Optimizing Multicast Performance in Large-Scale WLANs

Ai Chen, Dongwook Lee, and Prasun Sinha
Department of Computer Science and Engineering, The Ohio State University
email: {chenai, leedon, prasun}@cse.ohio-state.edu

Abstract

Support for efficient multicasting in WLANs can enable new services such as streaming TV channels, radio channels, and visitor's information. With increasing deployments of large-scale WLANs, such services can have a significant impact. However, for a solution to be viable, the mutlicast services must minimally impact the existing unicast services which are currently the core services offered by most WLANs. This paper focuses on three objective functions motivated by different revenue functions and network scenarios: maximizing the number of users (MNU), balancing the load among APs (BLA), and minimizing the load of APs (MLA). We show that these problems are NP-hard and present centralized approximation algorithms and distributed approaches to solve them. Using simulations we evaluate the performance of these algorithms. We observe that the number of users can be increased by up to 36.9%*, and the maximum AP load and the total load can be reduced by up to* 52.9% *and* 31.1%*, respectively.*

1 Introduction

The goal of anytime-anywhere connectivity is becoming a reality with increasing deployment of large scale Wireless LANs. Networks deployed in industrial campuses, academic campuses, and cities are some scenarios that illustrate the scale at which WLANs are being deployed today. The city-wide network in Chaska, Minnesota is one such example that provides WLAN coverage in a 15 sq miles area since Oct 2004[1]. A similar network is operational in Taipei[2] that consists of 2300 APs and provides coverage to 50% of the city's population, and is planned to be extended to provide coverage to 90% of the city's population in the near future. While unicast services are essential for providing Internet access to individual users through WLANs, efficient multicast support from the network can be leveraged for distribution of live or stored multimedia content. Such services will enable distribution of multimedia rich content such as local news, visitor's information, and local TV channels.

While introducing media-rich multicast streaming in WLANs, it is critical to ensure that the multicast services use the resources efficiently and the unicast services get minimally effected. However, the 802.11 standard can not efficiently maximize resource usage, since uncontrolled association causes multiple APs with overlapping regions to transmit the same multicast packets, thereby wasting resources for unicast services. In this paper, we study how to provide efficient multimedia multicast service to users through controlling the user-to-AP association in WLANs. Improving the efficiency of multicast services make it feasible to introduce multicasting support while minimally impacting unicast users.

Although association control has already been considered by both the research community and the industry, previous research on association control in WLANs primarily focused on unicast traffic [11, 8, 26, 2, 14]. The problem of unbalanced AP load under signal strength based association is discussed in [11]. In [8, 26], new metrics are studied to associate with APs instead of signal strength for only unicast traffic. These works do not consider load-balancing between APs. Recent work [2, 14] has explored the idea of association control to balance the network load and provide max-min fairness among users. However, they do not consider load balancing problems for multicast traffic. *To the best of our knowledge, we are the first to study the association control for enabling efficient multicast streaming sevices in WLANs.*

We study three different objectives supported by different revenue functions that can be used by the WLAN service provider depending on the expected network scenario. We believe these three objectives are most interesting for the WLAN service provider and users.

- **Maximize Number of Users (MNU):** Under high load scenarios, it may not be feasible to satisfy all the users' multicast requests. For such scenarios, it may be critical to maximize the number of satisfied multicast flows. If the service provider charges customers

[1] www.chaska.net
[2] http://english.taipei.gov.tw

based on the duration of the multicast flows, then this objective function will be of interest.

- **Balance Load among APs (BLA):** In case the users' requests can be all met, it is critical to balance the multicast load so that the available fraction of time for unicast is also balanced. More precisely the objective here is to minimize the maximum multicast duration of an AP. This will lead to fairer share of the unicast bandwidth in case of uniform distribution of users across APs. If the revenue function gives higher weightage to unicast traffic and the revenue function for unicast traffic is convex, then BLA will likely lead to improved revenue. It is known that convex functions are maximized when the resources are uniformly distributed [12].

- **Minimize Load of APs (MLA):** The objective is to reduce the summation of multicast load across all the APs in the WLAN. This will maximize the total available time for unicasting. Under revenue models where there is a flat rate per byte of unicast data, and a scenario with a sufficient number of users requesting unicast traffic, MLA may be the desired objective function for the service provider.

We make the following contributions in this paper. First, we show the NP-hardness of the three problems even if we restrict the problem so that multicast/broadcast packets are always transmitted at the basic rate. Second, we reduce the three problems to other known problems and present centralized approximation algorithms. For MNU, MLA and BLA, we present approximation algorithms with approximation factors of 8, $\log_{\frac{8}{7}}(n) + 1$ and $ln(n)$ respectively. Third, we present distributed approaches to solve the problems, although we believe that in smaller WLANs (of the order of 100 APs) centralized algorithms are still feasible to execute. Note that distributed solutions are preferred in large networks, as centralized solutions will lead to more frequent changes in associations causing increased signaling traffic over the wireless links. Fourth, through simulations we study the performance of the proposed distributed and centralized solutions for the three objectives. We observe that the number of users can be increased by up to 36.9%, and the maximum AP load and the total load can be reduced by up to 52.9% and 31.1%, respectively. We also evaluate the optimality of MLA, BLA, and MNU algorithms and find our centralized and distributed algorithms are very efficient compared with the optimal solutions.

The rest of the paper is organized as follows. Section 2 summarizes relevant related work. Section 3 defines the network model, the problems, the notations, and the terminology used in the paper. The centralized and distributed algorithms for MNU, BLA, and MLA are described in Sections 4, 5, and 6 respectively. Section 7 presents a detailed evaluation of our approach and comparison with signal strength based approach using simulations. The future work is described in Section 8. Finally, Section 9 concludes the paper.

2 Related Work

In this section, we outline related works in the areas of MAC layer multicast/broadcast and controlled association in wireless networks.

MAC Layer Multicast Protocols: IEEE 802.11 MAC protocol implements multicast using broadcast. As the 802.11 broadcast is unreliable, several protocols [15, 22, 23, 24, 21, 6, 20, 10] have been proposed to improve reliability. Kuri and Kasera [15] proposed a reliable multicast protocol for WLANs. Tang and Gerla [22, 23] extended the broadcast mechanism of 802.11 that tries to confirm that at least one receiver receives the broadcast packet in ad hoc networks. In [24], Tang and Gerla proposed BMW (Broadcast Medium Window) protocol which implements broadcast based on unicast and lets receivers overhear packets. In [21], Sun *et al.* proposed BMMM (Batch Mode Multicast MAC) protocol to implement reliable MAC layer multicast. Some MAC layer multicast/braodcast protocols, such as BPBT [6], RMAC [20], and 80211MX [10], use busy-tone to implement multicast reliability, while Chaporkar *et al.* [3, 4] proposed algorithms for maximizing throughput for MAC layer wireless multicast using busy tones. Our association control algorithms are independent of the MAC layer protocol, and the efficiency of the MAC layer protocol can increase the efficiency of our algorithms.

Association Control: In 802.11 networks user nodes often use signal strength as the key metric in selecting the AP. The problem of unbalanced AP load under signal strength based association was discussed in [11]. In [8, 26], new metrics were studied to select a unicast AP instead of signal strength. Packet error rate and number of users were used in [8]. The authors showed deployability and robustness of their AP selection architecture. Association time, system load, and signal/noise ratio (SNR) together were used to initiate handoff in [26]. The authors argued that their approach can provide Quality of Service (QoS) guarantee. However, these works did not consider load-balancing between APs.

