_n$ of BSs/APs in the geographical region, $\mathcal{S}_n = \{1, 2, \ldots, S_n\}$. The BSs/APs of each network have different coverage from those of other networks. Different networks have overlapped coverage in some areas. As a result, the geographical region can be described by a set $\mathcal{K}$ of service areas, $\mathcal{K} = \{1, 2, \ldots, K\}$. Each service area $k \in \mathcal{K}$ is covered by a unique subset of networks BSs/APs as shown in Figure 1. Each BS/AP, $s \in \mathcal{S}_n$, has a transmission capacity of C_n Mbps. There are M MTs in the geographical region, denoted by set $\mathcal{M}$, $\mathcal{M} = \{1, 2, \ldots, M\}$. During a given period, MTs in a service area k moves within this area, but do not make a handoff to another service area. Each MT, $m \in \mathcal{M}$, has its own home network, but can also get service from other available networks using the multi-homing capability. The set of MTs which lie in the coverage area of the sth BS/AP of the nth network is denoted as $\mathcal{M}_{ns} \subseteq \mathcal{M}$. The subset of MTs whose home network is network n is denoted by $\mathcal{M}_{ns1}$, while the subset of MTs whose home network is not network n is denoted by $\mathcal{M}_{ns2}$. That is, $\mathcal{M}_{ns1} \cup \mathcal{M}_{ns2} = \mathcal{M}_{ns}$, and $\mathcal{M}_{ns1} \cap \mathcal{M}_{ns2} = \phi$. An MT $m \in \mathcal{M}_{ns1}$ is referred to as a *subscriber* of network n, while an MT $m \in \mathcal{M}_{ns2}$ is referred to as a *user* of network n. An MT using its multi-homing capability can receive its required bandwidth from all wireless access networks available at its location. The bandwidth allocated from network n to an MT m through BS/AP s is denoted as b_{nms}, where $n \in \mathcal{N}$, $m \in \mathcal{M}_{ns}$ and $s \in \mathcal{S}_n$. Let B_{ns} be a vector of bandwidth allocation from network n through BS/AP s to each MT within

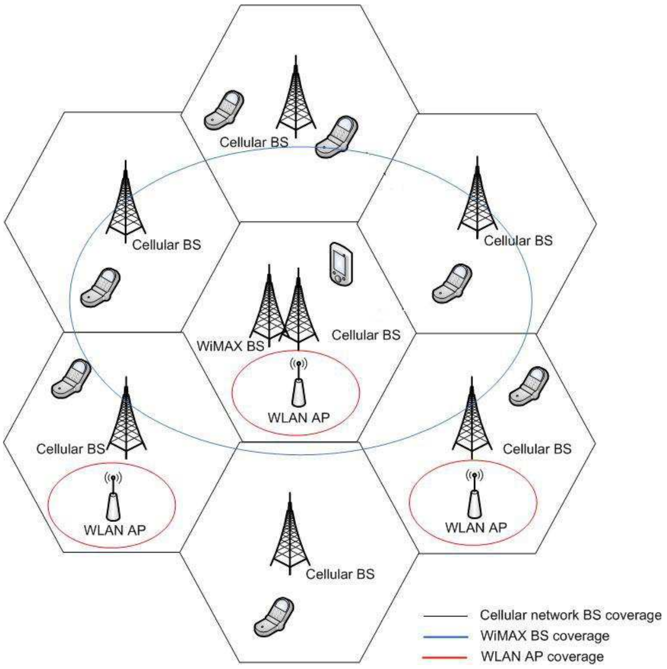


Fig. 1. The network coverage areas

its coverage area, $B_{ns} = (b_{nms} : m \in \mathcal{M}_{ns})$, with $b_{nms} = 0$ if MT m is not in the coverage area of network n BS/AP s.

The networks cooperatively support both CBR and VBR services. A CBR call of MT m requires a constant bandwidth B_m from all available wireless access networks BSs/APs in its service area. On the other hand, a VBR call of MT m requires a bandwidth allocation within a maximum value $B_m^{\max}$ and a minimum value $B_m^{\min}$. When there are sufficient resources in a given service area, the VBR call is allocated its maximum required bandwidth $B_m^{\max}$. When all networks BSs/APs in a given service area reach their capacity limitation, the bandwidth allocation for the VBR call is degraded towards $B_m^{\min}$ in order to support more calls. The set of MTs in the geographical region with CBR service is $\mathcal{M}_{r1}$, while that for MTs with VBR service is $\mathcal{M}_{r2}$, and both are subsets of $\mathcal{M}$.

A connection level only resource allocation is considered in this work. The objective is to find the optimum resource allocation to a set of MTs in a particular service area from each of the available networks BSs/APs in that service area in a given period. This can be performed according to the average connection level statistics in the different service areas [6]. As a result, a static system is studied without arrivals of new calls and departures of existing ones. Also, a call admission control procedure is assumed to be in place [9], so that feasible resource allocation solutions exist.

IV. PROBLEM FORMULATION

In this section, the problem of the multi-service resource allocation in the heterogeneous wireless access medium is formulated. A distributed solution for such a problem is then proposed.

