Resource and Bandwidth Allocation in Hybrid Wireless Mobile Networks

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Benjamin John Bengfort

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Computer Science

September 2010

Fargo, North Dakota

North Dakota State University
Graduate School

Title

RESOURCE AND BANDWIDTH ALLOCATION

IN HYBRID WIRELESS MOBILE NETWORKS

By

BENJAMIN JOHN BENGFORT

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

ABSTRACT

Bengfort, Benjamin John, M.S., Department of Computer Science, College of Science and Mathematics, North Dakota State University, September 2010. Resource and Bandwidth Allocation in Hybrid Wireless Mobile Networks. Major Professor: Dr. Weiyi Zhang.

In the lead up to the implementation of 802.16 and 4G wireless networks, there have been many proposals for addition of multi-hop MANET zones or relay stations in order to cut the cost of building a new backbone infrastructure from the ground up. These types of Hybrid Wireless Networks will certainly be a part of wireless network architecture in the future, and as such, simple problems such as resource allocation must be explored to maximize their potential. This study explores the resource allocation problem in three distinct ways. First, this study highlights two existing backbone architectures: max-coverage and max-resource, and how hybridization will affect bandwidth allocation, with special emphasis on OFDM-TMA wireless networks. Secondly, because of the different goals of these types of networks, the addition of relay stations or MANET zones will affect resource availability differently, and I will show how the addition of relay stations impacts the backbone network. Finally, I will discuss specific allocation algorithms and policies such as top-down, bottom-up, and auction-based allocation, and how each kind of allocation will maximize the revenue of both the backbone network as well as the mobile subscribers while maintaining a minimum Quality of Service (or fairness). Each of these approaches has merit in different hybrid wireless systems, and I will summarize the benefits of each in a study of a network system with a combination of the elements discussed in the previous chapters.

ACKNOWLEDGEMENTS

My work at NDSU could not have been accomplished successfully if not for the inspiration, support, and brilliance of my fellow researchers and classmates: Ranjana Sharma, Phil Reindl, Will Voorhees, and Paul Loree. I would also like to thank my committee members, Dr. You and Dr. Kong for their support in finalizing my paper. I should also give special note to Tim and Pat Hanna, without whom I would not have survived winters in North Dakota and a seventy-five mile commute. Two years ago, Dr. Nygard took the time to listen to a random applicant, and it is him I must thank for accepting me into the program in the first place, even though I had no background in computer science. Thank you also to my adviser and mentor, Dr. Weiyi Zhang, without whom this thesis could never have been written; it is a high point in my career that I got to work for him. Finally, to my dear and brilliant wife--even though our work kept us apart, your support has made that separation successful, and this thesis is the result.

TABLE OF CONTENTS

Abstract iii

Acknowledgements iv

List of Tables vii

List of Figures viii

Introduction 1

Related Work 7

- Hybrid Wireless Networks 7
- Resource Allocation 10
- Other Foundational Research 15

Real World Data: A Speedtest.Net Experiment 18

Dynamic Resource Allocation via Hybrid Wireless Networks 23

- Application Aware Allocation 25
- Congestion and Interference Avoidance 27
- Best QoS-First Allocation 28

Impact of Relay Station Hybridization on Backbone Networks 31

Allocation Methodology and Notation 36

- Simulated Network Environment 37
- Network Operation Over Time 39

Calculating Profit 41

Approaches to Resource Allocation 43

Approach 1: Equality Maximization Top-down Allocation 43

Approach 2: Bottom-up Allocation 50

Approach 3: Auction Based Allocation 55

Approach Evaluation 58

Conclusion 61

References 64

LIST OF TABLES

Table | Page

1. Context and Benefits of Hybridization 24
2. Application Data Usage Requirements 26
3. Equality Maximization Top-Down Allocation Algorithm 43
4. Top Down, Fairness-First Allocation Results 46
5. Top Down, Highest-First Allocation Results 47
6. Top Down, Lowest-First Allocation Results 48
7. Bottom-Up Allocation Algorithm (Profit to Nodes) 51
8. Bottom Up, Profit to Nodes Ratio Results 53
9. Bottom Up, Profit to Cost Ratio Results 54
10. Auction Allocation Results 57

LIST OF FIGURES

Figure Page

1. Download Speeds 19
2. Upload Speeds 20
3. Latency 21
4. AT&T Style Network 32
5. Verizon Style Network 33
6. Proposed Network Diagram 38
7. Heuristic Step Weight Function 42
8. Top-Down Allocation Network Graph 45
9. Top Down Allocation with Equal Weights Comparison 49
10. Top Down Allocation with Heuristic Step Weights Comparison 49
11. Bottom-Up Allocation with Profit to Node Ratio Network Graph 52

INTRODUCTION

Popular discourse concerning the two biggest wireless network providers in the United States goes something like this: "AT&T has great data rates, and gorgeous smart phones, but you can't make a call, and don't even think about trying to leave the city." "Sure, I can hear you now, anywhere in the country, and call quality is great-- but can I get some high speed bandwidth on my Verizon phone?" Perhaps these sentiments are best expressed in the recent advertising wars between the two companies-- Verizon showing off its coverage maps, and AT&T simply riding on the "there's an app for that" iPhone sales pitch. Customers are not incorrect in their observations about the comparative network performance. In the race to build infrastructure and thereby gain market share, the largest wireless companies have pursued divergent strategies, often choosing between geographic coverage and high-speed throughput; but, as we have approached market saturation, it has become clear that both throughput and coverage are essential to the success of a backbone network. Now, as we near a second phase of infrastructure networks, with WiMAX, LTE, and 4G networks slowly replacing the old infrastructure, the result has been the same cost without a significant increase in customers. In order to fill in the gaps left by the choice of throughput vs. performance, the hybridization of wireless networks has recently become a widespread notion.

