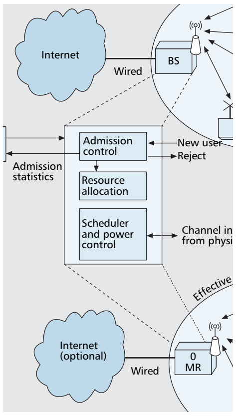
RADIO RESOURCE MANAGEMENT AND PROTOCOL ENGINEERING FOR IEEE 802.16

DYNAMIC RESOURCE ALLOCATION IN OFDMA WIRELESS METROPOLITAN AREA NETWORKS

SYED HUSSAIN ALI, KI-DONG LEE, AND VICTOR C. M. LEUNG,
THE UNIVERSITY OF BRITISH COLUMBIA



The authors present important resource allocation problems in IEEE 802.16 wireless metropolitan area networks employing orthogonal frequency division multiple access.

ABSTRACT

In this article we present important resource allocation problems in IEEE 802.16 wireless metropolitan area networks employing orthogonal frequency division multiple access. We first highlight the unique aspects of these networks and identify challenges and opportunities provided by the physical and medium access control layers. Next, we concentrate on four interrelated resource allocation problems: dynamic subcarrier allocation, adaptive power allocation, admission control, and capacity planning. We describe solution techniques, provide preliminary results, and discuss open problems for future research.

INTRODUCTION

Orthogonal frequency-division multiple access (OFDMA) is a physical layer specification for IEEE 802.16 [1] fixed wireless metropolitan area networks (WMANs), which is later extended to support combined fixed and mobile subscriber stations (SSs) in IEEE 802.16e [2]. These specifications support non-line-of-sight (NLOS) operation in the 2–11 GHz bands. Furthermore, the corresponding medium access layers support mesh networking as an option in addition to the required point-to-multipoint (PMP) topology. In the mesh topology, an SS can be directly connected to another SS and/or to the base station (BS), and it can route traffic on behalf of another SS by acting as a relay station. Both OFDMA and mesh topology offer very attractive solutions in providing high performance and flexible deployment for fixed and mobile broadband wireless access. However, NLOS and SS mobility exacerbate channel impairments due to path loss, shadowing and multipath fading, so that it is critical to design efficient methods to combat these impairments and to develop effective resource allocation solutions that exploit mesh mode operations.

OFDMA builds on orthogonal frequency-division multiplexing (OFDM), which is immune to intersymbol interference and frequency selective fading, as it divides the frequency band into a group of mutually orthogonal subcarriers, each having a much lower bandwidth than the coherence bandwidth of the channel. In multi-user environments, OFDMA provides another degree of freedom by allowing dynamic assignment of subcarriers to different users at different time instances, to take advantage of the fact that at any time instance channel responses are different for different users at different subcarrier frequencies. Thus, dynamic subcarrier assignments (DSA) to multiple users can be employed to improve the system data rate. Since the achievable data rate is a function of the power allocation (i.e., the data rate increases with transmit power and vice versa), it is expected that adaptive power allocation (APA) can further improve the system data rate. DSA and APA have attracted considerable research interests [3, 4, references therein].

In OFDMA systems, mesh topology allows each SS to have many more options in choosing relays from neighboring SSs. Compared to PMP topology, multihop routing (MHR) over a mesh topology allows SSs to exploit more reliable channels at lower transmission costs. This is because even if the channel condition between an SS and the BS is bad, the channel condition between the SS and one of its neighboring SSs may be good, and the SS may successfully route its traffic via the neighbor SS at a relatively lower cost than the BS. The opportunities for lower transmission costs by MHR increase with the node density in the network. This effect causes the system capacity to increase with the number of nodes, and is commonly referred as *multi-user diversity* or *user cooperation diversity*. Also, an SS may choose more than one neighboring SSs for multipath routing, which increases the reliability of packet delivery over error-prone wireless channels at the expense of increased consumption of network resources. In general, the additional flexibility that mesh topology brings to an OFDMA system makes resource allocation problems more complex, resulting in new and interesting research topics in the area of resource allocation [5].

The extension of fixed WMANs to serve

1536-1284/07/$20.00 © 2007 IEEE

mobile SSs in 802.16e [2] provides new challenges and opportunities in different resource allocation problems. A mobile SS experiences significant channel variations over the duration of its connection; therefore, an efficient DSA is even more critical for system performance than for a static assignment. Furthermore, mobility results in handoffs that need to be taken into account in efficient call admission control (CAC) and capacity planning (CP) solutions.

In mesh networks, the CAC problem is made more interesting by the option of setting up a new connection or handing off an ongoing connection to/from an SS via some neighboring SS, in the situation where the channel condition to the BS is not good. This optional multihop routing of connections can improve the utilization of network resources as well as the overall energy efficiency of all SSs, ultimately improving the quality of service (QoS). Furthermore, in systems where a large fraction of users experience *group mobility*, which is commonly observed in mass transportation systems (e.g., bus or train), the bulk arrivals of handoff requests have a highly nonstationary characteristic that degrades traffic performance from that of a stationary system with the same offered traffic. Thus, it is desirable to develop advanced techniques for CAC and CP that will not lower system utilization in the presence of group mobility.