Let $u_{nms}(b_{nms})$ denote a utility function of network n allocating bandwidth b_{nms} to MT m through BS/AP s. The utility function is defined as

$$u_{nms}(b_{nm}) = \ln(1 + \eta_1 b_{nms}) - (1 - p_{nms})\eta_2 b_{nms} \qquad (1)$$

where $p_{nms} \in [0, 1]$ is a priority parameter set by network n on its resources in BS/AP s for MT m, η_1 and η_2 are used for the scalability of b_{nms}. The first term in the right hand side of the utility function represents the attained network utility from the allocated resources b_{nms} [6], which is a concave function of b_{nms} [10]. This term originates from the concept of proportionally fair resource allocation [11]. The second term in the right hand side represents the cost the user pays for the allocated resources. This is a linear function of b_{nms}, the more the allocated resources, the higher the cost. Hence, the utility function of (1) involves a tradeoff between the attained network utility and the cost that the user has to pay on the network resources. Utility function (1) is a concave function of b_{nms} [10]. The priority parameter p_{nms} assigned by network n BS/AP s to MT m is used to establish service differentiation among different users by the network, and is given by

$$p_{nms} = \begin{cases} 1, & \forall m \in \mathcal{M}_{ns1} \\ \beta, & \forall m \in \mathcal{M}_{ns2} \end{cases} \qquad (2)$$

where $\beta \in [0, 1)$. From (2), the utility function of (1) for a network subscriber accounts only on the attained network utility by that subscriber, while a user of the network suffers from a tradeoff between the attained network utility and the cost that the network sets on its own resources. As a result, each network gives a higher priority in allocating its resources to its subscribers as compared to the other users. When all networks BSs/APs in a given service area reach their capacity limitation, resource allocation to the MTs with VBR service is reduced in order to support more calls. The subscribers of each network should be able to enjoy the resources of their own home network as long as they can. As a result, it is desirable to differentiate the allocation performed by the network to its own subscribers and the allocation performed by that network to the other users. This is achieved through the priority parameter p_{nms} to give a higher cost on the network resources for the other users as compared to its own subscribers. Each network sets a priority parameter value $p_{nms} \in [0, 1)$ on its resources for the users in its BS/AP coverage area, while making $p_{nms} = 1$ for its own subscribers. As a result, the subscribers of each network with VBR service enjoy the maximum required bandwidth using their home network resources for the longest possible time. The VBR allocation is degraded by a network to its own subscribers only in order not to violate the minimum required bandwidth of the other users.

The resource allocation objective of each network BS/AP is to maximize the total satisfaction for all the MTs within its coverage area, given by

$$U_{ns}(B_{ns}) = \sum_{m \in \mathcal{M}_{ns}} u_{nms}(b_{nms}), \qquad \forall s \in \mathcal{S}_n, \forall n \in \mathcal{N} \qquad (3)$$

where $U_{ns}(B_{ns})$ is the total utility of network n BS/AP s.

For the whole geographical region, the overall resource allocation objective of all the networks is to find the optimum

TABLE I
SUMMARY OF IMPORTANT SYMBOLS

Symbol	Definition
b_{nms}	Allocated bandwidth from network n to MT m through BS/AP s
B_{ns}	Bandwidth allocation vector from network n BS/AP s to all MTs within its coverage area
B_m	Required bandwidth of MT m with CR service
$B_m^{\min}$	Minimum required bandwidth of MT m with VBR service
$B_m^{\max}$	Maximum required bandwidth of MT m with VBR service
C_n	Transmission capacity of network n BSs/APs
$h(\cdot)$	Dual function
$\mathcal{K}$	Set of service areas in the geographical region
$L(\cdot)$	Lagrangian function
M_{nkr}	The number of subscribers of network n in service area k with service r
$\mathcal{M}$	Set of MTs in the geographical region
$\mathcal{M}_{r1}$	Set of MTs with CBR service in the geographical region
$\mathcal{M}_{r2}$	Set of MTs with VBR service in the geographical region
$\mathcal{M}_{ns}$	Set of MTs in the coverage area of network n BS/AP s
$\mathcal{M}_{ns1}$	Set of subscribers of network n in the coverage area of BS/AP s
$\mathcal{M}_{ns2}$	Set of users of network n in the coverage area of BS/AP s
$\mathcal{N}$	Set of available networks in the geographical region
$\mathcal{S}_n$	Set of BSs/APs of network n in the geographical region
λ_{ns}	Lagrange multiplier corresponding to capacity constraint of network n BS/AP s
ν_m	Lagrange multiplier corresponding to the required bandwidth constraint of MT m with CBR service
$\mu_m^{(1)}$	Lagrange multiplier corresponding to the maximum required bandwidth constraint of MT m with VBR service
$\mu_m^{(2)}$	Lagrange multiplier corresponding to the minimum required bandwidth constraint of MT m with VBR service

allocation b_{nms}, $\forall n \in \mathcal{N}$, $\forall m \in \mathcal{M}$, $\forall s \in \mathcal{S}_n$ that maximizes the total utility in the region, given by

$$U = \sum_{n=1}^{N} \sum_{s=1}^{S_n} U_{ns}(B_{ns}). \tag{4}$$

For each network n BS/AP s in the geographical region, the allocated resources should be such that the total load in its coverage area is within the network BS/AP capacity limitation C_n, that is

$$\sum_{m \in \mathcal{M}_{ns}} b_{nms} \leq C_n, \qquad \forall s \in \mathcal{S}_n, \forall n. \tag{5}$$