High-speed Internet devices that are rapidly becoming unburdened from the tether of Ethernet wired networks have fueled the boom in wireless infrastructure. It used to be that there were really only two, highly disparate types of devices-- mobile phones that required data for voice and simple messaging, and then much larger laptop devices that used air cards to achieve Internet connectivity, and had similar bandwidth requirements to broadband cable connections. Today, there are a number of devices that fill the gap between those two extremes, creating a much richer network fabric-- and creating more and more difficulties for bandwidth allocation. These powerful personal wireless devices—a middle ground of smart phones, media players, and content readers--rely on a bevy of technologies to power their apps, video streams, and on demand content services-- and the backbone technology that transports data between these devices is a combination of Wireless LAN (WLAN), cellular 3G networks, WiMAX networks, wireless mesh networks, and mobile ad hoc networks (MANET). Internet Smart Phones, devices such as the iPhone, Android phone, and Palm phone combine traditional cell phone voice and data with rich Internet data like video and browsing, relying on multiple radios, both 802.11 and 3G, in order to maintain consistent, high bandwidth connectivity.

There are even devices that don't resemble traditional cell phones at all. The Kindle Reader, iPad, and netbooks are scaled down versions of laptops that don't have the traditional cell phone data and voice usage, but still make use of 3G cellular networks to achieve Internet connectivity. Amazon's Kindle, for example, relies on a proprietary "whisper-net" based on Sprint's cellular network. Multiple and nontraditional radio devices show a second perspective in the need for hybridization-- vertical as opposed to horizontal integration of wireless infrastructure layers. While they certainly require more network resources than simple phones, they use less than laptops requiring always on broadband connections.

The combination of 802.11 radios with 3G radios in order to insure that the device can get the proper amount of bandwidth sidesteps the issue of network availability by using an "any or all circuits approach." Bandwidth requirements only continue to increase, however, because of new, powerful "apps"—the Internet applications that are standard on the new breed of mobile device. These applications, including web browsing, apps, VOIP, and streaming media, require larger and larger amounts of bandwidth and backbone infrastructure resources to support the power of the mobile devices, and herein lays the challenge for modern networks. Previously, it was easy for network providers to simply allocate data channels specifically for high requirement laptops in predictable high usage areas. These users had a distinct preference to hard line Internet connections, or local wireless networks, which generally have a surplus of bandwidth; so it was easy to predict the locations where the most intense mobile wireless usage would be – areas like airports and train stations where no local area network existed. Coffee shops and libraries already had free 802.11 connectivity; therefore urban areas that contained these types of places were not high priority for dedicated data channels. AT&T and T-Mobile went so far as to offer 802.11 Internet service in places like Starbucks and Barnes and Noble for their mobile customers in order to reduce the saturation of their cellular towers. Now, modern wireless networks provide "last mile" broadband connectivity, and must provide multiple entry points into the wide area network, where there are no clear geographic constraints like those that laptop users faced.

Static wireless networks are necessarily constrained in their ability to support the requirements of its users because of a myriad of factors, including the mobile nature of those users, limitations in resources across segments of the network and routing inefficiencies between different devices. Prioritizing an 802.11 connection over a 3G connection lowers the burden on traditional cellular networks somewhat, but not completely. The emergence of a hybrid 802.11 and 802.16 network (Sprint 4G wireless at home) to replace cable only means that bandwidth allocation issues are still a priority. Other services like Clear Wireless High Speed Broadband for mobile computing on laptop computers with 802.11 radios, show a trend that in the near future, the emergence of hybrid wireless networks (HWN) is necessary to dynamically adapt to increasing wireless broadband demand, and not only provide cost minimization to expansion in backbone networks but also to increase resource availability.

Hybrid Wireless Networks are intended to achieve better results by dynamic network construction rather than stand alone static infrastructure or ad hoc networks. A hybrid wireless network can be defined as the combination of an infrastructure network such as a WiMAX, WLAN, or 3G Cellular networks with ad hoc components like MANETs or multi-hop relay stations in order to expand infrastructure coverage either horizontally (expanding coverage) or vertically (expanding bandwidth) at a low cost. HWNs allow mobile users the benefits of several types of networks-- performance and mobility in a seamless fashion. Hybrid wireless architecture has, in the recent past, mostly attempted to cope with the increase in demand by a subsequent increase in the availability of resources from the infrastructure portion of the network-- but these attempts still are unable to meet the demand. In this project, I propose to investigate a resource allocation method that will equitably provide bandwidth to all users while maximizing utility ("profit") for all portions of the HWN.

In this paper we will consider the hybrid network made up of a base station (BS) that connects to relay stations (RS) that in turn provide connectivity to mobile devices. The base station has a fixed amount of bandwidth that it can provide to each relay station, and each relay station has a set amount of demand based on the number and requirements of each of its mobile nodes. It soon becomes clear that when demand exceeds availability, the BS has to allocate its resources in the most efficient, equitable, and profit maximizing manner. The resource allocation problem itself is not a new one, but this paper proposes solutions in the specific context of a Hybrid Wireless network.

The rest of my thesis is organized as follows: after discussing other relevant work in the study of resource allocation, we will consider current bandwidth allocation from the two major network providers in a real world data throughput experiment. Following the results of my experiment, we will discuss how dynamic resource allocation can be achieved through context aware application that is only available in hybrid wireless network architecture. It soon becomes apparent that when a backbone network attempts to maximize either coverage or throughput, the hybridization of a wireless network, specifically through the addition of relay stations has different goals, which is discussed in the third section. Finally, we look at specific resource allocation algorithms in the final chapter. Our view of resource allocation won't consider such limitations as channel allocation (instead using OFDMA subchanneling to allocate exactly the amount of resources), inefficient routing, or load unbalance, but these factors will certainly be important as this research is continued. Instead, I will approach this problem as a knapsack problem, where every item has a value and a weight, and we need to maximize the value of the items, but minimize the weight of the knapsack. Similarly, the base station must maximize its potential use of bandwidth, and limit the number of users that do not receive allocated resources. It is my contention that by optimizing the mode of resource allocation for a specific network, HWN providers can maximize profitability and make better choices when choosing how to expand their existing backbone architecture.

RELATED WORK

A survey of the current literature on this topic reveals an interesting dichotomy: while there are plenty of papers considering hybrid wireless networks as low cost solutions to network expansion, and similarly many papers discussing resource allocation in next generation wireless networks, few consider the possibility of resource allocation in hybrid wireless networks. As such, this section of my paper is split into categories to discuss current literature on the relevant topics.