Recent work [2, 14] has explored the idea of association control to balance the network load and provided max-min fairness among users. The authors in [2] proved that balancing the network load is equivalent to achieving the max-min fairness, and presented algorithms that achieve a constant-factor approximation to max-min fair bandwidth allocation. In [14], analytical model was formulated for the AP selection as an optimization problem to maximize different utility functions. The authors provided the optimal association results for some simple cases. However, these works primarily focused on unicast traffic.

For broadcast service in wireless Mesh networks, the optimal association algorithm was studied in [16], where minimum cost based greedy selection of an access point can decrease the size of the broadcast tree. The authors proposed the concept of multi-association, where the AP for unicast traffic and the AP for broadcast traffic are independently chosen by exploiting multiple coverages that are typical in mesh networks.

3 Preliminaries

In this section, we present the network model and outline the problems that are addressed in this paper.

3.1 Network Model

We consider a WLAN with a set of users U and a set of access points A. In the 802.11 standard the data rate of the link is chosen dynamically based on the signal quality to keep the BER (bit error rate) below a fixed threshold. For example 802.11 g/a supports discrete levels of data rates ranging from 6 Mbps to 54 Mbps. The maximum possible data rate on a link from an AP a to a user u is denoted by $r_{a,u}$. If an AP multicasts packets to the users associated with it, the AP will use the lowest rate among these users' maximum possible data rates on the links to this AP, which makes sure that every user can receive the packets. We assume that MAC layer multicast/broadcast can support multi-rate transmission[3] although in the 802.11 standard MAC layer broadcast packets are always transmitted at the basic data rate. However, if the basic data rate is always used for multicast/broadcast, MNU, BLA and MLA problems are still NP-hard because our NP-hardness proofs for these problems do not require multi-rate transmission, and our algorithms still get better performance than the strongest signal approach.

Each user node and each AP has a single radio. We assume that the radio channels of the neighboring APs are configured such that they do not interfere. Although IEEE 802.11b/g only has 3 non-overlapping channels, the newer IEEE 802.11a standard operates in the 5 GHz spectrum that supports 12 non-overlapping channels in US/Canada. Although interference modeling is difficult to obtain in a real network (See discussions in Section 8), our solutions implicitly optimize interference as discussed in the next subsection (Section 3.B). The APs are connected using a wired LAN to one or more gateways that provide connectivity to the Internet. We assume that users are *quasi-static*, which means that they often tend to stay at one place for a relatively long time period before changing their location. This assumption is supported by recent studies of user mobility patterns in deployed WLANs [1, 13]. Each user may request one multicast stream from the WLAN. This is similar to our TV services where a user typically watches only one TV channel at any time. Another such example is the video watching tool at CNN.COM which normally allows only one video streaming session per client at any time. A user requesting a multicast stream is referred to as a multicast user, and a user requesting unicast service is referred to as a unicast user. If a user can only be a unicast user or a multicast user, we do not need to do any other modification to the 802.11 standard except for that the association algorithm of 802.11 is replaced by our association algorithm for multicast users. If a user can be both a unicast as well as a multicast user, the network framework discribed in [16] can be applied, where the APs are synchronized through a time-synchronization protocol and each user independently selects one AP for unicast and another one for multicast services.

[3]Recent research [7] has implemented the multi-rate broadcast/multicast.

3.2 Problem Statement

This paper focuses on algorithms for selecting the optimal AP for receiving multicast flows. We study three different objective functions and propose approximation algorithms for solving them. The suitability of the objectives will depend on the users demands, users distribution, and the network providers revenue function. We first define the term *multicast load* that is used by some of the objective functions.

Definition 1 ***Multicast Load**: the multicast load of an AP is the fraction of time that the AP is busy in transmitting multicast flows; the total multicast load of a network is the sum of all APs' multicast load in the network.*

Maximize Number of Users (MNU): When there is a heavy demand for multicast flows, all the user's requests can not be met. For such scenarios, we define the goal to be maximization of the number of users that get multicast service from the network. Although the network revenue can be a function of both unicast and multicast flows, under typical network revenue models, the number of satisfied users will result in higher revenue.

An example revenue model is when the unicast services have a monthly charge, but the multicast services are charged based on the time for which multicast streams are served to users. This is like the Pay-per-view service offered by most cable and satellite TV services today. Under such a model, increasing the number of satisfied multicast users will increase the total revenue for the service provider.

Balance Load among APs (BLA): In order to reduce the impact of multicast services on unicast flows, it is critical to reduce the size of the multicast period. This can be achieved by balancing the multicast load of the APs. More precisely, the objective here is to minimize the maximum multicast load among all APs.

Consider a revenue model where one multicast flow is included in the basic monthly charges. Assume that the revenue function for unicast flows is convex, i.e., marginally decreasing with increasing bandwidth. Convex revenue functions are well known for achieving fairness among flows. Then balancing the multicast load will typically lead to fairness among the unicast flows and a higher total revenue, assuming uniform distribution of unicast users across the APs.

Minimize Load of APs (MLA): In order to free up the maximum amount of total time for unicast services, the total multicast load needs to be minimized. Although this can lead to uneven distribution of multicast load, for some revenue models this may be of interest.

Consider a revenue model where one multicast flow is included in the basic monthly charges. However, the unicast services are charged per byte. In scenarios where there is a high demand for unicast traffic, maximizing the total amount of unicast traffic will maximize the revenue while satisfying the multicast users.

Note that solutions to the BLA and MLA problems will implicitly optimize maximum interference from an AP and total interference from all the APs, resulting due to multicast transmissions.

An Example: We use the example scenario shown in Figure 1 to describe these three problems. The WLAN consists of two APs, a_1, a_2, and, 5 users, $u_1, u_2, \ldots, u_5$. Suppose that the maximum load that APs a_1 and a_2 can support for multicast traffic is 1 unit. The maximum data rates to the five users u_1, u_2, u_3, u_4 and u_5 from a_1 are $3, 6, 4, 4,$ and 4 respectively. AP a_2 can communicate only with users $u_3, u_4,$ and u_5 and the maximum data rates are $5, 5,$ and 3 $Mbps$, respectively. Suppose that users u_1 and u_3 request multicast sessions s_1, and users $u_2, u_4,$ and u_5 request for multicast session s_2.

If the multicast data rates of s_1 and s_2 are both 3 $Mbps$, this WLAN can not support all the users for multicast because u_1 and u_2 can only be associated with a_1, and a_1 can not provide multicast service to u_1 and u_2 simultaneously. If both u_1 and u_2 are supported by a_1 then the total load on a_1 will be $\frac{3}{3} + \frac{3}{6} > 1$, which is infeasible. In such scenarios, the objective of maximization of number of users (**MNU**) is relevant. One of the optimal solutions is that $u_2, u_4,$ and u_5 are associated with a_1 and u_3 is associated with a_2. This results in a load of $\frac{3}{4}$ at AP a_1 and a load of $\frac{3}{5}$ at AP a_2.

Suppose the data rate of s_1 and s_2 are both 1 $Mbps$ and the objective is to balance the multicast load among APs (**BLA**) by minimizing the maximum multicast load among the APs. In the optimal solution $u_1, u_2,$ and u_3 are associated with a_1, and u_4 and u_5 are associated with a_2. The load of a_1 will thus be $\frac{1}{3} + \frac{1}{6} = \frac{1}{2}$ and the load of a_2 will be $\frac{1}{3}$.

Suppose the data rate of s_1 and s_2 is 1 $Mbps$ and the objective is to minimize the total load of all APs for multicast streams (**MLA**). In the optimal solution all users are associated with a_1, which results in a total AP load of $\frac{1}{3} + \frac{1}{4} = \frac{7}{12}$.