For a CBR service, the total allocated resources from different available wireless access networks to a given MT should satisfy the MT application required bandwidth, i.e.,

$$\sum_{n=1}^{N} \sum_{s=1}^{S_n} b_{nms} = B_m, \qquad \forall m \in \mathcal{M}_{r1} \tag{6}$$

while for a VBR service, the total allocated resources from different available wireless access networks to a given MT should be within the application maximum required bandwidth $B_m^{\max}$ and the application minimum required bandwidth $B_m^{\min}$, i.e.,

$$B_m^{\min} \leq \sum_{n=1}^{N} \sum_{s=1}^{S_n} b_{nms} \leq B_m^{\max}, \qquad \forall m \in \mathcal{M}_{r2}. \tag{7}$$

To summarize , the resource allocation problem in the heterogeneous wireless access environment for MTs with multi-homing capabilities, for two classes of service, can be expressed by the following optimization problem

$$\begin{aligned} \max_{B_{ns} \geq 0} \quad & \sum_{n=1}^{N} \sum_{s=1}^{S_n} U_{ns}(B_{ns}) \\ s.t. \quad & (5) - (7). \end{aligned} \tag{8}$$

Following the utility function definitions in (1) and (3), the objective function of problem (8) is concave and the problem has linear constraints. Therefore, problem (8) is a convex optimization problem, which makes a local maximum a global maximum as well [10]. While problem (8) can be solved in a centralized manner with a central resource manager, this is not a practical solution when different networks are operated by different service providers. Hence, a distributed solution of (8) is desirable.

The constraints introduced in (6) and (7) are in fact coupling constraints and, as a result, it is difficult to obtain a distributed solution of (8) at each network. A distributed solution can be developed using full dual decomposition of (8) [12] - [17]. The constraint defined in (7) can be rewritten in the following form

$$\sum_{n=1}^{N} \sum_{s=1}^{S_n} b_{nms} \leq B_m^{\max}, \qquad \forall m \in \mathcal{M}_{r2} \tag{9}$$

$$\sum_{n=1}^{N} \sum_{s=1}^{S_n} b_{nms} \geq B_m^{\min}, \qquad \forall m \in \mathcal{M}_{r2}. \tag{10}$$

The Lagrangian function for (8) using the constraints of (9) and (10) can be expressed as

$$\begin{aligned} & L(B_{ns}, \lambda, \nu, \mu^{(1)}, \mu^{(2)}) = \sum_{n=1}^{N} \sum_{s=1}^{S_n} U_{ns}(B_{ns}) \\ & + \sum_{n=1}^{N} \sum_{s=1}^{S_n} \lambda_{ns}(C_n - \sum_{m \in \mathcal{M}_{ns}} b_{nms}) + \sum_{m \in \mathcal{M}_{r1}} \nu_m (B_m - \\ & \sum_{n=1}^{N} \sum_{s=1}^{S_n} b_{nms}) + \sum_{m \in \mathcal{M}_{r2}} \mu_m^{(1)} (B_m^{\max} - \sum_{n=1}^{N} \sum_{s=1}^{S_n} b_{nms}) \\ & + \sum_{m \in \mathcal{M}_{r2}} \mu_m^{(2)} (\sum_{n=1}^{N} \sum_{s=1}^{S_n} b_{nms} - B_m^{\min}) \end{aligned} \tag{11}$$

where $\lambda = (\lambda_{ns} : n \in \mathcal{N}, s \in \mathcal{S}_n)$ is a matrix of Lagrange multipliers corresponding to the capacity constraint of (5) with $\lambda_{ns} \geq 0$, $\nu = (\nu_m : m \in \mathcal{M}_{r1})$, $\mu^{(1)} = (\mu_m^{(1)} : m \in \mathcal{M}_{r2})$, $\mu^{(2)} = (\mu_m^{(2)} : m \in \mathcal{M}_{r2})$ are vectors of Lagrange multipliers corresponding to the required bandwidth constraints of (6), (9) and (10) respectively, with $\mu_m^{(1)}, \mu_m^{(2)} \geq 0$. The dual function can be expressed as

$$h(\lambda, \nu, \mu^{(1)}, \mu^{(2)}) = \max_{B_{ns} \geq 0} L(B_{ns}, \lambda, \nu, \mu^{(1)}, \mu^{(2)}) \tag{12}$$

and the dual problem corresponding to the primal problem of (8) is

$$\min_{(\lambda, \mu^{(1)}, \mu^{(2)}) \geq 0, \nu} h(\lambda, \nu, \mu^{(1)}, \mu^{(2)}). \tag{13}$$