Hybrid Wireless Networks

Khadivi, et al discusses hybridization between WLAN (802.11) and cellular networks via a combination with multi-hop ad hoc zones. The use of ad hoc relaying is intended to reduce the probability of dropping via a proposed upward vertical hand-off between wireless network types, or between base-stations—not just as a coverage extension measure. The multi-hop routing zone facilitates the hand-off by providing the opportunity for continuous coverage back to the original access points. This ingenious scheme makes use of relay stations, not just cooperative mobile nodes, to facilitate multi-hop traffic. However, the mechanism for upward vertical handoffs requires a continuous connection to determine the best bandwidth available, a critical metric in determining routing (Khadivi, Todd, Samavi, Saidi, & Zhao, 2008).

The model of a hybrid ad-hoc and cellular network required a focus on the hand-off between the two network areas, with a benefit of extension of coverage and reduction in infrastructure costs as described in (Yamanaka & Shimohara, 2007). This paper focused on channel-based allocation, especially in the overlapping coverage area between ad-hoc and infrastructure zones. The use of relay stations placed in the overlapping areas was crucial the facilitation of hand-offs and in fact, the relay stations themselves create the multi hop zones so that base stations themselves do not have to be in range of each other. This paper did not necessarily discuss mobile to mobile routing, or other multi-hops, but implied that these additional features could be added with ease.

One of the most popular hybridizations, the combination of 802.11 and 802.16 wireless networks is notably discussed by (Prasath, Raghu, & Ma, 2009). Prasath et al use a MAC protocol instead of a scan & connect model, and a separation of traffic categories (real-time vs. non-real-time as defined by packet loss tolerance) to different connections through QoS (Quality of Service) connection management. The role of the base station in signaling identification and control packets as the basis of their integration, and in fact, WLAN base-stations play the critical role in handing off control to WiMAX stations. This paper can be seen as the logical follow-on of the previous discussion of the integration between 802.11 and 3G as in (Salkintzis, 2004).

Wireless Overlay Networks; a hierarchical structure of WLAN, multi-hop, and WAWN, is used to solve connectivity for a large number of mobile users and to connect low coverage/high bandwidth network zones with high coverage/low bandwidth networks by (Siddiqui & Zeadally, 2006), who also discuss the challenges between horizontal and vertical hand-offs. They propose that some high-level mobility management protocol is necessary to facilitate hand-offs between hybrid zones, as well as to provide minimum QoS and bandwidth allocation in the situation where mobile users move between network zones or different application requirements.

A possible high-level mobility management protocol is described as HYWINMARC by (Chaudhry, Akbar, Kim, Hong, & Yoon, 2006). This architecture is described as set of network policies and mobile agents used to autonomically manage a hybrid of ad-hoc MANET zones and mesh backbone networks. They suggest that any such architecture should be a combination of self-managing, self-configuring, and self-healing protocols to balance the network load. However, they don't necessarily discuss hand-offs between zones or dynamic MANET architectures.

Salem, et al shows that the hybridization of node-to-node ad hoc networks with infrastructure networks (instead of the use of multi-hop via relay stations) is possible through end-to-end connections with the base-station and a charging and reward scheme to facilitate cooperation (Salem, Buttyan, Hubaux, & Jakobsson, 2006). The nodes themselves do not participate in the resource allocation process, but rather are treated as individuals by the base-station, to ensure a fair cooperation mechanism and to prevent malicious nodes from operating unfairly in the network. Similarly, Weyland, Staub, and Braun compare the hybridization of cellular networks with node cooperative multi-hop zones vs. single hop networks. They found that in order to facilitate node cooperation, motivation-based schemes seem to have more of an impact on this network type rather than enforcement or single-hop schemes (Weyand, Staub, & Braun, 2006). They propose a profit based reward mechanism, where nodes are metered and charged for their own traffic in a decentralized basis, but are compensated for the packets of others that they forward. They also note that this scheme requires a centralized authentication scheme to prevent mal-actors from abusing the cooperation of others.

Finally, it is shown by (Chen & Wei, 2010) that the throughput capacity in hybrid, multi-channel networks with up to $O\log(n)$ channels suffers no degradation, thereby proving that a resource-allocation scheme can be optimized for a fixed amount of bandwidth.

Resource Allocation

Power and bandwidth are the two physical layer resources being allocated via two algorithms: a fairness-constrained greedy revenue algorithm in an adaptive power allocation module (APA), and a utility-constrained greedy approximation for call admission control (CAC) in non-hybrid 802.16 networks (Rong, Qian, & Lu, 2007). The authors suggest a separation of uplink and downlink systems and applications with separate allocation schemes in order to consider fairness and QoS because link directionality has different requirements. Optimization of both APA and CAC is an attempt to increase the revenue of a successful system of both providers and subscribers, whereas previously considered optimization problems only considered the utility of the provider. Also notable is the authors' use of OFDMA-TDD (Orthogonal Frequency Division Multiple Access--Time Division Duplex) to fully use subchannels on both the uplink and downlink schedule to precisely allocate resources to mobile nodes.

Another type of optimization via separation is achieved via distinguishing between real-time and non-real-time traffic (as defined by packet-loss tolerance) and allocating resources, in this case channels, to each type of traffic (Tzeng, 2006). Tzeng suggests that allocating to channel-type creates a boundary between the two traffic types that can be either fixed or movable. In movable boundary schemes, traffic is monitored in real-time, and the boundary between traffic types is moved to accommodate the type of traffic in the network, with a specific preference towards real-time traffic.

Another common mechanism for allocation comes in the form of game theoretic approaches. (Niyato & Hossain, 2007), (Touati, Altman, & Galtier, 2006), and (Salles & Barria, 2005) use microeconomic approaches to solve bandwidth allocation. Niyato & Houssain set up a non-cooperative game during the call admission procedure between a single base station and the new, requesting subscriber. New subscribers queue for admission to the network, and after the game achieves equilibrium they are allocated resources as the result of the game. Equilibrium is reached by target levels of satisfaction of QoS, while the conflict is the limitation of resources: whereas the base station wants to maximize revenue by admitting as many subscribers as possible, the subscriber wants to maximize the amount of resources available to it. Touati, Altman, and Galtier take a slightly different approach, utilizing a Generalized Nash Bargaining Solution rather than a non-cooperative game. Although it discusses more general network architectures, not just wireless systems, the paper proposes to use a fairness model defined by the users' utility differentiated across applications, rather than just an assigned amount of throughput and Nash equilibrium is found through per node quadratic utility functions. Performance in this case is measured via optimality and fairness rather than profit or some form of min-max optimization. Salles and Barria suggest a similar microeconomic scheme by attempting to maximize Rawlsian fairness and Paretian efficiency by a weighted fair queuing algorithm. Quantitative, piecewise, linear utility functions are assigned to different applications types including multimedia, video, VOIP, MPEG, TCP, etc. and are aggregated to determine their weights during traffic analysis. These weights are then used to determine allocation and queuing schemes to each application type.