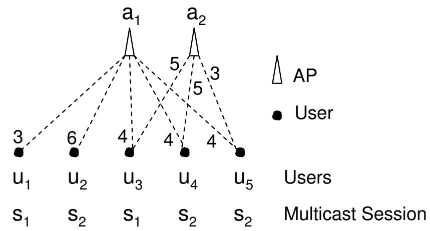


Figure 1. An Example Network Scenario: Users u_1 and u_3 request for multicast session s_1, and users u_2, u_4, u_5 request for multicast session s_2.

4 Maximize Number of Users (MNU)

In this section, we show that MNU is an NP-hard problem and present centralized and distributed algorithms. Due to resource limitation for multicasting at the APs, the network may not be able to support the request from all users. Therefore, the network should try to maximize the number of users that it supports.

We show that MNU is an NP-hard problem, by showing a reduction from the Subset Sum problem, which is described in the Appendix. Because the Subset Sum problem is NP-hard, MNU problem is also NP-hard. Note that MNU is trivially in P, if there is only one multicast session in a WLAN. For a single session, all APs can choose to transmit at the lowest rate that does not violate the maximum multicast period.

4.1 Centralized MNU

In order to solve this problem, we present a reduction from the MNU problem to the Maximum Coverage with Group Budgets (MCG)[5] problem. Here, we give the definition of the cost version.

Definition 2 ***Maximum Coverage with Group Budgets (MCG) - cost version:*** *There are m subsets $S_1, S_2, \ldots, S_m$ of a ground set X. There are l sets $G_1, G_2, \ldots, G_l$, each G_i being a subset of $\{S_1, \ldots, S_m\}$. Each G_i is a group and the groups are disjoint from each other*[4]*. A cost $c(S_j)$ is associated with each set S_j. Further, each group G_i is given*

[4]If they are not disjoint from each other, we can make them disjoint by making copies of sets in $S_1, \ldots, S_m$.

a budget B_i and the overall budget is B. The objective is to find a subset H of $\{S_1, \dots, S_m\}$ to maximize the size of the union of sets in H under the limitation that the total cost of the sets in H is at most B, and for any group G_i, the total cost of the sets in $H \bigcap G_i$ is at most B_i.

Theorem 1 *MNU can be reduced to MCG problem.*

Proof: The set of all users becomes the set X in the instance of MCG. Corresponding to each AP, we create $m \times |S|$ subsets of X in MCG, where m is the number of discrete transmission rates that the WLAN supports and S is the set of multicast sessions. Thus, each subset corresponds to an AP, a single transmission rate, and a multicast session. The cost of a subset is the ratio of the corresponding multicast session's data rate and the transmission rate. All such subsets that are related to AP a_i form the group G_i. The budget B_i for the group G_i is the fraction of the time AP a_i spends on multicast transmissions. For our problem, there is no overall budget limitation for the whole network, i.e., $B = \infty$, as we assume that the capacity of the wired network is not the bottleneck. □

Example – MNU: If the data rate of s_1 and s_2 is 3 $Mbps$ in the WLAN shown in Figure 1, we can reduce the MNU problem for the WLAN to the MCG problem shown in Figure 2. One of the optimal solutions for this MCG problem is $H = \{S_4, S_5\}$.

$X = \{u_1, u_2, u_3, u_4, u_5\}$

$S_1 = \{u_3\}$	$c(S_1) = 3/4$
$S_2 = \{u_1, u_3\}$	$c(S_2) = 3/3 = 1$
$S_3 = \{u_2\}$	$c(S_3) = 3/6 = 1/2$
$S_4 = \{u_2, u_4, u_5\}$	$c(S_4) = 3/4$
$G_1 = \{S_1, S_2, S_3, S_4\}$	$B_1 = 1$
$S_5 = \{u_3\}$	$c(S_5) = 3/5$
$S_6 = \{u_4\}$	$c(S_6) = 3/5$
$S_7 = \{u_4, u_5\}$	$c(S_7) = 3/3 = 1$
$G_2 = \{S_5, S_6, S_7\}$	$B_2 = 1$

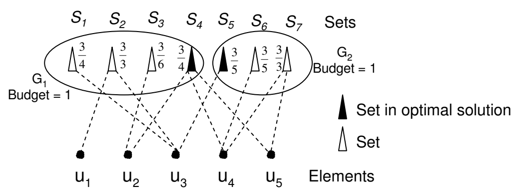


Figure 2. The reduction from MNU problem for the WLAN in Figure 1 to MCG problem. The data rate of s_1 and s_2 is 3 $Mbps$. One of the optimal solutions for this MCG problem is $H = \{S_4, S_5\}$.

We use the above reduction to reduce any arbitrary instance of MNU to an instance of the MCG problem. In [5], the authors presented a greedy algorithm for MCG as it is an NP-hard problem. Because there is no overall budget limitation for our problem, we adapt the algorithm in [5] and present the modified algorithm below. The boundary of the algorithm is also different from the one in [5].

Algorithm **Centralized MNU**

1. $H \leftarrow \phi, X' \leftarrow X$.
2. **repeat**
3. $\quad flag \leftarrow 0$
4. $\quad$ **for** $i = 1, 2, \dots, n$ **do**
5. $\quad\quad$ **if** $c(H \bigcap G_i) < B_i$ **then**
6. $\quad\quad\quad k \leftarrow argmax_j \frac{|S_j \bigcap X'|}{c(S_j)}$ $(S_j \in G_i)$
7. $\quad\quad\quad A_i \leftarrow S_k$
8. $\quad\quad\quad flag \leftarrow 1$
9. $\quad\quad$ **else** $A_i \leftarrow \phi$
10. $\quad$ **endfor**
11. $\quad$ **if** $flag = 0$ **then Break**
12. $\quad r \leftarrow argmax_i \frac{|A_i \bigcap X'|}{c(A_i)}$
13. $\quad H \leftarrow H \bigcup A_r,\ X' \leftarrow X' - A_r$
14. $\quad$ **if** $X' = \phi$ then **break**
15. **endrepeat**
16. **output** H

Figure 3. Centralized Solution for MNU

The algorithm greedily picks up subsets with minimum cost for every additional element until either all elements have been covered or until each group's budget has been violated by the last selected subset for the group. In the pseudo-code presented in Figure 3, H represents the set of selected subsets at any step. The set X' denotes the elements of X which have not yet been cover by the subsets in H. The statements from line 3 to line 14 are repeatedly executed until all the group budgets are exceeded or all elements of X get covered. The variable *flag* is used for this purpose in line 11. In the *for* loop, *Centralized MNU* finds a set S_j in every group G_i whose budget has not been exceeded, and S_j is the set which is the most cost-effective set in the group G_i, i.e., $\frac{|S_j \bigcap X'|}{c(S_j)} = \max_{D \in G_i} \frac{|D \bigcap X'|}{c(D)}$. Then in line 12, *Centralized MNU* finds the most cost-effective set in the sets selected in the *for* loop. This set is added into H and the elements in this set is removed from X' in line 13. Eventually, we get the output H.

Obviously, H does not obey the group budget requirements. We assume the cost of any single set S_j in any group G_i is not more than the budget of G_i. We partition H into two subsets H_1 and H_2. H_2 contains those sets S_j which when added to H caused the budget of some group G_i to be violated. $H_1 = H - H_2$. Observe that H_1 and H_2 by themselves do not violate the budget constraints and one of these two sets must be covering at least $1/2$ the number of elements covered by H. Out of H_1 and H_2, we select the one which covers the most number of elements. The final solution directly maps to the solution to the MNU problem. Then, we have the following theorem.