As the primal problem of (8) is a convex optimization problem, a strong duality exists [10]. The optimal values for the primal and dual problems are equal. As a result, it is appropriate to solve (8) through its dual problem of (13). The maximization problem of (12) can be simplified to

$$\begin{aligned} h(\lambda, \nu, \mu^{(1)}, \mu^{(2)}) \quad &= \quad \sum_{n=1}^{N} \sum_{s=1}^{S_n} \max_{B_{ns} \geq 0} \{U_{ns}(B_{ns}) \\ &- \quad \lambda_{ns} \sum_{m \in \mathcal{M}_{ns}} b_{nms} - \sum_{m \in \mathcal{M}_{r1}} \nu_m b_{nms} \\ &- \quad \sum_{m \in \mathcal{M}_{r2}} (\mu_m^{(1)} - \mu_m^{(2)}) b_{nms}\}. \end{aligned} \tag{14}$$

Consequently, each network BS/AP can solve its own utility maximization problem, expressed as

$$\begin{aligned} \max_{B_{ns} \geq 0} \{U_{ns}(B_{ns}) - \lambda_{ns} \sum_{m \in \mathcal{M}_{ns}} b_{nms} - \sum_{m \in \mathcal{M}_{r1}} \nu_m b_{nms} \\ - \sum_{m \in \mathcal{M}_{r2}} (\mu_m^{(1)} - \mu_m^{(2)}) b_{nms}\}. \end{aligned} \tag{15}$$

The optimum allocation B_{ns} for fixed values of $\lambda, \nu, \mu^{(1)}$ and $\mu^{(2)}$ can be calculated by each network BS/AP by applying the Karush - Kuhn - Tucker (KKT) conditions on (15) [10], and we have

$$\frac{\partial u_{nms}(b_{nms})}{\partial b_{nms}} - \lambda_{ns} - \nu_n - (\mu_m^{(1)} - \mu_m^{(2)}) = 0. \tag{16}$$

Using the utility function of (1), (16) results in

$$b_{nms} = [\frac{1}{\lambda_{ns} + \nu_m + (1 - p_{nms})} - 1]^{+},$$

$$\forall m \in \mathcal{M}_{r1} \tag{17}$$

$$b_{nms} = [\frac{1}{\lambda_{ns} + (\mu_m^{(1)} - \mu_m^{(2)}) + (1 - p_{nms})} - 1]^{+},$$

$$\forall m \in \mathcal{M}_{r2} \tag{18}$$

where the notion $[\cdot]^{+}$ is a projection on the positive orthant to account for the fact that $B_{ns} \geq 0$. The optimum values of $\lambda, \nu, \mu^{(1)}$ and $\mu^{(2)}$ that give the optimum allocation b_{nms} of (17) and (18) can be calculated by solving the dual problem of (13). For a fixed allocation B_{ns}, the dual problem can be simplified to

$$\begin{aligned} &\sum_{n=1}^{N} \sum_{s=1}^{S_n} \min_{\lambda \geq 0} \{\lambda_{ns}(C_n - \sum_{m \in \mathcal{M}_{ns}} b_{nms})\} \\ &+ \sum_{m \in \mathcal{M}_{r1}} \min_{\nu} \{\nu_m (B_m - \sum_{n=1}^{N} \sum_{s=1}^{S_n} b_{nms})\} + \\ &\sum_{m \in \mathcal{M}_{r2}} \min_{\mu^{(1)} \geq 0} \{\mu_m^{(1)} (B_m^{\max} - \sum_{n=1}^{N} \sum_{s=1}^{S_n} b_{nms})\} \\ &+ \sum_{m \in \mathcal{M}_{r2}} \min_{\mu^{(2)} \geq 0} \{\mu_m^{(2)} (\sum_{n=1}^{N} \sum_{s=1}^{S_n} b_{nms} - B_m^{\min})\}. \end{aligned} \tag{19}$$

For a differentiable dual function, a gradient descent method can be applied to calculate the optimum values for $\lambda, \nu, \mu^{(1)}$ and $\mu^{(2)}$ [10], given by

$$\lambda_{ns}(i+1) = [\lambda_{ns}(i) - \alpha_1 (C_n - \sum_{m \in \mathcal{M}_{ns}} b_{nms}(i))]^{+} \tag{20}$$

$$\nu_m(i+1) = \nu_m(i) - \alpha_2 (B_m - \sum_{n=1}^{N} \sum_{s=1}^{S_n} b_{nms}(i)) \tag{21}$$

$$\mu_m^{(1)}(i+1) = [\mu_m^{(1)}(i) - \alpha_3 (B_m^{\max} - \sum_{n=1}^{N} b_{nm}(i))]^{+} \tag{22}$$

$$\mu_m^{(2)}(i+1) = [\mu_m^{(2)}(i) - \alpha_4 (\sum_{n=1}^{N} \sum_{s=1}^{S_n} b_{nms}(i) - B_m^{\min})]^{+} \tag{23}$$

where i is the iteration index and α_j with $j = \{1, 2, 3, 4\}$ is a sufficiently small fixed step size. Convergence towards the optimum solution is guaranteed since the gradient of (19) satisfies the Lipchitz continuity condition [10]. As a result, the resource allocation b_{nms} of (17) and (18) converges to the optimum solution.