Agent-based allocation is suggested as an alternative to game theoretic approaches in (Manvi & Venkataram, 2005) although similarly not solely in wireless networks. The agent-based approach by Manvi and Venkataram considers bandwidth usage as "bursty" and that all other QoS parameters revolve around the available bandwidth. Since available bandwidth is static, agents must do resource optimization, by re-routing bandwidth from congested lines during the network run-time. Crucially, these agents don't consider the initial state, and are only partially aware of network topographical information when making decisions, making this scheme applicable to a wide array of networks and easily scalable.

To return to wireless network-specific resource allocation: (Lin, Lin, Lai, & Wu, 2009) and (Thulasiraman & Shen, 2010) attempt to account for latency and interference in wireless networks in order to allocate resources efficiently. Lin, et al. proposes a Highest Urgency First policy, which considers a grant per subscriber station (GPSS) operation rather than grant per connection (GPC). In this scheme, the base station grants bandwidth to a specific mobile subscriber rather than to a channel or connection so that the mobile subscriber can request varying levels of QoS depending on his requirements. Working within the confines of the minimum reserved rate, the maximum sustained rate, and the maximum latency, the base station can assign an urgency parameter to mobile nodes that considers fairness, latency, modulation, and priority, and via OFDMA-TDD can assign exactly the amount of physical layer resources to the highest urgency mobile nodes in decreasing order. Thulasiraman and Shen have a slightly different take on the constraints of allocation in hybrid-wireless networks, where hybridization essentially is the deployment of relay stations and multi-hops to achieve low cost additional geographic coverage. Here, the resources available are subcarriers and power, and the major limiting factor of bandwidth across multiple hops is interference, not a bandwidth availability cap. They propose a spatial reuse of subcarriers to improve throughput, and avoid interference. I considered this work similar to the agent-based dynamic rerouting around congested areas, with the difference being that congestion is referred to as interference.

Elias, et al. proposes a dynamic resource allocation scheme, as opposed to what they consider current static mechanisms (Elias, Martignon, Capone, & Pujolle, 2007). Interestingly, their scheme is not maximization or utility based; instead users subscribe for a guaranteed transmission rate for short-term contracts. Allocation of remaining resources then occurs at the network edges in a bottom-up fashion using traffic shaping schemes. This is an accounting model more than a maximization model, and in fact, the paper defines performance as a combination of network load and revenue generated from contracts for increased bandwidth. Similarly, Ahn and Kim discuss bandwidth adaption--the optimal distribution of unallocated bandwidth to cells that require additional resources (Ahn & Kim, 2003). They propose that adaption necessarily generates more hand-offs, and therefore use a binary linear integer optimization for distributing those resources.

I would like to reserve some special attention to multicast specific resource allocation, as multicast applications are integral parts of the networks we are discussing, both for emergency services, as well as control and signaling in hybrid networks. In (Kuo & Lee, 2010), Kuo and Lee discuss recipient maximization in 802.16 networks. They define an interesting problem in multicast applications that is not necessarily relevant in other resource allocation schemes--namely that different users have different bit-error rates and therefore require different resources even for the same data from the base station! They also suggest hybridization via the addition of multi-hop relay stations could assist in maximizing multicast recipients, but otherwise propose a mechanism called dynamic station selection to solve the problem. DSS solves the maximization problem by creating an envelope tree from the network topology of node utility functions and assigning resources to the node with the largest utility (the root) down the tree until budget is exhausted. In (Lee & Cho, 2007), Lee and Cho discuss discrete bandwidth allocation in the context of lexicographical optimal fairness for multicast networks. Fairness here is described as a non-increasing convex function, which essentially maximizes the minimum components. Instead of assigning the bandwidth from the root of the tree downwards as in (Kuo & Lee, 2010), they assign bandwidth to the leaf nodes first towards the root.

Other authors propose that in a wireless network, simple throughput allocation is not the issue, but rather simple transmission scheduling will provide resource allocation on the physical layer. In (Huang, Subramanian, & Agrawal, 2009), Huang, et al. again proposes the use of OFDM to allow physical layer allocations through the use of subchannels. Here, resource allocation is a maximization based on capacity, and a logarithmic utility function can be derived as the number of nodes approaches the bandwidth capacity of the base station. In (Nascimento, Rodriquez, Mumtaz, Gameiro, & Politis, 2008), Nascimento, et al. suggest that resources are radio signal, power, channels, time slots, and spatial beams, OFDM scheduling is used to achieve dynamic resource allocation, and they propose several allocation algorithms including Max C/I, proportional fair, and round robin.

Other Foundational Research

Klasing, et al. discusses some of the difficulties in bandwidth and resource allocation in wireless networks, for instance radio resources are disturbed by interference and physical constraints. Additionally complexity is injected because of the necessity to schedule radio transmission in order to satisfy traffic demands. Wireless networks can essentially be seen as a multi-commodity flow problem, modeled as independent sets (Klasing, Morales, & Perennes, 2008). Similarly, Savkin, et al. presents a problem of hybrid dynamical system of both continuous and discrete states, where nodes create interference. They approach the solution to node interference as transmission scheduling (Savkin, Matveev, & Rapajic, 2005).