In this article we discuss four closely related resource allocation problems in OFDMA WMANs: DSA, APA, CAC, and CP. We briefly introduce a simple but practical linear-integer-programming-based DSA formulated by a maximal bipartite matching problem and review a few solution techniques. Furthermore, we present some guidelines for CP in the presence of group mobility in OFDMA networks. We report some preliminary results, identify open problems, and provide directions for future research.

This article is organized as follows. The next section presents the structure and properties of IEEE 802.16 OFDMA WMANs and the relationships between different resource allocation problems. We discuss DSA and APA problems, their solutions, and some open problems. We present the CP and CAC problems with several open research issues. We then conclude the article.

STRUCTURE AND PROPERTIES OF OFDMA WMANS

NETWORK STRUCTURE OF IEEE 802.16 WMANS

IEEE 802.16 [1] specifies the physical and medium access control layer structures and protocols for WMANs with support for PMP and optional mesh modes. In the PMP mode (Fig. 1a), the network has a cellular structure where each cell has a central BS that routes traffic between SSs located within the cell and the global network. In contrast, in the mesh mode (Fig. 1b), traffic can be routed through other SSs acting as relay stations by means of MHR, or exchanged directly between SSs. Within a mesh WMAN, a node that has a direct connection to the backhaul network outside the mesh WMAN is termed a mesh BS. All the other nodes of a mesh WMAN are termed mesh SSs (MSSs). There are two possible network structures in the mesh mode: full mesh and infrastructure mesh. In full mesh, MSSs communicate only with each other by MHR and they do not have access to any backhaul network. In the infrastructure mesh, access to a backhaul network is provided through one or more mesh BSs that act as mesh routers (MRs). In Fig. 1b MSSs numbered 3, 5, and 6 form a full mesh network in which these three nodes communicate among themselves without any access to the backhaul network, while MSSs numbered 1, 2, 4, and 6 form an infrastructure mesh network with node 0 acting as the mesh BS or MR to provide access to the backhaul network on behalf of other MSSs.

RESOURCE ALLOCATION

The model shown in Fig. 1 also depicts different resource allocation problems for WMANs: CP, CAC, DSA, and APA. These resource allocation problems are related to different timescales and different granularity levels in terms of the resources being allocated. CP is a static optimization problem that allocates the overall bandwidth capacity for each BS or each link based on long-term traffic statistics. Solutions are usually computed offline and have the longest time horizon in terms of validity compared to other resource allocation problems. CAC decisions are made at the call or connection level, to allocate bandwidth to new connections based on the available capacity determined from CP. Thus, CAC is performed online in real time, and the decision epochs have a medium time horizon (i.e., the next call arrival or departure). Finally, DSA/APA allocates carrier/power on a per frame or burst level for the admitted connections. Therefore, DSA/APA solutions have the shortest time horizon in terms of their validity, and the corresponding algorithms and processes should operate online in real time. Hence, efficient dynamic control is most important for DSA/APA. This model also contains feedback from the CAC block to the CP block in order to tune the latter with the help of real-time admission statistics. Moreover, the scheduler and power control blocks rely on real-time channel statistics from the physical layer.

In the PMP mode, the BS plays a central role in resource allocation. In this mode CAC and DSAs/APAs are closely related as admission of a new call requires updates in subcarrier and power allocations to accommodate the requirements of the new and existing calls. Also, a portion of subcarrier and power resources may be held in reserve for handoff calls to provide them with the higher call-level QoS commonly required for user satisfaction.

In the mesh mode resource allocation can be fully distributed, fully centralized, or a combination of both. In general, resource allocation in mesh mode is more complex and requires coordination between multiple nodes. The mesh mode not only retains the relationships between different resource allocation problems of PMP mode but also imposes more complicated dependencies between them. In mesh mode DSA can be implemented with or without frequency reuse.

There are two possible network structures in the mesh mode: full mesh and infrastructure mesh. In full mesh, MSS's communicate only with each other by MHR and they do not have access to any backhaul network. In the infrastructure mesh, access to a backhaul network is provided through one or more mesh BSs that act as mesh routers.

DSA with spatial frequency reuse will increase overall capacity. Furthermore, intelligent APA may cause less interference in neighboring cells, thus increasing capacity. MHR via neighboring MSSs can improve performance and capacity at the expense of increased latency.