V. A DISTRIBUTED MULTI-SERVICE RESOURCE ALLOCATION ALGORITHM

The proposed decomposition method for the optimization problem of (8) has two levels. The first one is a lower level where sub-problems are solved at each network BS/AP to find the optimum resource allocation b_{nms}. The sub-problems are defined in (15), which has the optimum solution of (17) for a CBR service and (18) for a VBR service. The other is a higher level, where the master problem exists. The master problem is defined in (19) and the optimum solution is obtained using the iterative method defined in (20)-(23). The master problem is to set the dual variables $\lambda, \nu, \mu^{(1)}$ and $\mu^{(2)}$ to coordinate the sub-problem solution at each network BS/AP.

Following the classical interpretation of λ_{ns} in economics as the price of resources [12], λ_{ns} gives the price of network n link resources in BS/AP s. Thus, λ_{ns} serves as an indication of the capacity limitation experienced by network n link resources in BS/AP s. When the total traffic load on network n BS/AP s ($\sum_{m \in \mathcal{M}_{ns}} b_{nms}$) reaches the capacity limitation (C_n), the link access price value (λ_{ns}) increases to denote that it is expensive to use that link. On the other hand, ν_m is a coordination parameter used by MTs with CBR service, while $\mu_m^{(1)}$ and $\mu_m^{(2)}$ are coordination parameters used by MTs with VBR service. As a result, ν_m is used by MT m for coordination among different available networks, to ensure that the required bandwidth is met. While $\mu_m^{(1)}$ and $\mu_m^{(2)}$ are used to ensure that the allocated resources for an MT with VBR service lie within the specified required bandwidth range.

The Lagrange multiplier λ_{ns} can be calculated at each network BS/AP based on its capacity limitation and the total load experienced in the coverage area. The Lagrange multiplier ν_m is calculated at each MT with CBR service, while the Lagrange multipliers $\mu_m^{(1)}$ and $\mu_m^{(2)}$ are calculated at each MT with VBR service. The multipliers ν_m, $\mu_m^{(1)}$ and $\mu_m^{(2)}$ are calculated based on the allocated bandwidth from different wireless access networks BSs/APs and its required bandwidth. Each BS/AP starts with an initial feasible value for its link access price. Similarly, each MT starts with an initial feasible value for its coordination parameter. Each BS/AP performs its bandwidth allocation to a given MT based on its link access price value, priority parameter value and the coordination parameter value for that MT. Each BS/AP then updates its link access price value. Also, the value of ν_m and the difference $\mu_m^{(1)} - \mu_m^{(2)}$ are updated and broadcasted by the MTs to the different available wireless access networks through the different interfaces of the MT, in order to perform coordination among the resource allocation from different networks so that the required bandwidth can be met eventually.

VI. NUMERICAL RESULTS AND DISCUSSION

This section presents analytical results for problem (8) using the proposed distributed algorithm. A geographical region that is entirely covered by an IEEE 802.16e WMAN BS and partially covered by a 3G cellular network BS and an IEEE 802.11b WLAN AP is considered [6]. As a result, $\mathcal{N} = \{1, 2, 3\}$ with the WMAN, cellular network and WLAN indexed as 1,2 and 3 respectively. Three service areas can be distinguished, $\mathcal{K} = \{1, 2, 3\}$. In area 1, service from the WMAN BS only is available. In area 2, the WMAN and the cellular network BSs services are available. In area 3, services from all three networks BSs/AP are available. For the priority mechanism, different networks can set different costs on their resources using the priority parameter p_{nm}. Since the cellular network has the lowest capacity among all the available networks, it sets the highest cost on its resources so that it can devote its resources to its own subscribers. The WMAN and the WLAN both have a high capacity, yet the WMAN BS covers a larger area with more MTs, hence the WMAN sets a higher cost on its resources than the WLAN with its limited coverage area. Let M_{nkr} denotes the number of subscribers of network n in service area k with service r, where $r = 1$ represents a CBR service while $r = 2$ represents a VBR service. The system parameters are listed in Table II, where the required resources units are in Mbps, and the given priority parameters are for the networks users.

Figures 2-5 shows various bandwidth allocation results versus the number of ongoing CBR connections for the WLAN subscribers in area 3 (M_{331}).

TABLE II
SYSTEM PARAMETERS

Parameter	Value	Parameter	Value	Parameter	Value
C_1	20	p_{2m}	0.5	M_{132}	5
C_2	2	p_{3m}	0.8	M_{221}	8
C_3	11	M_{111}	10	M_{222}	8
B_m	0.256	M_{112}	10	M_{231}	5
$B_m^{\min}$	0.256	M_{121}	7	M_{232}	5
$B_m^{\max}$	0.512	M_{122}	7	M_{332}	5
p_{1m}	0.6	M_{131}	5	$\eta_{1,2}$	1