Theorem 2 *The algorithm Centralized MNU is an 8-approximation algorithm for MNU problem with no total budget limitation.*

Proof: Define $X(H)$ as the number of the elements covered by the subsets of X in H. Let OPT be some fixed optimal solution to the given problem instance. In [5], it was proved that $X(H) \geq \frac{1}{4}X(OPT)$. As either H_1 or H_2 must contain at least half the elements covered by H, *Centralized MNU* is an 8-approximation algorithm for MNU problem. □

Example – Centralized MNU: We run *Centralized MNU* algorithm on the MCG problem shown in Figure 2. S_4 is selected in the first round because it has the maximum value of $\frac{|S_4 \bigcap X'|}{c(S_4)} = \frac{3}{3/4} = 4$ among all S_i $(1 \leq i \leq 7)$. After that, $H = \{S_4\}$, $X' = \{u_1, u_3\}$. In the second round, S_2 is selected because it has the maximum value of $\frac{|S_2 \bigcap X'|}{c(S_2)} = \frac{2}{1} = 2$ and $c(H \bigcap G_1) = c(S_4) = 3/4 < B_1 = 1$. After that, we get output $H = \{S_2, S_4\}$ because $X' = \phi$. Now, $c(H \bigcap G_1) = c(S_2) + c(S-4) = 7/4 > B_1 = 1$. We divide H into $H_1 = \{S_4\}$ and $H_2 = \{S_2\}$. Eventually, we get output H_1 because H_1 cover more elements than H_2. Therefore, u_2, u_4, u_5 are associated with a_1 and 3 users get multicast streams. If we use strongest signal based approach, u_1, u_2, u_5 can only be associated with a_1 and u_3, u_4 can only be associated with a_2. If u_1, u_3 are associated with APs first, u_2, u_4, u_5 can not be associated with APs because of the load limitation of APs. So, only 2 users get multicast service.

4.2 Distributed MNU

We provide a simple distributed algorithm to maximize the number of users. Intuitively, because the total resource of the network (APs) for multicast is fixed, every user should increase the total load minimally in order to attempt increasing the total number of users. Due to lack of global view, the distributed approach has to take decisions based only on local information obtained from the APs.

A user periodically sends a query message to each of its neighboring APs. Then, each AP responds with a message containing information about the current multicast sessions being transmitted and the data rate of such transmissions. The user also knows the maximum data rate for the link between itself and its neighboring APs. If a user is currently associated with some AP a, this user also needs to know the load of a if it leaves AP a. According to the information from the neighboring APs, the user calculates the total load of its neighboring APs if it can associate with it without violating the maximum multicast load for that AP. The user then associates with the neighboring AP that results in minimum increase in total load. If there are several APs that result in the same minimum increase in total load, the user can associate with the one with the strongest signal.

Example – Distributed MNU: Consider that the data rate of s_1 and s_2 is 3 $Mbps$ in the WLAN in Figure 1, and users use the distributed algorithm in the order u_1, u_2, u_3, u_4, u_5. First u_1 associates with a_1. Then, u_2 can not associate with a_1 because of the load limitation of a_1. After that, u_3 associates with a_1, which results in the minimum total load 1 of u_3's neighboring APs a_1, a_2. Similarly, u_4, u_5 are associated with a_2. Eventually, 4 out of the 5 users receive their multicast service.

Lemma 1 *The algorithm Distributed MNU converges when the network becomes static if the users in an AP's transmission range make their local decisions one by one.*

Proof: Because the users in an AP's transmission range make decision one by one, each user always operates on the most up-to-date information about the multicast sessions. First, we consider the scenario where there are no new users joining the network. If a user has been associated with an AP and wants to change its association, it should reduce the total load of all of its neighboring APs, which also means the total load of the whole network will be reduced. As the number of discrete levels of data rates, number of APs, and number of users are limited, the total load of the whole network should be eventually reduced to a final value in finite steps. If a new user joins the network, the total load of the whole network also should reach a final value in limited steps. The network is static, and the total number of users is finite. The number of new users joining the network is also finite. Therefore, the distributed algorithm converges for a static network if the users make decision one by one. □

However, if the users in an AP's transmission range make their local decisions simultaneously, the algorithm *Distributed MNU* may not converge. The example scenario is shown in Figure 4. AP a_1 can communicate with u_1, u_2 and u_3 with the rates $5, 4$ and 4 $Mbps$, respectively; AP a_2 can communicate with u_2, u_3 and u_5 with the rates $4, 4$ and 5 $Mbps$, respectively. Users u_1 and u_2 are associated with a_1, and u_3 and u_4 are associated with a_2. All users request the same multicast session s_1 with the rate 1 $Mbps$. So, the current total load of a_1 and a_2 is $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. Now, u_2 and u_3 make the local decision simultaneously. If only u_2 changes its association and it associates with a_2, the total load of a_1 and a_2 is reduced to $\frac{1}{5} + \frac{1}{4} = \frac{9}{20}$. If only u_3 changes its association and it associates with a_1, the total load of a_1 and a_2 is reduced to $\frac{1}{4} + \frac{1}{5} = \frac{9}{20}$. Therefore, both u_2 and u_3 will change their associations, which actually does not change the total load. Next, if u_2 and u_3 make local decisions simultaneously again, u_2 and u_3 will be associated with a_1 and a_2 respectively, again. Therefore, the algorithm does not converge.

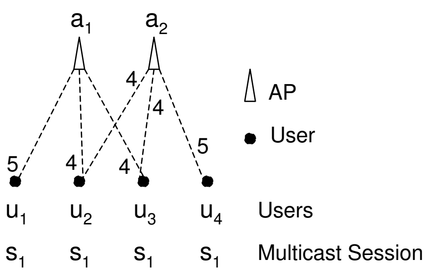


Figure 4. Negative Example of Converge for Simultaneous Local Decisions: All users request the same multicast session. Users u_1 and u_2 are associated with a_1, and u_3 and u_4 are associated with a_2. u_2 and u_3 always make their local decisions simultaneously.

5 Balance Load among APs (BLA)

In this section, we prove the NP-hardness of the BLA problem and present centralized and distributed algorithms. The objective is to minimize the maximum load among the APs.

We present a reduction from Minimum Makespan Scheduling problem [25] to the BLA problem, which is described in the Appendix, to prove the NP-hardness of BLA. Based on the reduction, BLA is NP-hard because the minimum makespan scheduling problem is NP-hard. Note that BLA is a P problem if there is only one multicast session. As there are constant number of discrete transmission rates, each of these transmission rates can be checked in sequence for feasiblity of being the maximum transmission rate. For a given value of the transmission rate, all APs are assigned the same rate (as the optimization function only concerns the maximum). Among all the transmission rates the highest rate (when assigned to all APs) that provides service to all users, is the solution.

5.1 Centralized BLA

In order to solve the BLA problem, we present a reduction to the Set Cover with Group Budgets (SCG) [5] problem. We first give the definition of the cost version of SCG.

Definition 3 ***Set Cover with Group Budgets (SCG):*** *There is a set $S = \{S_1, S_2, \ldots, S_m\}$ of subsets of a ground set X. The set S is partitioned into groups $G_1, G_2, \ldots, G_l$. A cost $c(S_j)$ is associated with each set S_j. The objective is to find a subset H of S such that all elements of X are covered by sets in H and $max_{i=1}^{l} c(H \bigcap G_i)$ is minimized.*

Theorem 3 *The problem of balancing load among APs (BLA) can be reduced to SCG problem.*

Proof: The reduction from BLA to SCG is similar to the reduction from MNU to MCG in section 4. We denote the set of all users as X. For each AP, we create multiple subsets, each corresponding to a particular combination of session number and transmission rate. The cost of each subset is obtained by dividing the rate of the corresponding session by the transmission rate associated with that subset. All of the subsets that are related to AP a_i form the group G_i. □

Example – BLA: If the data rate of s_1 and s_2 is 1 $Mbps$ in the WLAN in Figure 1, we can reduce the BLA problem for the WLAN to the SCG problem shown in Figure 5. The optimal solution of this SCG problem is $H = \{S_2, S_3, S_7\}$.