My paper proposes to essentially solve a knapsack problem, and (Lin F.-T. , 2008), (Fujimoto & Yamada, 2006), (Egeblad & Pisinger, 2009) are unique approaches to the knapsack problem. Lin discusses imprecise weight coefficients, where decision makers can only make rough estimates of weights, and a solution using genetic algorithms. Imprecision, in this case, is remedied by the use of fuzzy number sets, and is important to my work because weighting nodes to determine revenue can be seen as an imprecise calculation. Fujimoto and Yamada discuss a knapsack sharing problem with multiple players who each have their own profit weights on a set of items--and where some items are commonly profitable to a set of the players. In the sense of my paper, relay stations and base stations can all be seen as players, who possibly have nodes as common items. Egeblad and Pisinger discuss two--and three--dimensional knapsack problems, which at first seems irrelevant to my paper, but because they use a heuristic model and a simulated annealing search problem to find solutions to space constraints, this can be applied to the problem of localization in hybrid wireless networks, where relay stations don't know the overall bandwidth available in different regions of the network.

Another solution to the resource allocation/call admission procedure can be found via optimization rather than a dynamic programming solution state. Wu and Bertsekas propose an approximate dynamic programming technique in order to reduce the computational demand of a large state space. Networks formulated as Markov decision models are approximated to a smaller state space so that off-line calculation is possible to optimize the network space, thereby reducing the performance load during real-time network operations (Wu & Bertsekas, 2007).

Finally, many of these papers rely on Markov chains to do performance analysis or network simulation. I utilized the Manoj et al. discussion on multidimensional Markov chains for performance evaluation, specifically because the paper focused on Markov chains in Twill networks--throughput enhanced wireless local loops that are hybrid multi-hop and backbone networks (Manoj, Ranganath, & Murthy, 2006).

REAL WORLD DATA: A SPEEDTEST.NET EXPERIMENT

In order to understand real world bandwidth allocation, I conducted an experiment to gather actual data rates from both Verizon and AT&T. In this experiment, I used two modern smart phones (Internet enabled, dual radio): a Motorola Droid running the Android 2.1 operating system on the Verizon network, and an Apple iPhone running the iPhone 3G operating system on the AT&T network.

I made use of an app (Internet applet) that was available both for the iPhone and for Android phones: Speedtest.Net. The Speedtest app simply finds a server in geographic proximity to the device, and then tests the available bandwidth in three ways. First, latency is tested by measuring, in milliseconds, how long a packet takes to travel to the server and back; in essence a ping test without the use of the echo protocol. Next, a file of significant size (a video, audio, or large text file) is downloaded to the phone. The app calculates the speed at which the download is taking place every second, then outputs the average download speed in kilobytes per second. Finally, the file is uploaded back to the server, and the upload speed calculated. Using these measurements a clear picture of the available network bandwidth is obtained.

My experiment consisted of 24 bandwidth tests--two an hour for 12 hours. The results, detailed in Figures 1-3, showed different available bandwidths across the two different networks, and highlighted how the different network architectures responded to heavy bandwidth allocation requests.

Analysis of download speeds, displayed in Figure 1, show that the Verizon network allows more consistent available bandwidth. However, the available bandwidth is much less, almost 50% that of the throughput that AT&T provides to the iPhone. There is something to be said for consistency; while AT&T allows blazing fast speeds skyrocketing to 3 MBps, it can bottom out just as often. Except during peak hours in the afternoon, Verizon maintained a fairly consistent distribution of bandwidth. These very small-scale results are consistent with the popular notion of AT&T vs. Verizon performance that can be found extensively in the media and in technology news (Laporte, 2010).

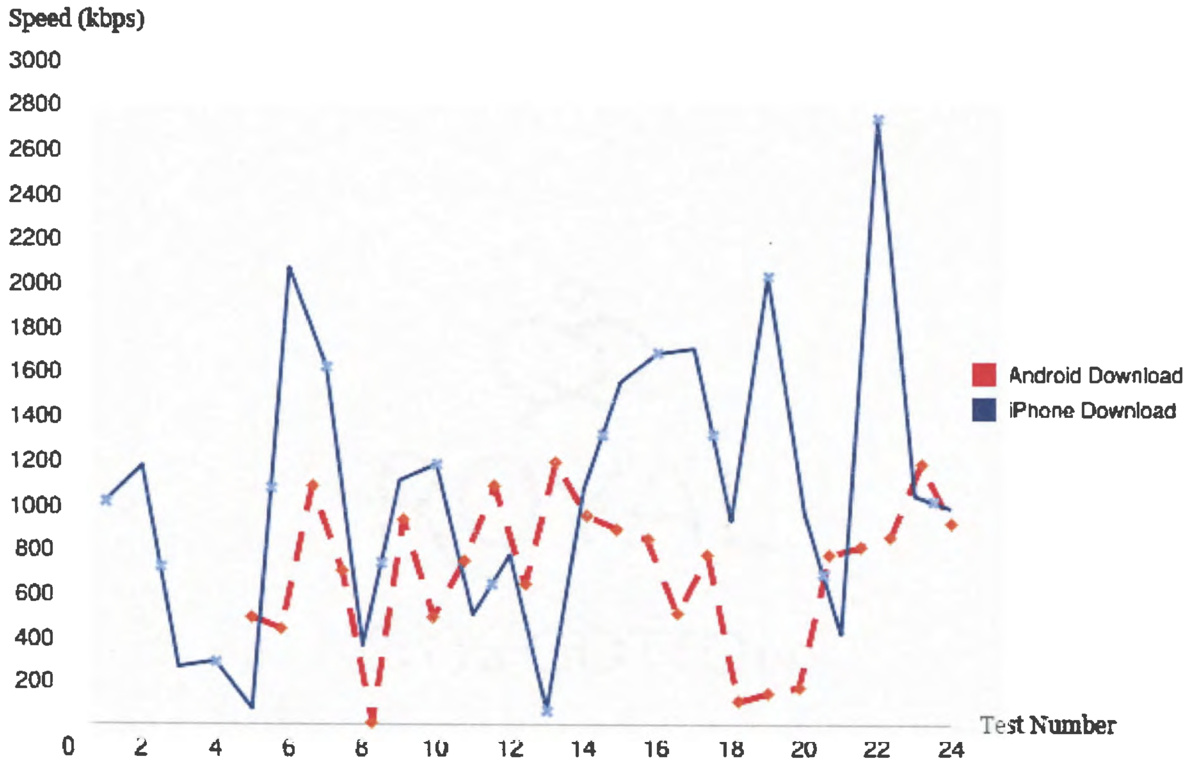

Figure 1. Download Speeds

The upload speeds show another interesting effect of the backbone architectures of the two networks, as charted in Figure 2. The Verizon network surprisingly surpasses the available upload bandwidth of AT&T. This could be because Verizon makes no differentiation between upload and download channels--they are both the same consistent lower data rate. However, the data shows that the AT&T network attempts to maximize the download speeds (by far the biggest usage of data) by favoring download channels over the upload ones. AT&T uploads are so close to zero, that I would hazard that they might even use similar channels to the control channel (which is used for incoming call notification, push messaging, geolocation, and SMPP).