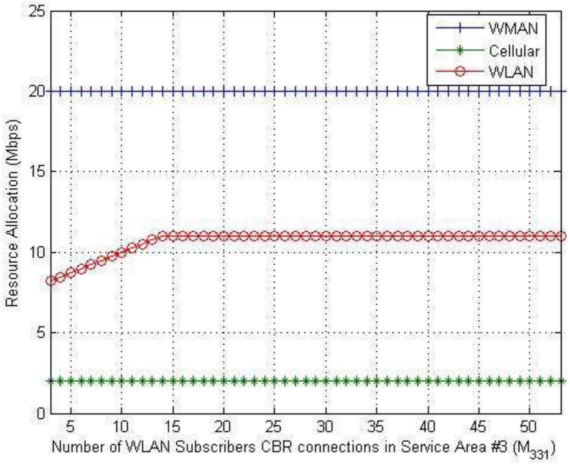


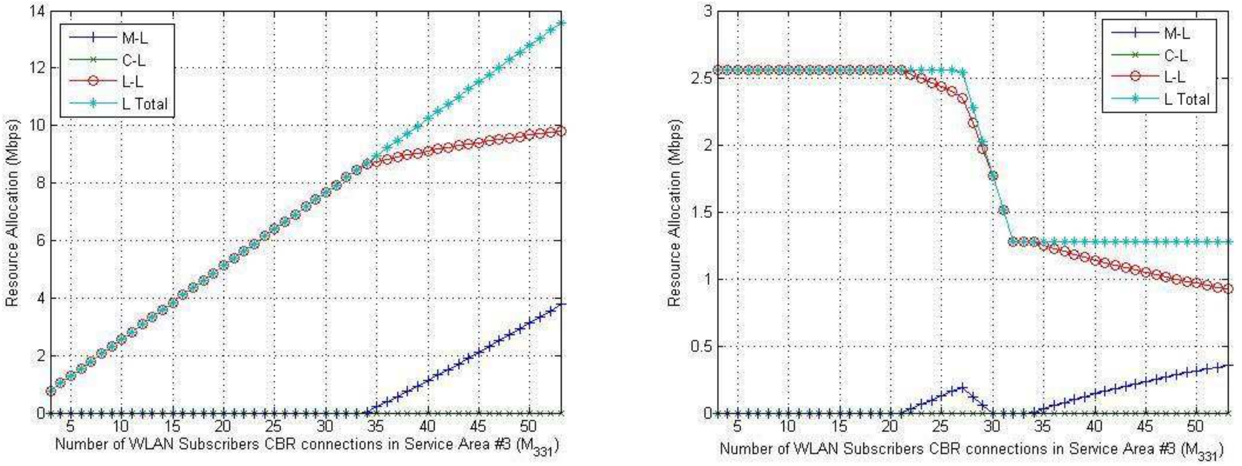
Fig. 2. Total bandwidth allocation by each network

Figure 2 shows the total bandwidth allocation by each network BS/AP. The WMAN and the cellular network BSs reach their capacity limitation, independent of M_{331}. The WLAN AP increases its total allocation with M_{331} to accommodate more subscribers. At $M_{331} = 14$, the WLAN AP also reaches its capacity limitation.

In the following, we study the total bandwidth allocation from each network BS/AP to different subscribers in service area 3.

Figure 3a shows the total bandwidth allocation by each network BS/AP for the CBR WLAN subscribers in area 3. With the priority mechanism, the WLAN AP supports its own subscribers to avoid the high cost of the WMAN and the cellular network. As a result, the WLAN AP bandwidth allocation (L-L) increases with M_{331} to accommodate more subscribers, while the allocation from the WMAN (M-L) and the cellular network (C-L) BSs is equal to zero. However, for $M_{331} > 34$, the WLAN AP does not have sufficient bandwidth to support its subscribers. As a result, the WMAN BS increases its allocation to support the WLAN subscribers. The support comes from the WMAN BS as its bandwidth have a lower cost than those from the cellular network.

Figure 3b shows the total bandwidth allocation by each network BS/AP for the VBR WLAN subscribers in area 3. For $M_{331} \geq 22$, the WLAN AP decreases its allocation (L-L) to support its CBR subscribers. As a result, the WMAN BS increases its allocation (M-L) to keep the total allocation constant at the maximum required bandwidth (512 kbps for each VBR call). However, for $M_{331} > 27$, any further increase in the WMAN BS allocation would degrade the WMAN BS allocation to its VBR subscribers. The priority mechanism



(a) CR

(b) VR

Fig. 3. Total bandwidth allocation by each network to the (a) CBR and (b) VBR WLAN subscribers

does not allow this, since it gives a higher priority on the WMAN BS bandwidth to the WMAN subscribers. As a result, the WMAN BS decreases its allocation to the VBR WLAN subscribers, and the total allocation is degraded towards the minimum required bandwidth. For $M_{331} > 34$, the WLAN AP decreases the allocation to support the CBR WLAN subscribers. As a result, the WMAN BS increases its allocation in order not to violate the minimum required bandwidth (256 kbps for each VBR call).