$X = \{u_1, u_2, u_3, u_4, u_5\}$
$S_1 = \{u_3\}$ $c(S_1) = 1/4$
$S_2 = \{u_1, u_3\}$ $c(S_2) = 1/3$
$S_3 = \{u_2\}$ $c(S_3) = 1/6$
$S_4 = \{u_2, u_4, u_5\}$ $c(S_4) = 1/4$
$G_1 = \{S_1, S_2, S_3, S_4\}$
$S_5 = \{u_3\}$ $c(S_5) = 1/5$
$S_6 = \{u_4\}$ $c(S_6) = 1/5$
$S_7 = \{u_4, u_5\}$ $c(S_7) = 1/3$
$G_2 = \{S_5, S_6, S_7\}$

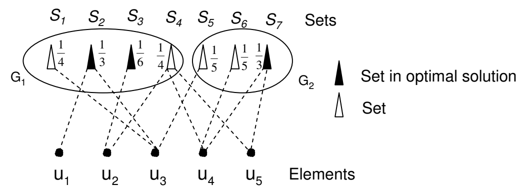


Figure 5. The reduction from BLA problem for the WLAN in Figure 1 to SCG problem. The data rates of s_1 and s_2 are both 1 $Mbps$. The optimal solution of this SCG problem is $H = \{S_2, S_3, S_7\}$.

SCG problem is also NP-hard. In [5], the authors gave an algorithm for the cardinality version of SCG based on the greedy algorithm for MCG. Our algorithm is similar. The algorithm is shown in Figure 6. Assume the number of the elements in the ground set X is n.

Theorem 4 *The algorithm Centralized BLA is an $(log_{8/7}n + 1)$-approximation algorithm for BLA problem.*

Proof: In each iteration of running *Centralized MNU*, the total cost of the sets added from any group G_i is bounded by B^*. Because we iterate *Centralized MNU* $log_{8/7}n + 1$ times in algorithm *Centralized BLA*, the total cost of the sets

Algorithm **Centralized BLA**

1. Guess the optimal value B^* and assume there is an optimal cover H^* such that $max_{i=1}^{l} c(H^* \bigcap G_i) \leq B^*$
2. Create an instance of MCG by having a budget of B^* on each group G_i. Run *Centralized MNU*, which covers at least $\frac{1}{8}$th the elements in X. Remove the covered elements from X and all subsets of X, and run *Centralized MNU* again. Iterating *Centralized MNU* $log_{8/7}n + 1$ times results in a solution that covers all elements. Let $H = \{all\ sets\ S_i\ added\ when\ repeating\ Centralized\ MNU\}$.
3. Output H.

Figure 6. Algorithm Centralized BLA

added from any group G_i is bounded by $(log_{8/7}n + 1)B^*$ when all elements in X are covered. Therefore, *Centralized BLA* is an $(log_{8/7}n+1)$-approximation algorithm for BLA. □

To implement the algorithm *Centralized BLA*, there is an issue of how to guess B^*. Let the maximum cost among all subsets of X in all groups be c_{max}. B^* also should be less than 1. Therefore, we can try several (a constant number) values of B^* between c_{max} and 1 to get the best result.

Example – Centralized BLA: We run *Centralized BLA* algorithm on the SCG problem shown in Figure 5. Let $B* = 1/2$ and create an instance of MCG problem. Then run *CentralizedMNU*, and get the output $\{S_4\}$. After that, remove u_2, u_4, u_5 from every S_i $(1 \leq i \leq 7)$ and create a new instance of MCG problem. Run *CentralizedMNU* again, and get output $\{S_2\}$. Therefore, all users are associated with a_1.

5.2 Distributed BLA

As the objective is to balance the load among APs, a user should attempt to minimize the maximum load of the neighboring APs. The following is the distributed algorithm for BLA.

A user periodically sends a query message to each of its neighboring APs. Then, each AP responds with a message containing information about the multicast sessions that this AP supports and the rates for the supported multicast sessions. The user also knows the maximum data rate for the link between itself and its neighboring APs. If a user is currently associated with some AP a, this user also needs to know the load of a if it leaves AP a. According to the information from the neighboring APs, the user calculates the new load of a neighboring AP if it is associated with this AP. For each AP it computes the new vector of loads of neighboring APs if it decides to join that AP. Each load vector is sorted in non-increasing order of the loads of APs in that vector. The user then determines to receive the desired flow from the AP that locally minimizes the sorted new load vector.[5]

Example – Distributed BLA: Assume that the data rates of s_1 and s_2 are both 1 $Mbps$ in the WLAN in Figure 1, and users run the distributed algorithm in the order u_1, u_2, u_3, u_4, u_5. First u_1, u_2 is associated with a_1. After that, u_3 makes the decision. If u_3 is associated with a_1, it's neighboring APs' load vector in non-increasing order is $(1/2, 0)$; if u_3 is associated with a_2, the load vector is $(1/2, 1/5)$. Therefore, u_3 is associated with a_1. Next, if u_4 is associated with a_1, its neighboring APs' load vector with non-increasing order is $(7/12, 0)$; if u_4 is associated with a_2, the load vector is $(1/2, 1/5)$. Hence, u_4 is associated with a_2. Similarly, u_5 is associated with a_2. Eventually, the load of a_1 is $1/2$ and the load of a_2 is $1/3$, which is also the optimal solution.

Lemma 2 *The distributed algorithm for BLA converges when the network is static if the users in an AP's transmission range make decision one by one.*

Proof: The proof is similar to the proof of Lemma 1. If a user has been associated with an AP and wants to change its association, it should reduce the vector of neighboring APs' loads, which also means the global vector of all APs' loads in the network is reduced. Because the number of different data rates, the number of APs, and the number of users are all finite, the vector of all APs' loads in the network will eventually settle down to a final value in limited number of steps. If there is a new user who joins the network, the sequence of all APs' loads in the network also should reach a final value in limited steps. Therefore, the distributed algorithm converges when the network is static if the users in an AP's transmission range make decision one by one. □

However, if the users in an AP's transmission range make their local decisions simultaneously, the distributed algorithm for BLA may not converge. The example scenario is same as the scenario for the distributed algorithm for MNU shown in Figure 4.

6 Minimize the Load of APs (MLA)

In this section, we prove the NP-hardness of MLA and describe our centralized and distributed algorithms. The objective of MLA is to reduce the total network load.

We show that MLA is an NP-hard problem, by showing a reduction from the Set Cover problem, which is described in the Appendix. MLA is NP-hard as set cover problem is NP-hard [9].

[5] We define two sequences with non-increasing order to be equal if each pair of values at the same position of these two sequences are equal. If two sequences are not equal, we compare the first pair of unequal elements at the same position and the sequence with the smaller element is smaller than the other sequence.

6.1 Centralized MLA

In order to solve the MLA problem, we reduce it to the Set Cover problem.

Theorem 5 *The problem of minimizing the load of APs (MLA) can be reduced to set cover problem.*

Proof: We regard the set of all of the users as the ground set X. The construction is same as the construction in the proof of Theorem 3 except that there are no groups since we are only concerned the total multicast load of a network, not each AP's load for MLA problem."

□

Example – MLA: If the data rate of s_1 and s_2 are 1 $Mbps$ in the WLAN in Figure 1, we can reduce the MLA problem for the WLAN to the set cover problem shown in Figure 7. The optimal solution of this set cover problem is $H = \{S_2, S_4\}$.