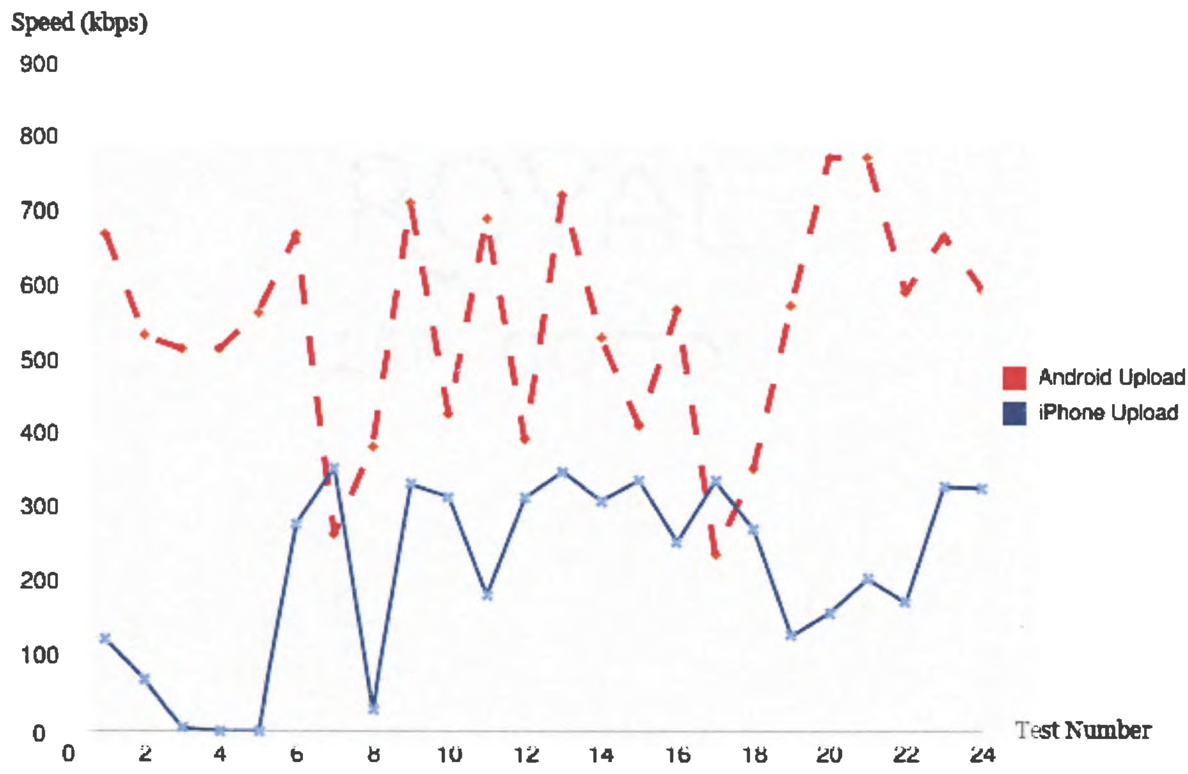

Figure 2. Upload Speeds

On average, the latency of both networks is very similar, as expected. The latency results of the experiment are detailed in Figure 3. The backbone networks are optimized in terms of routing, and latency has very little to do with bandwidth since ping is a very small packet size. However, AT&T spikes coincide with the drops in bandwidth in the download graph--meaning there was almost no available bandwidth during those tests. Again, the consistency of the Verizon network and the lower latency with spikes of inoperability for AT&T is consistent with the popular language concerning network architectures in the media (Vogelstein, 2010).