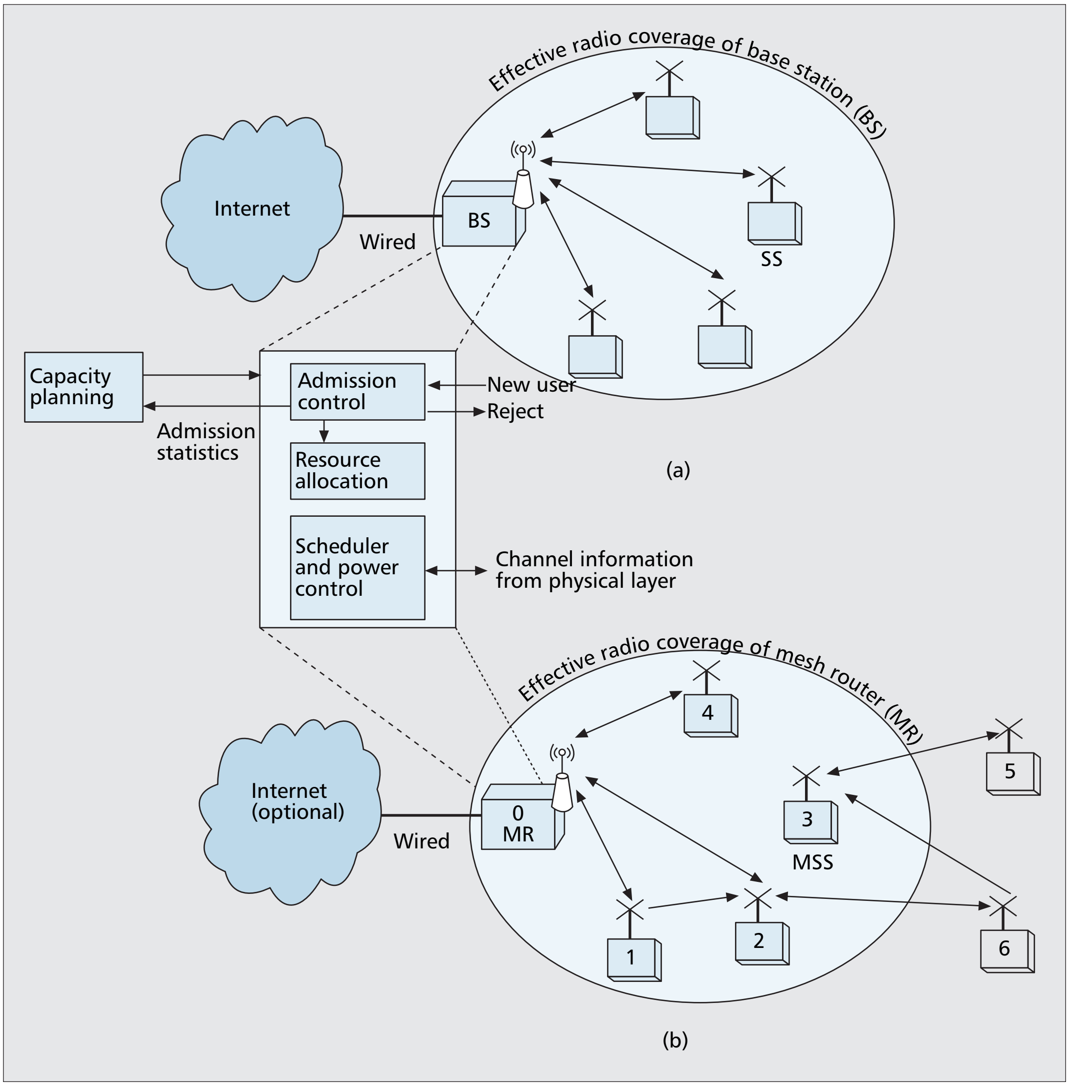


Figure 1. *A reference model of interfaces between units responsible for resource allocation: a) point-to-multipoint network; b) multihop network (with mesh mode support).*

DSA with spatial frequency reuse will increase overall capacity. Furthermore, intelligent APA may cause less interference in neighboring cells, thus increasing capacity. MHR via neighboring MSSs can improve performance and capacity at the expense of increased latency.

SCHEDULING IN OFDMA: SUBCARRIER ASSIGNMENT AND POWER ALLOCATION

OPTIMIZATION-THEORETIC STUDIES

Most optimization-theoretic approaches to solve the DSA and APA problems assume continuous data rates and an infinite number of subcarriers in the OFDMA system (e.g., [3]). These ideal conditions render the DSA and APA problems tractable and enable the system data rate to be maximized through a greedy algorithm. In the case of greedy assignments, a subcarrier is assigned to only one user that has the maximum channel gain on that subcarrier. However, this trivial solution is highly unfair because some users close to the BS can monopolize all the subcarriers while other users starve, not receiving any assignment. Thus, it is desirable that the DSA problem should have fairness constraints on individual users' performance, or the objective function should be designed so that it implicitly enforces some fairness among users, as in [3, 5]. In general, the problem can be formulated as the minimization of the overall transmit power under fairness constraints on users' data rates or the maximization of a particular fairness criterion. Where the problem is convex, the solution is straightforward. However, in most cases, since the actual problem is combinatorial in nature, it is attractive to use efficient heuristic solution techniques.

The role of APA is to allocate power levels on the respective subcarriers so that the data rate is maximized subject to a certain constraint on the transmit power. The solution of this problem can be found using the Lagrange multiplier technique, even though the problem is combinatorial or non-convex. According to the solution, more power is allocated when the channel gain is high and vice versa. This solution algorithm is popularly known as water-filling. The joint DSA and APA solution determines subcarrier and power allocations to maximize system performance. In principle, allocating different power levels to individual subcarriers should improve performance. However, previous studies [3, 6] have shown that performance improvements are marginal over a wide range of signal-to-noise ratios (SNRs) due to the statistical effects. Therefore, a simpler solution involving DSA with equal power per subcarrier is preferred over a more complex joint DSA and APA solution.