$X = \{u_1, u_2, u_3, u_4, u_5\}$
$S_1 = \{u_3\}$ $c(S_1) = 1/4$
$S_2 = \{u_1, u_3\}$ $c(S_2) = 1/3$
$S_3 = \{u_2\}$ $c(S_3) = 1/6$
$S_4 = \{u_2, u_4, u_5\}$ $c(S_4) = 1/4$
$S_5 = \{u_3\}$ $c(S_5) = 1/5$
$S_6 = \{u_4\}$ $c(S_6) = 1/5$
$S_7 = \{u_4, u_5\}$ $c(S_7) = 1/3$

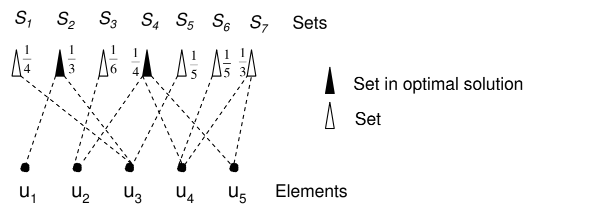


Figure 7. The reduction from MLA problem for the WLAN in Figure 1 to the set cover problem. The data rate of s_1 and s_2 is $1\ Mbps$. The optimal solution of this set cover problem is $H = \{S_2, S_4\}$.

We use the greedy solution [25] to the set cover problem in our simulations. However, it should be mentioned that the layer algorithm, which is bounded by a constant, can also be used if for any user the number of APs that it can associate with is bounded by a constant [25].

The greedy algorithm for set cover is well known. The cost version of greedy set cover algorithm is shown in Figure 8, which can be directly used to solve MLA problem after reducing it to an instance of the set cover problem.

Example – Centralized MLA: We run *CostSC* algorithm on the set cover problem corresponding to Figure 7.

Algorithm **CostSC**

$H \leftarrow \phi, X' \leftarrow X.$
while $X' \neq \phi$ do
  $A \leftarrow S_i$ s.t. $\frac{|S_i \bigcap X'|}{c(S_i)} = max_{D \in S} \frac{|D \bigcap X'|}{c(D)}$
  $H \leftarrow H \bigcup A, X' \leftarrow X' - A$
endwhile
output H

Figure 8. Algorithm CostSC [25]

S_4 is selected in the first round because it has the maximum value of $\frac{|S_4 \bigcap X'|}{c(S_4)} = \frac{3}{1/4} = 12$ among all S_i $(1 \leq i \leq 7)$. After that, $H = \{S_4\}$, $X' = \{u_1, u_3\}$. In the second round, S_2 is selected because it has the maximum value of $\frac{|S_2 \bigcap X'|}{c(S_2)} = \frac{2}{1/3} = 6$. After that, we get output $H = \{S_2, S_4\}$ because $X' = \phi$. Therefore, all users are associated with AP a_1, which is also the optimal solution.

The following theorem's proof is given in [25].

Theorem 6 *The algorithm CostSC is an* $(\ln n + 1)$-*approximation algorithm for set cover problem.*

6.2 Distributed MLA

Because the objective of MLA is to minimize the total load of the APs in the network, intuitively, a user should be associated with the AP which increases the total load minimally. Therefore, we use the same distributed algorithm for MLA as the one for MNU.

Example – Distributed MLA: Consider that the data rate of s_1 and s_2 is 1 $Mbps$ in the WLAN in Figure 1, and users use the distributed algorithm in the order u_1, u_2, u_3, u_4, u_5. First u_1, u_2 is associated with a_1. After that, u_3 is associated with a_1 because the total load of u_3's neighboring APs a_1 and a_2 is $\frac{1}{3} + \frac{1}{6} = \frac{1}{2}$ if u_3 is associated with a_1 and the total load is $\frac{1}{3} + \frac{1}{6} + \frac{1}{5} = \frac{7}{10}$ if u_3 is associated with a_2. Similarly, u_4, u_5 are associated with a_1. Eventually, all users are associated with AP a_1, which is also the optimal solution.

7 Performance Evaluation

In this section we report on performance studies of the proposed association algorithms for multicast using simulations in the Network Simulator ns2 [18]. The simulation source code can be downloaded from http://www.cse.ohio-state.edu/∼chenai/ICDCS07/. The simulation results show the average performance of our algorithms, while our analysis in the previous sections only shows the performance of our algorithms in the worst cases. We compare the performance of the three algorithms, MLA, BLA, and MNU, with the signal strength based association algorithm (SSA), which always lets a user associate with the AP providing the strongest signal among all the neighbor APs of this user. We have simulated the proposed algorithms over $1.2km^2$ area with upto 200 APs and 400 users randomly located in the area. The radio propagation range of both AP and user is 200m. The transmission rates and their distance thresholds are shown in Table 1. The users collect information of neighbor APs using active scanning [19]. Every user joins one multicast session. The APs operate in IEEE 802.11a infrastructure mode. We use 0.9 as the load limitation of multicast for every AP. Unless otherwise specified, we use 5 multicast sessions. Each user selects one of the multicast sessions at random. These simulation settings are used for all algorithms unless mentioned otherwise. We depict the average, min and max values for 40 random scenarios in the figures.

Rate (Mbps)	6	12	18	24	36	48	54
Distance Threshold (m)	200	145	105	85	60	40	35

Table 1. Transmission Rate vs. Distance Threshold [17]

Minimize Load of APs: Figure 9 shows the total load (the summation of all AP's multicast loads) with respect to the number of users, APs, and sessions, respectively. Figures 9(a) and 9(c) show that the total AP load increases, as the number of users and the number of sessions increase because of increased multicast demand. The total AP load, however, has an inverse relationship to the number of APs as shown Figure 9(b). The reason is that the resulting increased density of APs allows for higher transmission rate between APs and users.

We can observe that the centralized and distributed MLA algorithms perform better than SSA through simulations in Figure 9. The total multicast load of the centralized MLA and the distributed MLA perform 31.1% and 30.1% better than that of SSA at 400 users, respectively in Figure 9(a). The distributed algorithm performs only slightly worse (up to 5%) than the centralized algorithm.

Balance Load among APs: Figure 10 shows the maximum load among APs with respect to the number of users, APs, and sessions, respectively. The centralized and distributed BLA algorithms have upto 52.9% and 50.5% lower maximum load than SSA at 400 users, respectively (Figure 10 (a)). Moreover, unlike the SSA algorithm, for the distributed and centralized BLA algorithms, the maximum load increases slowly with the number of users or sessions (Figures 10 (a) and 10 (c)). Figure 10 (b) shows that the maximum load decreases as the number of APs increases, since the multicast load can be shared by more APs. We observe that the centralized and distributed BLA algorithms have similar performance.

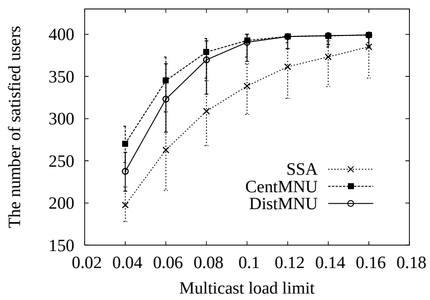
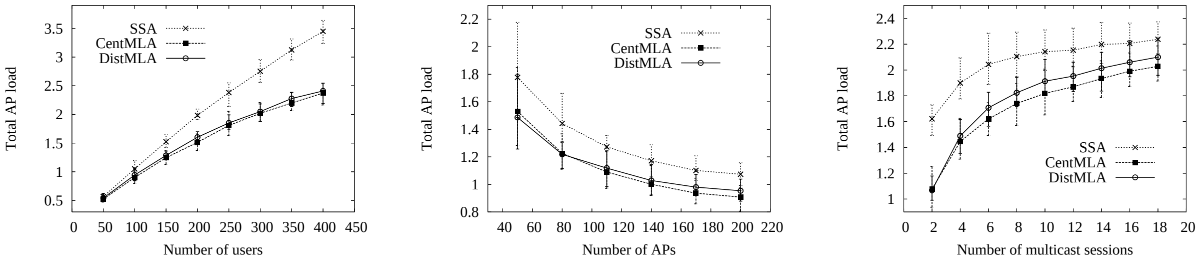


Figure 11. The number of satisfied user with respect to multicast load limit (budget) with 400 users, 100 APs, and 18 multicast sessions.