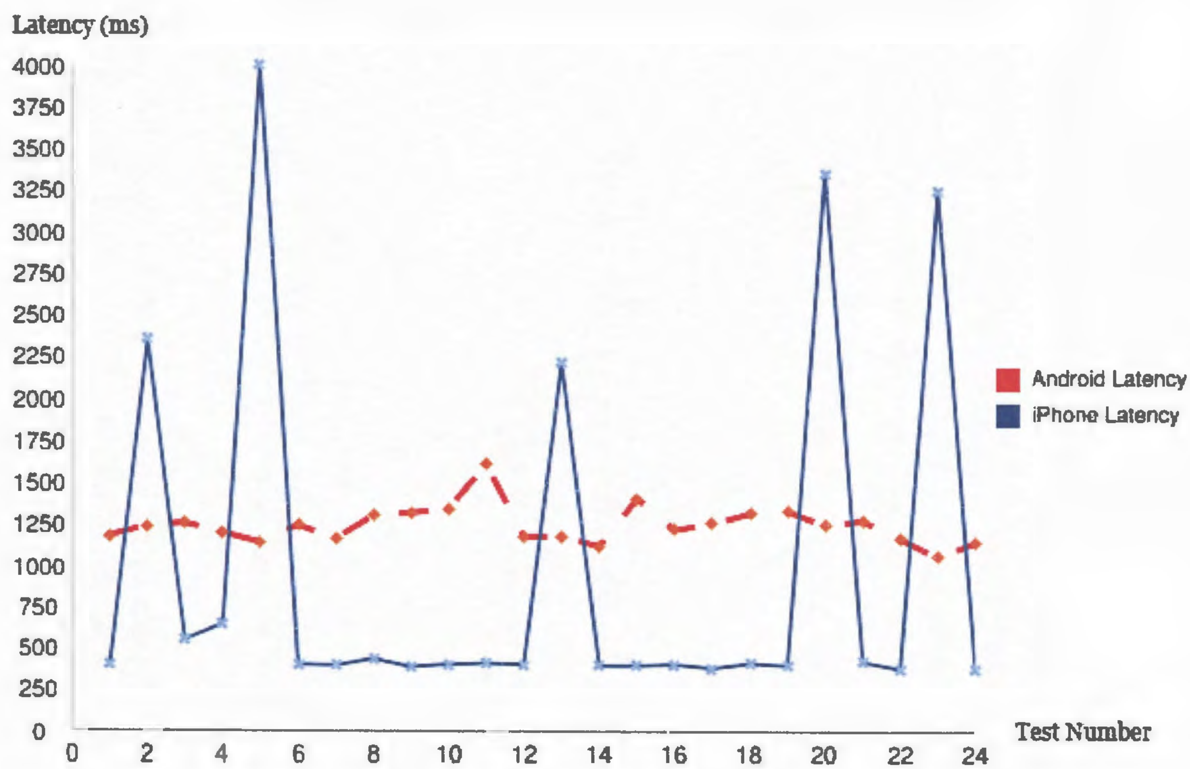

Figure 3. Latency

In the chapter entitled "Impact of Relay Station Hybridization on Backbone Networks", I will discuss two types of backbone networks: a backbone network that uses lower cost base stations at a much greater geographic dispersion, and a backbone network that implements high throughput, more expensive base stations at the cost of lower coverage. I believe that my experiment results show that Verizon has implemented the lower cost base stations, trying to maximize coverage, while AT&T has implemented the higher cost, fewer base station approach. Although this idea has been made popular by advertising and technology news reporting, the data clearly indicates that the goal of the addition of relay stations is different for each type of network--Verizon is attempting to add additional throughput, and AT&T is attempting to minimize the latency spikes that plague its users. Before we discuss backbone architectures, however, we will first take a look at how hybrid wireless networks make dynamic resource allocation possible in the next chapter.

DYNAMIC RESOURCE ALLOCATION VIA HYBRID WIRELESS NETWORKS

Hybrid Wireless Networks (HWNs) are essentially mixed hand-off solutions between various wireless network architectures. Their goal is simple: provide extended coverage, both in terms of geography and throughput, to mobile devices without the cost of building large scale infrastructures. Most of the literature in the review focuses on routing techniques and hand-off mechanisms to make hybrid wireless networks possible, but we have already seen an effective, if not dynamic solution to hybrid wireless networks: dual radio mobile devices like the Android phones or Apple devices. These devices provide a hybrid solution between 802.11 wireless local area networks (WLAN) and 3G cellular networks (wireless wide area networks--WWAN). Specifically they provide an application specific network context: phone calls, text messages, and VOIP (Skype) are forced to connect through the WWAN. However, other bandwidth applications-- specifically TCP/IP-- can go through either the 3G network or through a WLAN, thereby providing the mobile user with an increase in bandwidth whenever they are in range of a high-speed or broadband wireless connection. When this idea is extrapolated to a larger context, we see that hybridization does more than just provide increased coverage at a low cost: it in fact allows for *dynamic resource allocation* that can be application-specific, context-specific, or even network-specific as in the case of dual radio devices.

In this paper, we can define "Dynamic Resource Allocation" as follows: a form of radio resource allocation on both the physical and transport layer that is context and application aware in order to maximize throughput through all sections of a hybrid wireless network.

In fact, there are different hybridizations by combining different wireless network types, and each combination has different contexts and benefits. We have just highlighted allocation in an application context. Additionally, hybridization can also have saturation, geographic, and routing specific contexts that may provide dynamic resource maximization. Table 1 describes in particular, the benefits and contexts of the most common types of network hybridization.

Table 1. Context and Benefits of Hybridization

Hybridization	Benefit	Context
3G Cellular and WLAN	When high-speed connection is available, bandwidth-intensive applications like video receive the benefit.	Application-specific: some applications are built for 3G data, and others for high-speed only.
WiMAX and WLAN	WiMAX provides last mile connectivity to WLAN networks, but as the user moves away from their local area network they do not lose coverage.	Saturation-specific: most nodes stay behind a WLAN, reducing interference and saturation for optimal allocation by the base station.
Base Stations and multi-hop Relay Stations	Relay stations are comparatively low cost to base stations and therefore provide extended geographic coverage at a low cost.	Geography-specific: Relay stations decide what base-station to hand-off mobile users to, based on saturation and mobility.
Infrastructure and mobile-mobile ad hoc networks.	Even more low cost than a Relay Station to extend coverage, but also provides additional routes that can join more diverse hybridizations.	Routing-specific: mobile nodes can find optimal allocation by establishing end-to-end connections with infrastructure nodes offering the best QoS.

True hybrid wireless networks--the integration of *all* wireless network types--benefit the most because they can dynamically allocate resources in all contexts. Although some high level proposals have been made to try to integrate various network types into a single large-scale protocol, this area hasn't been fully studied; however, because of the benefits of true hybridization in terms of resource allocation, I believe it will become very relevant in the future. For now, we shall simply focus on how true hybrid networks provide dynamic resource allocation through context-aware provisioning.

Application Aware Allocation

In our first example, we showed how dual-radio devices take advantage of high-speed WLAN connections for TCP/IP traffic. In fact different wireless traffic schemes are better suited for different applications. For instance the low bandwidth, bursty nature of SMPP has allowed it to be relegated to the control channel in many infrastructure-based networks. Both SMPP and Multicast applications could therefore be maximized from the multiple-route, multi-hop ad hoc networks, especially in regions of heavy mobile user saturation. There is no reason for a base-station to establish a connection with a mobile node for these types of connections, thereby freeing up bandwidth for other applications. Similarly VOIP and voice traffic should be treated differently than TCP/IP--where base stations prioritize voice. Call quality and other real-time traffic with low-bit error tolerance and a medium bandwidth requirement can be maximized through subchannel allocation.