PRACTICAL SCHEMES

Practical systems like IEEE 802.16 WMANs have a finite number of subcarriers and discrete data rates, which do not satisfy the ideal assumptions that enable optimization-theoretic solutions. Therefore, such solutions only provide guidelines on how a real-time suboptimal solution can be obtained. In the following we give an outline of the problem structure for a practical system.

Consider a system with N subcarriers, K users, and M discrete data rates, where a higher data rate is used when a higher SNR is realized. Thus, the joint DSA and APA algorithm will minimize the overall transmit power and contain three dimensional binary decision variables. APA does not significantly improve system performance [3, 6]; therefore, we can reduce the decision variables by one dimension by assuming that power is equally distributed among all allocated subcarriers. The resulting problem is aimed at maximizing the overall system data rate. Let $r_{k,n}$ denote the resulting best data rate for the kth user on the nth subcarrier, and $\mathbf{R}$ be the corresponding matrix with $K \times N$ entries. The Hungarian method can be used to solve this new assignment problem if matrix $\mathbf{R}$ becomes a square matrix. In general, $K \leq N$; therefore, $\mathbf{R}$ is not a square matrix. However, if the minimum data rate constraint for each user can be expressed in terms of an integer number of subcarriers, $\mathbf{R}$ can easily be converted into a square matrix of dimension $N \times N$ by repeating rows in $\mathbf{R}$. The resulting problem formulation, with decision variable $\gamma_{k,n} \in \{0, 1\}$, is as follows:

$$\begin{aligned} &\text{maximize} \sum_{k=1}^{N}\sum_{n=1}^{N} r_{k,n}\gamma_{k,n}, \\ &\text{subject to} \sum_{k=1}^{N} \gamma_{k,n} = 1 \quad \forall n \in \{1, \dots N\}, \\ &\qquad\qquad \sum_{n=1}^{N} \gamma_{k,n} = 1 \quad \forall k \in \{1, \dots N\}. \end{aligned} \tag{1}$$

The above problem is known as a *maximal bipartite matching problem* [7] where *users* and *subcarriers* form two disjoint sets of vertices of a graph. Every edge of the graph has one endpoint in the users set and the other endpoint in the subcarriers set, and a weight that represents the achievable data rate. The objective is to find a mapping between subcarriers and users such that only one edge connects two graph nodes, and the sum of weights is maximum for all selected edges.

Level (l)	Modulation (coding rate)	Information bytes/OFDM symbol	Rate at 5 MHz (Mb/s) (r_l)	Required SNR (dB) (η_l)
1	QPSK (1/2)	24	4.03	5
2	QPSK (3/4)	36	6.04	8
3	16-QAM (1/2)	48	8.06	10.5
4	16-QAM (3/4)	72	12.09	14
5	64-QAM (1/2)	72	12.09	16
6	64-QAM (2/3)	96	16.12	18
7	64-QAM (3/4)	108	18.14	20

■ **Table 1.** *Modulation and coding parameters for IEEE 802.16 WMANS.*

The Hungarian method transforms the problem in Eq. 1 into a combinatorial optimization problem, and solves it in $O(N^3)$. There are several less complex heuristic methods that iteratively solve this problem. For example, for a fully connected bipartite graph, [7] gives an iterative solution that finds the maximum element in the augmented matrix $\mathbf{R}$ at every iteration, assigns an available subcarrier to the corresponding user, and removes the associated row-column pair from future selection. This procedure is repeated until all the subcarriers are assigned.

To compare this iterative heuristic method with the Hungarian method, we perform a simple simulation for an OFDMA system with 240 subcarriers. Scheduling is performed every frame based on the SNR levels of all users on all subcarriers. Table 1 lists the OFDMA WMAN [2] modulation and coding parameters for different levels of SNR (η_l) and corresponding data rates (r_l) for $l = 1, \dots, 7$. For comparison purposes, we consider the average uncoded data rate (measured in megabits per second) as a performance metric.

In the simulation we construct the matrix $\mathbf{R}$ using continuously varied $r_{k,n}$ values based on the instantaneous SNR values by expressing data rate as a piece-wise linear continuous function of SNR using the points defined in Table 1. The reason for considering continuous data rates is to avoid ties in solving problem in Eq. 1. The continuous SNR values are mutually independent exponential random variables with mean values uniformly distributed between [0, 16] dB. After finding the optimal subcarrier allocations through the solution of Eq. 1, we map the SNR values to discrete data rates through a step function. A generic SNR value x maps to a data rate level r_l in Table 1 if $\eta_l \leq x < \eta_{l+1}$ where $\eta_{l+1} = \infty$ for $l = 7$.