Maximize Number of Users: Figure 11 shows the number of satisfied users with respect to the multicast load limitation. We define the multicast load in Definition 1. As the multicast load limitation increases, the number of satisfied user increases as well. The satisfied number of users of the centralized and distributed MNU algorithms are 36.9% and 20.2% higher than that of SSA at the load limitation 0.04.

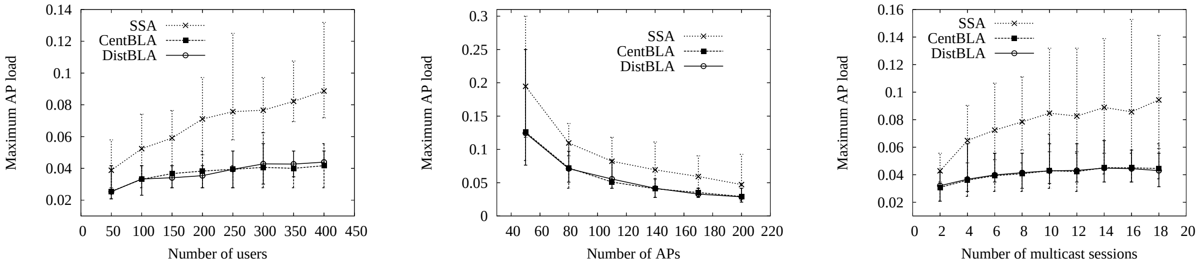
Optimal Solutions: In Figure 12, we evaluate the optimality of MLA, BLA, and MNU algorithms. Because there has been no other research work considering the problems mentioned in this paper, we have implemented ILPs for MLA, BLA, and MNU problems based on the ILP of set cover problem to compute the optimal solutions. Note that MLA, BLA, and MNU are NP-hard problems. As ILP takes exponential time to arrive at solutions, we limit our evaluation to small networks. The total AP loads of the centralized and distributed MLA algorithms are 25% and 22.2% higher than the one of the optimal solution at 30 users in Figure 12 (a). The maximum loads among APs of the centralized and distributed BLA algorithms are 12% and 22.6% higher than the one of the optimal solution at 40 users in Figure 12 (b). Although on average the MNU algorithm performs much closer to the optimal algorithm than the SSA algorithm, the maximum number of unsatisfied users for the centralized and distributed MNU algorithms are 5 and 8 respectively (Figure 12(c), 50 users[6]) but for the optimal solution it is 1. If networks are small, it is possible in some scenarios for the distributed algorithms to perform even better than the centralized algorithms (Figure 12(a)). The reason is that in small networks, the distributed algorithms have relatively more global information than in case of large networks.

[6]These details are not visible in the graph due to overlapping error bars



(a) (b) (c)

Figure 9. Total AP load for multicast sessions. Figure (a) varies the number of users for 200 APs, Figure (b) changes the number of APs with 100 users, and Figure (c) changes the number of session with 200 APs and 200 users.



(a) (b) (c)

Figure 10. Maximum load among APs for multicast sessions. Figure (a) varies the number of users for 200 APs, Figure (b) changes the number of APs with 100 users, and Figure (c) changes the number of session with 200 APs and 200 users.

8 Discussions and Future Work

In this section we outline some open issues that we are current investigating.

Distributed Convergence: We have shown that in some cases, simultaneous association decisions by multiple nodes may not necessarily result in global optimization of the objective functions. We are currently working on local coordination mechanisms to guarantee optimization of the global objectives at each step. An idea that we are currently exploring uses explicit *locks* from neighboring APs before committing to a change in association. We are exploring issues such as deadlocks, communication overhead and delayed association for such approaches.

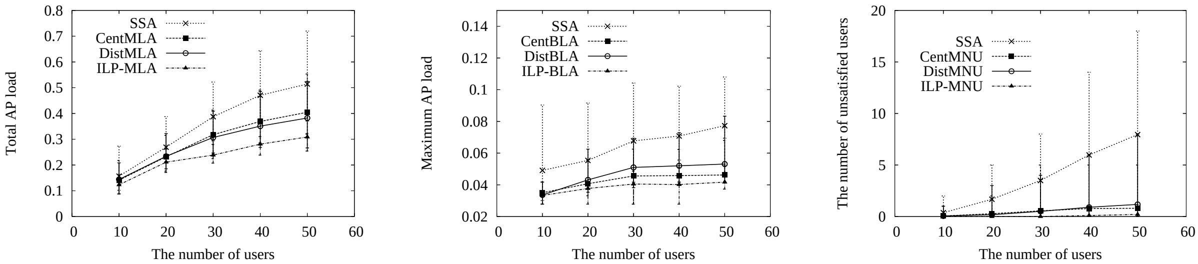
Adaptive Power Control: Adaptive power control provides an additional degree of flexibility that has not not been explored in this work. We are currently working on approximation algorithms based on a generalized network model that allows nodes to choose from a finite set of discrete power levels.

Explicit Interference Modeling: The approximation algorithms need to be modified to explicitly account for interference from neighboring users and APs[7]. In addition, along with such explicit interference models, it is necessary for the nodes to dynamically maintain a list of interfering nodes. To keep track of interfering sources, naive solutions proposed in the literature use explicit beaconing at high power levels, which adds extra overhead. We are currently investigating extensions to our work that use explict interference modeling along with practical ways to keep track of interfering sources"

9 Conclusion

Multicast services must be deployed with minimal impact to existing unicast services in WLANs. The problem of enabling multicast based streaming services in large-scale WLANs has not received attention in the past. Motivated by recent reports of dense deployments of APs in WLANs, we study techniques for exploiting overlapping coverage from neighboring APs to optimize performance. Three objective functions motivated by different revenue functions

[7]Our solutions to MLA and BLA implicitly optimize interference



(a) (b) (c)

Figure 12. The optimal solution of MLA, BLA, and MNU problems with respect to the number of user with 30 APs: (a) total AP load, (b) maximum AP load among all APs, and (c) the number of unsatisfied user with multicast budget 0.042. 30 APs and 50 users are randomly located in $600m^2$ area.

and network scenarios are studied: maximizing the number of users (MNU), balancing the load among APs (BLA) and minimizing the load of APs (MLA). We show that these problems are NP-hard. We present centralized approximation algorithms and distributed algorithms for these problems. Using simulations we evaluate the performance of these protocols and find that compared with multicasting from associated APs chosen based on strongest signal, the number of users can be increased by up to 36.9%, and the maximum AP load and the total load can be reduced by up to 52.9% and 31.1%, respectively. We conclude that the impact of multicast services on unicast services in WLANs can be effectively reduced by association control mechanisms.

References

[1] A. Balachandran, G. Voelker, P. Bahl, and V. Rangan, "Characterizing User Behavior and Network Performance in a Public Wireless LAN," in *Proc. of ACM SIGMETRICS'02*, Marina Del Rey, June 2002.

[2] Y. Bejerano, S.-J. Han, and L. Li, "Fairness and Load Balancing in Wireless LANs Using Association Control," in *Proc. of Mobicom*, Sept. 2004.