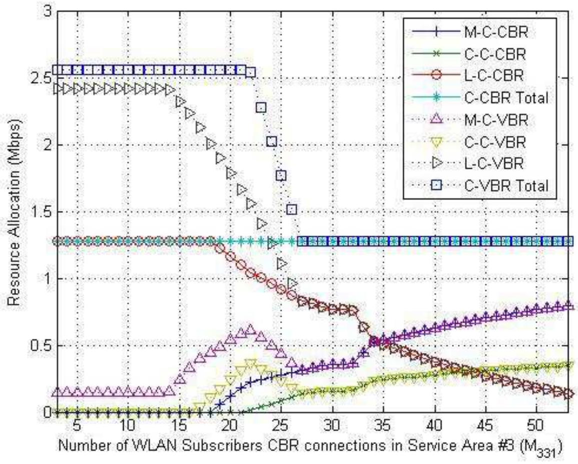


Fig. 4. Total bandwidth allocation by each network to the cellular network subscribers in area 3

Figure 4 shows the total bandwidth allocation by each network BS/AP to the cellular network subscribers in service area 3. The total bandwidth allocation (C-CBR Total) to the CBR cellular network subscribers comes from the WLAN AP (L-C-CBR). The allocation from the cellular network (C-C-CBR) is zero, since it uses its bandwidth to support its subscribers in area 2. The allocation from the WMAN BS (M-C-CBR) is zero since it imposes a higher cost on its bandwidth than the WLAN. However, for $M_{331} > 18$, the WLAN AP starts to decrease its allocation for cellular network subscribers to support its own subscribers. This is compensated by an increase in the WMAN BS allocation, to keep the total allocation constant at the required bandwidth (256 kbps for each CBR call). For $M_{331} > 21$, more bandwidth are required from the WMAN BS to keep the total allocation constant, but in order to reduce the required amount of bandwidth from the WMAN due to the associated high cost, the cellular network increases its allocation to support its own subscribers. The total allocation is always constant at the required bandwidth. For the VBR subscribers, as M_{331} increases, the WLAN AP reduces its allocation to the VBR cellular network subscribers to support its own subscribers. As a result, the WMAN BS increases its allocation to keep the total bandwidth allocation (C-VBR Total) at its maximum required bandwidth (512 kbps for each VBR call). For $M_{331} > 17$, the cellular network BS increases its allocation to reduce the amount of bandwidth required from the WMAN BS due to its high cost. For $M_{331} > 22$, any further increase in the WMAN BS allocation would degrade the WMAN BS allocation to its VBR subscribers. As a result, the WMAN BS decreases its allocation. The cellular network BS also decreases its allocation to support its CBR subscribers in this area. Hence, the total allocation starts to degrade towards the minimum required bandwidth. For $M_{331} > 26$, the WMAN and the cellular network BS increase their allocation to compensate for the reduction in the WLAN AP allocation and keep the total bandwidth allocation constant at the minimum required bandwidth.

Figure 5 shows the total bandwidth allocation by each network to the WMAN subscribers in service area 3. For the CBR and VBR calls, most of the allocated bandwidth comes from the WMAN BS (M-M-CBR and M-M-VBR) as compared to the WLAN AP allocation (L-M-CBR and L-M-VBR), in order to reduce the associated cost of the WLAN bandwidth. The allocation from the cellular network BS (C-M-CBR and C-M-VBR) is zero, as it uses its bandwidth to support its subscribers in areas 2 and 3. For $M_{331} > 13$,

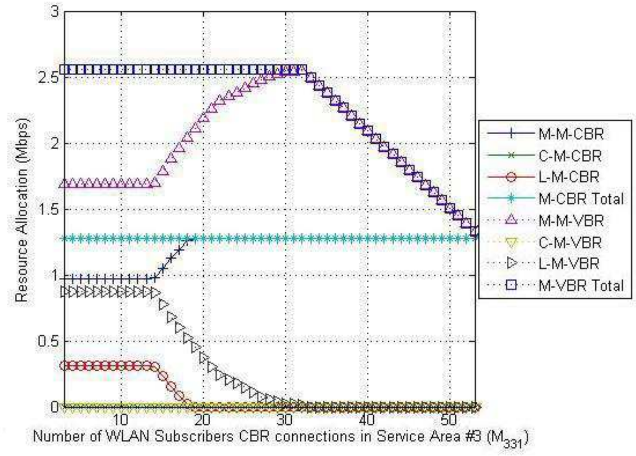


Fig. 5. Total bandwidth allocation by each network to the WMAN subscribers in area 3

the WLAN AP decreases its allocation to support its own subscribers. As a result the WMAN BS increases its allocation to support its own subscribers. For $M_{331} > 18$, all the required bandwidth to service CBR calls (M-CBR Total) in area 3 come from the WMAN BS. For $M_{331} > 32$, the WMAN BS reduces its allocation to the VBR WMAN subscribers towards the minimum required bandwidth to support the WLAN subscribers (refer to Figure 3).

From the results in Figures 3-5, service degradation of VBR calls starts from the cellular network subscribers because these users depend heavily on other networks to satisfy their bandwidth demands. Due to the priority mechanism, these networks allocate their bandwidth first to their own subscribers, leading to a reduced bandwidth allocated to the VBR calls of cellular network subscribers.