Figure 2 compares the average data rates achieved by the Hungarian and iterative heuristic methods at various user populations. The figure shows that although its complexity is only $O(N)$, the iterative heuristic method has only slightly inferior performance to the Hungarian method, which has a complexity of $O(N^3)$. More sophisticated iterative heuristics can achieve per-

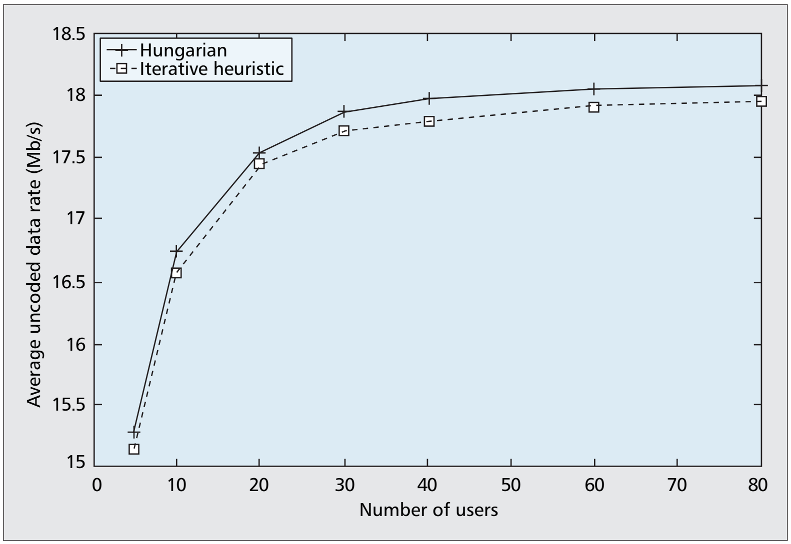


Figure 2. *Comparison of the Hungarian method with an iterative heuristic [7] to solve the bipartite graph matching problem.*

formance comparable to that of the Hungarian method.

OPEN PROBLEMS

Most published studies on DSA are based on single-cell scenarios where channels are assigned irrespective of users' physical locations. However, three particular properties of OFDMA mesh networks should be considered. First, OFDMA subcarriers are susceptible to interference from neighboring cells. This interference is significant at the cell boundary for the full intercell frequency reuse case. Thus, DSA and APA should take this factor into account. Second, MHR provides an opportunity to reduce intercell interference as shown in Fig. 3, where BS A can reduce its transmit power and hence the interference to neighboring cells if it routes traffic to MSS b via MSS a instead of directly, since MSS a is closer to BS A than MSS b. Although interference reduction in this case is at the expense of a slight increase in delivery delay, MHR is an attractive solution for delay-tolerant traffic. Also, by freeing resources at the BS, the multihop DSA can improve some other QoS metrics (e.g., new connection blocking and handoff dropping probabilities). Third, intracell spatial frequency reuse (i.e., the same subcarrier assigned to users who are sufficiently far apart within a cell) can increase the overall capacity of mesh mode.

Initial results of multihop DSA in mesh without intracell frequency reuse have been reported in [5]. Each mesh BS gathers resource requests from all the MSSs within a certain hop range. It determines the amount of granted resources for each link within the area under its control and communicates these grants to all the MSSs within the hop range. The grant messages contain parameters for predetermined algorithms that are run by the MSSs to compute their individual schedules.

Several open problems remain to be addressed pertaining to application of DSA and APA in mesh networks, whether full mesh or infrastructure mesh.

Multihop frequency reuse: Even in practical cellular networks that do have a regular hexagonal topology, frequency reuse is an NP-hard problem commonly solved by *graph coloring*. In multihop networks the frequency reuse issue in channel allocation is much more difficult and therefore challenging. This is largely because alternate MSSs can be selected as relays in MHR, so the actual network topology depends on the routing decisions. Therefore, developing reliable architectures as well as efficient techniques for frequency reuse is an interesting future item (e.g., [8] considers frequency reuse for wireless LANs). Whereas APA does not significantly improve system performance in PMP mode, in mesh mode APA could play an important role in (efficient) spatial reuse of frequency between MSS-to-MSS links by lowering interference to the neighborhood. This possibility will need further investigation.

Overhead and scalability of centralized scheduling: If a central scheduler is employed in a mesh network, many network resources will be consumed by the signaling overhead to provide local status to the central controller in a timely manner. This overhead as well as the computational load of the central scheduler grow with network size, making centralized scheduling not scalable. Thus, it is preferred to adopt scalable distributed scheduling techniques so that the network can experience lower signaling and scheduling burdens [5]. Developing reliable distributed architectures and algorithms for such scalable networks is attracting much research interest.

QoS and fairness for multiple service classes: As the major service shifts from voice to multimedia, the scheduler needs to account for the diverse bandwidth, link quality, and delay requirements of different classes of service. Thus, it is increasingly challenging to come up with computationally efficient scheduling mechanisms for mesh networks that maintain fairness for different users while maximizing network utilization.

CONNECTION ADMISSION CONTROL AND CAPACITY PLANNING

CONNECTION ADMISSION CONTROL

In cellular wireless networks the utilization of system resources by new calls is often kept below a threshold level to accommodate handoff connections because service providers are obligated to provide a minimum QoS to subscribers even as they aim to maximize bandwidth utilization. Thus, when a mobile SS engaged in a call is handed off to a new cell, it may receive a higher priority for channel allocation by the new BS than new calls originating in the cell. Even with resource reservation, connections can still be dropped due to fluctuations in the received SNRs at the mobile SSs, especially for those located near the edges of cells. In the case of OFDMA mesh WMANs, CAC becomes interesting due to the possibility of setting up a new connection or handing off an ongoing connection to/from an MSS via some neighboring MSS.