[3] P. Chaporkar, A. Bhat, and S. Sarkar, "An Adaptive Strategy for Maximizing Throughput in MAC Layer Wireless Multicast," in *Proc. of MobiHoc'04*, Roppongi, Japan, May 2004, pp. 256–267.

[4] P. Chaporkar and S. Sarkar, "Stochastic Control Techniques for Throughput Optimal Wireless Multicast," in *Proc. of Control and Decision Conference (CDC)*, Maui, Hawaii, Dec 2003.

[5] C. Chekuri and A. Kumar, "Maximum Coverage Problem with Group Budget Constraints and Applications," in *Proc. of the APPROX*, 2004.

[6] C. Y. Chiu, E. H. Wu, and G. H. Chen, "A Reliable and Efficient MAC Layer Broadcast (Multicast) Protocol for Mobile Ad Hoc Networks," in *Proc. of Global Telecommunications Conference*, 2004, pp. 2802–2807.

[7] C. T. Chou and A. Misra, "Low Latency Multimedia Broadcast in Multi-rate Wireless Meshes," in *Proc. of the 1st IEEE Workshop on Wireless Mesh Networks*, 2005, pp. 54–63.

[8] Y. Fukuda, T. Abe, and Y. Oie, "Decentralized access point selection architecture for wireless LANs," in *Proc. of Wireless Telecommunications Symposium*, Pomona, CA, USA, May 2004.

[9] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman Publishing Company, 1979.

[10] S. Gupta, V. Shankar, and S. Lalwani, "Reliable Multicast MAC Protocol for Wireless LANs," in *Proc. IEEE ICC03*, May 2003, pp. 93–97.

[11] G. Judd and P. Steenkiste, "Fixing 802.11 access point selection," in *ACM SIGCOMM 2002 Poster Session*, July 2002.

[12] F. P. Kelly, A. Maulloo, and D. Tan, "Rate Control in Communication Networks: Shadow Prices, Proportional Fairness and Stability," *Journal of the Operational Research Society*, vol. 49, pp. 237–252, Mar. 1998.

[13] D. Kotz and K. Essien, "Analysis of a Campus-wide Wireless Network," in *Proc. of ACM MobiCom'02*, Marina Del Rey, September 2002.

[14] A. Kumar and V. Kumar, "Optimal Association of Stations and APs in an IEEE 802.11 WLAN," in *Proc. of the National Conference on Communications*, IIT Kharagpur, Jan. 2005.

[15] J. Kuri and S. K. Kasera, "Reliable Multicast in Multi-Access Wireless LANs," in *Proc. of INFOCOM*, 1999, pp. 760–767.

[16] D. Lee, G. Chandrasekaran, and P. Sinha, "Optimizing broadcast load in Mesh Networks using Dual-Association," in *1st IEEE Workshop on Wireless Mesh Networks*, Santa Clara, CA, USA, 2005.

[17] M. H. Manshaei and T. Turletti, "Simulation-Based Performance Analysis of 802.11a Wireless LAN," in *Proc. of International Symposium of Telecommunications (IST)*, Isfahan, Iran, August 2003.

[18] ns2: Network Simultor, ," http://www.isi.edu/nsnam/ns/.

[19] I. Ramani and S. Savage, "SyncScan: Practical fast handoff for 802.11 infrastructure networks," in *Proc. of the IEEE Infocom Conference*, Miami, FL, USA, March 2005.

[20] W. Si and C. Li, "RMAC: A Reliable Multicast MAC Protocol for Wireless Ad Hoc Networks," in *Proc. of ICPP'04*, 2004, pp. 494–501.

[21] M. T. Sun, L. Huang, A. Arora, and T. H. Lai, "MAC Layer Multicast in IEEE 802.11 Wireless Networks," in *Proc. of ICPP*, 2002.

[22] K. Tang and M. Gerla, "MAC Layer Broadcast Support in 802.11 Wireless Networks," in *Proceedings of IEEE MILCOM*, 2000.

[23] K. Tang and M. Gerla, "Random Access MAC for Efficient Broadcast Support in Ad Hoc Networks," in *Proc. of WCNC*, 2000.

[24] K. Tang and M. Gerla, "MAC Reliable Broadcast in Ad Hoc Networks," in *Proceedings of IEEE MILCOM*, 2001, pp. 1008–1013.

[25] V. Vazirani, *Approximation Algorithms*, Springer Verlag, 2001.

[26] L. Wang, Z. Niu, Y. Zhu, H. Deng, and M. Yano, "Integration of SNR, load and time in handoff initiation for wireless LAN," in *Proc. of PIMRC*, September 2003, vol. 3, pp. 2032–2036.

Appendix

A. Reduce the Subset Sum Problem to the MNU Problem

The Subset Sum problem is defined as follows.

Definition 4 ***Subset Sum Problem:*** *Given a set of natural numbers $G = \{g_1, g_2, \ldots, g_k\}$ and a target number T, decide if a set $S \subseteq \{1, 2, \ldots, k\}$ exists, such that $\sum_{i \in S} g_i = T$.*

Theorem 7 *The Subset Sum problem can be reduced to the MNU problem.*

Proof: We present a reduction from an arbitrary instance of the subset sum problem. We construct a WLAN with a single AP. The maximum load this AP can support for multicast traffic is T. The WLAN supports k multicast sessions $s_1, s_2, \ldots, s_k$, where session s_i creates a load of g_i when transmitted at unit data rate. Corresponding to each session we add g_i users requesting session s_i. Each user has a link of unit data rate to the AP. If the maximum number of users this WLAN can support for multicast traffic is T, the Subset Sum problem has a positive answer; otherwise, it has a negative answer. Note that the maximum value of the load of an AP is 1 according to Definition 1 while T can be any natural number. However, we can make them less than 1 by dividing the maximum load T of AP and the load g_i of every multicast session s_i by a large enough number. □

B. Reduce the Minimum Makespan Scheduling Problem to the BLA Problem

Definition 5 ***Minimum Makespan Scheduling:*** *Given processing times for n jobs, $p_1, p_2, \ldots, p_n$, and an integer m, find an assignment of the jobs to m identical machines so that the completion time, also called the makespan, is minimized.*

Theorem 8 *The Miminum Makespan Scheduling problem can be reduced to the BLA problem.*

Proof: We construct a WLAN with m APs. Every AP only provides one transmission rate to users. The multicast sessions supported by this WLAN are $s_1, s_2, \ldots, s_n$. All APs can provide service to all users. The load requirement for a multicast session s_i is p_i ($1 \leq i \leq n$). The objective is to minimize the maximum value of an AP's load among all APs under the limitation that all users get multicast service. □

C. Reduce the Set Cover Problem to the MLA Problem

We first define the Set Cover problem, and then present a reduction from the set cover problem to MLA problem to show that MLA is an NP-hard problem.

Definition 6 ***Set Cover:*** *There are m subsets $S_1, S_2, \ldots, S_m$ of a ground set $X = \{u_1, u_2, \ldots, u_n\}$. A cost $c(S_j)$ is associate with each set S_j. The objective is to find a subset H of $S = \{S_1, \ldots, S_m\}$ which covers all elements of X and has the minimum total cost. If the cost of every subset S_j is a same value c, the set cover problem is a cardinality version.*

Theorem 9 *The cardinality version of set cover problem is reducible to MLA.*

Proof: We construct a WLAN with m APs, $a_1, a_2, \ldots, a_m$, and n users, $u_1, u_2, \ldots, u_n$. In this WLAN, all users request for the same multicast stream session with load requirement c. AP a_j can provide service to users in subset S_j ($1 \leq j \leq m$). The link between the AP and the user has a unit data rate. The objective is to minimize the total load of all APs under the limitation that all users receive multicast service. □