VII. CONCLUSION

In this paper, a distributed multi-service resource allocation algorithm in a heterogeneous wireless access environment is proposed. The algorithm has the following features: 1) It is a distributed algorithm in a sense that each network BS/AP solves its own utility maximization problem and performs its own resource allocation. Hence, no central resource manager is required. This is very essential for the algorithm to be implemented in a practical environment where different networks are operated by different service providers; 2) The algorithm supports different classes of services, namely CBR and VBR services; 3) Each MT can obtain its required bandwidth from all the available networks using its multi-homing capability; 4) The MTs play an active role in the allocation operation by coordinating the allocation from different networks such that the required bandwidth is satisfied, 5) A priority mechanism is employed to give a higher priority for each network subscribers on its resources as compared to other users. Cooperation among different networks is achieved to support all the CBR and VBR calls.

REFERENCES

[1] M. Ismail and W. Zhuang, “A Distributed resource allocation algorithm in heterogeneous wireless access medium,” *Proc. IEEE ICC 2011*, June 2011.

[2] M. Kassar, B. Kervella, and G. Pujolle, “An overview of vertical handover strategies in heterogeneous wireless networks,” *Computer Communications*, vol. 31, no. 10, pp. 2607-2620, June 2008.

[3] X. Pei, T. Jiang, D. Qu, G. Zhu, and J. Liu, “Radio resource management and access control mechanism based on a novel economic model in heterogeneous wireless networks,” *IEEE Trans. Vehicular Technology*, vol. 59, no. 6, pp. 3047-3056, July 2010.

[4] W. Shen, and Q. Zeng, “Resource management schemes for multiple traffic in integrated heterogeneous wireless and mobile networks,” *Proc. 17th Int. conf. ICCCN*, pp. 105-110, August 2008.

[5] I. Blau, G. Wunder, I. Karla, and R. Sigle, “Decentralized utility maximization in heterogeneous multicell scenarios with interference limited and orthogonal air interfaces,” *EURASIP J. Wireless Communications and Networking*, 2009.

[6] C. Luo, H. Ji, and Y. Li, “Utility based multi-service bandwidth allocation in the 4G heterogeneous wireless access networks,” *Proc. IEEE WCNC*, April 2009.

[7] D. Niyato and E. Hossain, “Bandwidth allocation in 4G heterogeneous wireless access networks: A noncooperative game theoretical approach,” *Proc. IEEE GLOBECOM’06*, 2008.

[8] K. Chebrolu, and R. Rao, “Bandwidth aggregation for real time applications in heterogeneous wireless networks,” *IEEE Trans. Mobile Computing*, vol. 5, no. 4, pp. 388-402, April 2006.

[9] W. Song, Y. Cheng, and W. Zhuang, “Improving voice and data services in cellular/WLAN integrated network by admission control,” *IEEE Trans. Wireless Communications*, vol. 6, no. 11, pp. 4025-4037, Nov. 2007.

[10] D. P. Bertsekas, *Non-linear programming*, Athena Scientific, 2003.

[11] H. Shen and T. Basar, “Differentiated Internet pricing using a hierarchical network game model,” *Proc. 2004 American Control Conference*, pp. 2322-2327 vol.3, 2004.

[12] L. S. Lasdon, *Optimization theory for large systems*, Macmillan series in operations research, 1970.

[13] M. Chiang, S. Low, A. Calderbank, and J. Doyle, “Layering as optimization decomposition: a mathematical theory of network architectures,” *Proc. IEEE*, vol. 95, no. 1, pp. 255-312, January 2007.

[14] D. Palomar and M. Chiang, “A tutorial on decomposition methods for network utility maximization,” *IEEE J. Select. Areas Communications*, vol. 24, no. 8, pp. 1439-1451, August 2006.

[15] D. Palomar and M. Chiang, “Alternative distributed algorithms for network utility maximization: framework and applications,” *IEEE Trans. Automatic Control*, vol. 52, no. 12, pp. 2254-2269, December 2007.

[16] L. Koutsopoulos and G. Losifidis, “A framework for distributed bandwidth allocation in peer-to-peer networks,” *Performance Evaluation*, vol. 67, no. 4, pp. 285-298, April 2010.

[17] J. Chen, W. Xu, S. He, Y. Sun, P. Thulasiramanz, and X. Shen, “Utility-Based Asynchronous Flow Control Algorithm for Wireless Sensor Networks,” *IEEE J. Selected Areas of Communications*, Vol. 28, No. 7, pp. 1116-1126, September 2010.

Muhammad Ismail (S’10) received the BSc. and MSc. in Electrical Engineering from Ain Shams University, Cairo, Egypt in 2007 and 2009, respectively. He is a research assistant and currently working towards his Ph.D. degree at the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research interests include distributed resource allocation, quality-of-service provisioning, call admission control, and cooperative networking.

Weihua Zhuang (M93-SM01-F08) is a Professor and a Tier I Canada Research Chair in wireless communication networks, at the Department of Electrical and Computer Engineering, University of Waterloo, Canada. Her current research focuses on resource allocation and QoS provisioning in wireless networks. She is the Editor-in-Chief of IEEE Transactions on Vehicular Technology, the Technical Program Committee (TPC) Symposia Chair of the IEEE Globecom 2011, and a Fellow of the Canadian Academy of Engineering (CAE).