In MHR, when a relay MSS moves closer to the BS, the connection can be further handed off from the relay MSS to the BS. The use of MHR can improve both resource utilization and QoS. MHR is known to reduce overall power utilization in the system, although efficient channel allocation remains to be addressed. The QoS improves because the improved efficiency of resource utilization increases the capability of call accommodation.

CAC for OFDMA WMANs has recently started to receive increasing attention among researchers. In [9] the authors studied two CAC schemes for such networks. One scheme sets a threshold level on the number of ongoing connections. A new connection is accepted as long as the total number of connections after admission does not exceed the threshold level. The second CAC scheme admits a connection with a certain probability based on the queue status. In [10] the authors proposed a framework for joint CAC, DSA, and MHR in OFDMA mesh networks. The framework supports QoS constraints, such as average delay and packet dropping probability. According to this framework, a new connection request submits its QoS requirements to the nearest BS. The nearest (originating) BS sends connection initiation requests to BSs along possible paths of the connection. These BSs schedule subcarrier assignment for the new connection in a distributed fashion and send allocated channel information back to the originating BS. The originating BS estimates the end-to-end delay, and if it satisfies the required QoS metric, the BS grants the connection.

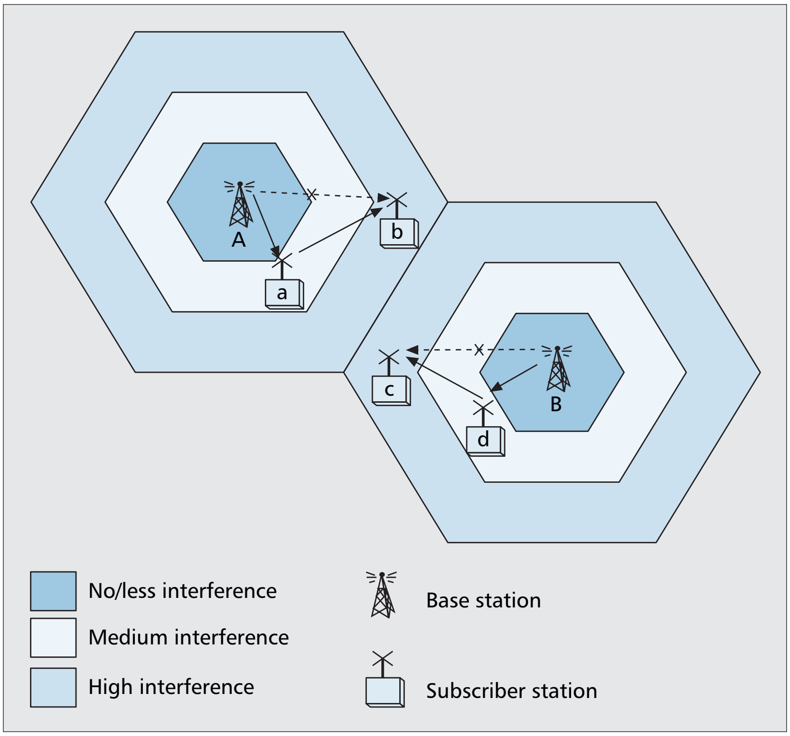


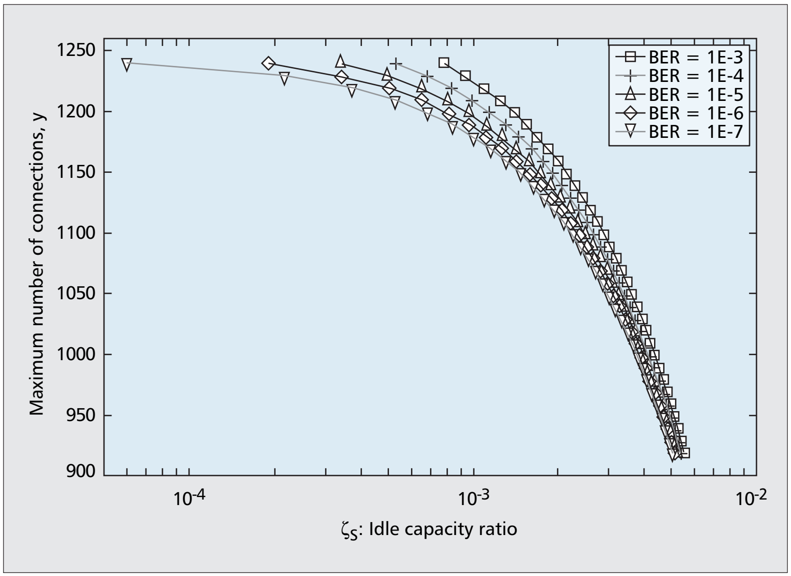
Figure 3. *The advantage of multihop routing in reducing interference to neighboring cells.*

CAPACITY PLANNING

In general, capacity planning is intended to dimension network resources based on long-term traffic demands to satisfy call-level QoS requirements, such as connection blocking probability and handoff dropping probability. There are several issues that must be considered in capacity planning of an OFDMA system: the effect of group mobility users on QoS, the effect of fluctuation in channel gains on QoS, and the like. In classical network dimensioning studies, handoff arrival processes have commonly been simplified as stationary or piecewise stationary processes (e.g., Markov modulated Poisson process, MMPP [11]). In addition, the MMPP model, which gives a more accurate representation of the nonstationary nature of handoff arrivals, has been shown to yield call blocking/dropping probabilities comparable to modeling handoff arrivals by a Poisson process. However, these models are not sufficient to characterize the nature of bulk arrivals, which are commonly observed in mass transportation systems in which handoffs from a group of active users may occur at the same time. Even though MMPP more closely reflects the nonstationary nature of an arrival process, by nature it is a counting process that cannot have more than one arrival in a given epoch [11].

In addition, in OFDMA systems it is also necessary to take into account the fluctuation of channel gains that affects the actual transmit rates users can experience. Because of the randomness of user mobility behavior, the average channel gain of a targeted group of users (in short, the average channel gain) in a wireless network changes over time, causing the average SNR of the user group to continuously fluctuate. Since the maximum achievable transmit rate is bounded by the SNR, ongoing connections may experience outage events; furthermore, the *outage probability* (defined later) increases with the number of connections admitted in the system. Therefore, it is necessary to take the fluctuating nature of SNR into account when performing CP. Several different optimization criteria have been used for CP, such as the average connection blocking probability, average delay, and utilization of bandwidth resources.

Several issues on CP in OFDMA are discussed in [12]. On one hand, evaluation of admission capacity without considering the degrading effect of group mobility users may produce results that are too optimistic. On the other hand, it is clear that CP based on group mobility analysis overestimates QoS. In general, dropping an ongoing call causes more unfavorable impacts on user satisfaction than blocking a new call. Although service providers would prefer to increase bandwidth utilization by optimistically admitting more calls, they are obliged to provide a minimum QoS to users, which means limiting the probabilities of call dropping and call blocking events. Various resource reservation or call queuing methods have been proposed to address these requirements. Therefore, CP must take into account these call-level QoS requirements as well as the mobility patterns of users, which affect the rate of handoff call arrivals. As group mobility could result in a large increase in handoff arrivals over a short time period, the possible adverse impacts of group



■ **Figure 4.** *Admission capacities as a function of bit error rates.*

mobility users must be properly taken into account in CP.

In this section we discuss one method of capacity planning for group mobility users presented in [12]. There are various types of CP formulations. The CP can be formulated as an optimization problem where the objective is to minimize the *outage probability* on ongoing connections subject to the constraint that the excess *idle capacity ratio* is not greater than a certain bound. The outage probability is defined as the average fraction of the total number of connections suffering from outages over a specific timeframe, whereas the idle capacity ratio is defined as the average fraction of the available capacity that is not utilized by user connections over the same time frame.

Let $P_O(y)$and $P_S(y)$ denote the outage probability and idle capacity ratio, respectively. Both of these are functions of nonnegative integer y which represents the number of connections. Let ζ_S be the upper bound on the idle capacity ratio. Thus, CP solves the following problem:

$$\begin{aligned}&\text{minimize } P_O(y)\\&\text{subject to } P_S(y) \le \zeta_S.\end{aligned} \tag{2}$$

Given that the outage probability is strictly increasing and the idle capacity ratio is strictly decreasing with respect to the number of active connections, the optimal solution of the above problem is given by

$$y_O^* = \lceil P_S^{-1}(\zeta_S) \rceil, \tag{3}$$

where $\lceil x \rceil$ is the smallest integer not less than x. Figure 4 shows the admission capacities vs. the desired idle capacity ratio for different levels of desired bit error rate (BER). The admission capacity decreases when BER decreases and when the targeted idle capacity ratio increases. The admission capacity increases approximately linearly with the decrease in desired idle capacity ratio (logarithm scale applied in the figure). It is observed that the differences between admission capacities at different values of BER decreases when the idle capacity ratio increases.

OPEN PROBLEMS

The CAC and CP problems in OFDMA mesh WMANs are attracting increasing attention. There are several open problems in these areas. Most notably, the impact on CAC and CP when MHR is used to support handoff connections by relaying through intermediate SSs has not been extensively studied. It has been found that MHR can increase the cell capacity, and the improved capacity can contribute to QoS improvements (e.g., reducing connection dropping and new connection blocking probabilities) at the expense of slight increases in end-to-end delays. Furthermore, the negative impact of bulk arrivals on CP can also be alleviated by applying MHR. It is also interesting to develop realistic traffic arrival models for these networks that take into account individual and group mobility of users as well as network- and user-centric QoS provisioning techniques.

CONCLUSION

This article has addressed four types of interrelated resource allocation problems in OFDMA WMANs: dynamic subcarrier allocation, adaptive power allocation, connection admission control, and capacity planning. Addressing all of these problems is crucial for the successful evolution of fixed and mobile WMANs. The dynamic subcarrier assignment problem is combinatorial in nature, and cannot be solved exactly for realistic problem instances, thus necessitating the development of efficient heuristics to generate near-optimal or suboptimal solutions in real time. Under certain conditions, dynamic subcarrier assignment can be formulated as a maximal bipartite matching problem, in which a polynomial-time algorithm, known as the Hungarian method, gives the optimal solution when modulation and coding levels are known. Although adaptive power allocation does not significantly improve the performance of cellular OFDMA systems, its application in mesh OFDMA WMANs can be potentially beneficial due to reduced interference at the cell edges and multiple available relays, and therefore merits further investigation. Admission control with priority to handoff connections can benefit from multihopping routing and needs to be further studied. For capacity planning, we have considered the group mobility scenario where handoffs follow bulk arrival processes. Capacity planning in a mesh network taking mobility into account remains a challenge.

ACKNOWLEDGMENTS

This work was supported by a University of British Columbia Graduate Fellowship, a grant from Bell Canada, and by the Canadian Natural Sciences and Engineering Research Council grant STPGP269872-03.

REFERENCES

[1] "IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems," IEEE, Tech. rep. 802.16, Oct. 2004.

[2] "IEEE Standard for Local and Metropolitan Area Networks Part 16 and Amendment 2," IEEE, Tech. rep. 802.16e, Feb. 2006.

[3] G. Song and Y. G. Li, "Cross-Layer Optimization for OFDM Wireless Networks — Part I: Theoretical Framework," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, Mar. 2005, pp. 614–24.

[4] M. Ergen, S. Coleri, and P. Varaiya, "QoS Aware Adaptive Resource Allocation Techniques for Fair Scheduling in OFDMA Based Broadband Wireless Access systems," *IEEE Trans. Broadcast.*, vol. 49, no. 4, Dec. 2003, pp. 362–70.

[5] K.-D. Lee and V. C. M. Leung, "Fair Allocation of Subcarrier and Power in an OFDMA Wireless Mesh Network," *IEEE JSAC*, vol. 24, no. 11, Nov. 2006, pp. 2051–60.

[6] J. Jang and K. B. Lee, "Transmit Power Adaptation for Multiuser OFDM Systems," *IEEE JSAC*, vol. 21, no. 2, Feb. 2003, pp. 171–78.

[7] Z. Zhang, Y. He, and E. K. P. Chong, "Opportunistic Downlink Scheduling for Multiuser OFDM Systems," *IEEE WCNC '05*, vol. 2, Mar. 2005, pp. 1206–12.

[8] H. Lim, C. Lim, and J. C. Hou, "A Coordinate-Based Approach for Exploiting Temporal-Spatial Diversity in Wireless Mesh Networks," *ACM Mobicom '06*, Sept. 2006.

[9] D. Niyato and E. Hossain, "Connection Admission Control Algorithms for OFDM Wireless Networks," *Proc. IEEE GLOBECOM '05*, 2005, pp. 2455–59.

[10] D. Niyato and E. Hossain, "A Radio Resource Management Framework for IEEE 802.16-Based OFDM/TDD Wireless Mesh Networks," *Proc. IEEE ICC '06*, June 2006.

[11] X. Lagrange and B. Jabbari, "Fairness in Wireless Microcellular Networks," *IEEE Trans. Vehic. Tech.*, vol. 47, no. 2, May 1998, pp. 472–79.

[12] K.-D. Lee and V. C. M. Leung, "Capacity Planning for Group-Mobility Users in OFDMA Wireless Networks," *EURASIP J. Wireless Commun. and Net.*, vol. 2006, 2006, pp. 1–12.

BIOGRAPHIES

SYED HUSSAIN ALI [S'05] (hussaina@ece.ubc.ca) received a B.E. degree in computer engineering from NED University of Engineering and Technology, Karachi, Pakistan in 1993, and an M.S. degree in computer engineering from KFUPM, Dhahran, Saudi Arabia, in 1998. From 1993 to 1996 he worked on several communication and networking projects in Pakistan. After his M.S., he was a network engineer for four and a half years at the College of Computer Sciences and Engineering, KFUPM. He is currently working toward his Ph.D. degree in the Department of Electrical and Computer Engineering, University of British Columbia (UBC), Vancouver, Canada. He is a recipient of a UBC graduate fellowship.

KI-DONG LEE [S'00, A'01, M'04] (kidongl@ece.ubc.ca) received a Ph.D. degree from KAIST, Korea, in 2001. After working four years at ETRI as a senior engineer, he joined UBC. His research interests include revolutionary radio resource management in present and evolutionary wireless networks. He received the IEEE ComSoc AP Outstanding Young Researcher Award and APORS Young Scholar Award, and was listed in Marquis *Who's Who in Science and Engineering*. He served as a co-guest editor for a special issue on next-generation hybrid wireless systems in *IEEE Wireless Communications*.

VICTOR C. M. LEUNG [F] (vleung@ece.ubc.ca) is a professor of electrical and computer engineering and the Telus Mobility Research Chair at UBC, where he pursues research on mobile and wireless networks. He was a lecturer in electronics in the Chinese University of Hong Kong in 1988 and a senior technical staff member at MPR Teltech Ltd. from 1981 to 1987. He received his B.A.Sc. and Ph.D. from UBC in 1977 and 1981, respectively. He is on the editorial boards of *IEEE Transactions on Wireless Communications* and *IEEE Transactions on Vehicular Technology*. He is a registered Professional Engineer in British Columbia, Canada.

For capacity planning, we have considered the group mobility scenario where handoffs follow bulk arrival processes. Capacity planning in a mesh network taking into account of mobility remains a challenge.