TCP/IP data is also not created equally. Web service requests for serialized data in mobile apps require far less bandwidth than full-page requests by browsers, which in turn require far less resources than streaming video or multimedia downloads. Most TCP/IP data is non-real time with high-bit error tolerance, meaning that the quality of service can be determined dynamically in real time with consideration for other application demands. Obviously, multi-media and other intensive bandwidth requirements should be relegated to high-speed wireless local area networks whenever possible. Table 2 describes some common mobile wireless application data requirements, with some average sizes:

Table 2. Application Data Usage Requirements

Application	Approx. Data
SMS via SMPP protocol	180 bytes
Multi-Media Messaging Service	300 kb
Voice	90 kbps
Mobile Internet Browsing (site designed with mobile stylesheets)	20 kb
Email (push)	7 – 30 kb
XML Data (application data)	20-40 kb
File attachment data (Images, Documents)	Varies
Internet Browsing (including images)	500 kb
Streaming Audio	64 kbps
Streaming Video	300 kbps

By maintaining an application-aware resource allocation policy, we can avoid saturation to the point of admission refusal and minimum QoS allocation. Consider the case of overlapping coverage areas that occur mostly in urban areas, and also have the highest saturation and mobile requirements. Application-aware polices will create a vertical hierarchy of hybrid network coverage, and seamlessly balance the available bandwidth by maximizing throughput for application type. Current, single-connection models force all allocation requests to be tunneled through a single allocation mechanism, usually based on the number of concurrent connections with other nodes. This results in packet scheduling and prioritization, which is inefficient because heuristic prioritization (calls are more important than web requests are more important than text messages) is not end-user specific. Indeed, allocation is wasted on low bandwidth requests, or simple messaging that doesn't require the full static allocation.

As overlapping hybrid coverage deteriorates, generally the amount of available bandwidth does as well. This may seem obvious, but consider that current single-connection-only models force users to choose which type of connection will maximize their application usage--rather than gaining the full benefit of all types of coverage. Instead, the decrease in overlapping coverage generally means a decrease in the number of mobile users present, which means a decrease in the amount of infrastructure resources in the area. Here application-aware resource allocation becomes even more important even if there is only one available network resource (although in most hybrid systems, this will not be the case because the entire point of a hybrid system is to extend network coverage) because the presence of a high demand application can tax the available network access point.

Congestion and Interference Avoidance

Saturation-aware allocation specifically attempts to avoid congestion and interference in a network region of densely populated nodes. Because the benefit of a hybrid wireless network is essentially an increase in the number of connective nodes in a wireless network, base stations become more aware of locally available resources, and don't have to allocate their own resources solely to provide the best connection. Relay stations, which generally facilitate hand-off between geographically unconnected base stations, can go further by choosing which base station has the maximum potential for resource allocation and distribute nodes appropriately. In fact, we've seen real-world application of this: at South by Southwest (SXSW), AT&T expected a heavy load of iPhone users with a high demand for bandwidth, and so they brought in relay stations to distribute the load to a wider geographic area, thereby reducing the saturation of one particular geographic area (Terdiman, 2010).

The more hops a hybrid wireless network includes will increase the chance of congestion-avoidance. Saturation-aware allocation will also allow interference avoidance, reducing the need for physical layer transmission scheduling. Although it appears that the saturation context is merely load balancing across a higher geographic area (horizontally allocating), there is in fact also potential for vertical allocation--as ad-hoc zones become more saturated, ad-hoc resources can be allocated from base station networks. Similarly, as 3G networks get saturated, base stations can hand-off some bandwidth to WLAN, even if the mobile node is not in range of the WLAN.

Best QoS-First Allocation

It seems that an increase in the number of network resources through hybridization would necessarily lead to an increase in traffic simply because of the amount of control packets that are added to the network. Another possibility is that the network would have to be a pre-aware hybridization rather than a deliberately dynamic one in order to allocate most efficiently. However, as a mobile network moves through the hybrid wireless space, it receives QoS offers from the various network access nodes that are application-, geographic-, and saturation-aware. The mobile node then has the opportunity to choose the best QoS to maximize its own utility. Policy implementations of this type then only need to be locally aware can be dynamically changed at run time. Short-term allocation contracts then allow mobile users to maximize the amount of bandwidth they have at any given time.

Static networks don't have the opportunity to allow users to perform context-aware resource allocation, simply because there is only one context--the wide area network that the user and all of his applications are participating in. Through horizontal and vertical hybridization of different network types, dynamic resource allocation can be achieved through context awareness. Different application types can choose different hybrid sources in order to maximize their potential QoS. Relay stations can perform congestion or interference aware routing to prevent bandwidth reductions simply because of saturation, and best QoS-first allocation is only possible if there are a number of QoS options for a mobile user to choose from. Resource allocation in the context of hybrid wireless networks then becomes more than just a simple division of radio resources, distributed to all mobile nodes, it is instead a dynamic, context aware maximization problem that is determined in real time as nodes move through coverage areas and as application and data saturation changes.

Because the nature of resource allocation changes depending on the network context, in the next chapter we will specifically look at two types of backbone networks, and how the addition of relay stations in those kinds of networks allows for different goals to be achieved. When considering dynamic resource allocation, and the previous chapter on the speed test experiments, it is important to take a look at such specificities before going into resource allocation algorithms; because of the difference in network architecture goals, there will also be differences in profit maximization for the various network types.

IMPACT OF RELAY STATION HYBRIDIZATION ON BACKBONE NETWORKS

Generally the deployment of infrastructure resources is limited by geographic coverage and resource availability, which are inversely related.

There are two types of backbone networks. The first is a high bandwidth backbone that adds more expensive base-station nodes to the network to provide more resources, but at the cost of geographic coverage. The second type of network seeks to maximize geographic coverage, and as a result cannot provide as much throughput. Bandwidth allocation is a difficult issue for two reasons--first, not all mobile devices are created equally: some devices demand significantly more bandwidth for applications that go beyond simple voice and data. Second, not all networks are created equally, as is especially apparent in the current customer war between AT&T and Verizon. During mobile network development, there is typically a trade-off between performance and coverage. AT&T has chosen to maximize performance in high population density regions, while Verizon has chosen more cost effective base station solutions in order to provide much wider geographic coverage as we noted in the chapter called “Real World Data: A Speedtest.net Experiment”.

Certainly this affects device choice across those networks. For instance, AT&T was the ideal network provider and partner for Apple during the release of their 3G iPhone. Simply put, no other network was capable of providing extensive data coverage for these new, bandwidth-expensive devices. The joke is now that iPhone users have great applications on their phone, but they can’t make a call! The

AT&T style network comprises fewer, more bandwidth intensive devices, with base stations that provide a higher amount of bandwidth. Figure 4 graphically represents such a network. Devices that require low amounts of bandwidth are typically not provided with service in favor of the more expensive, bandwidth hogging devices.

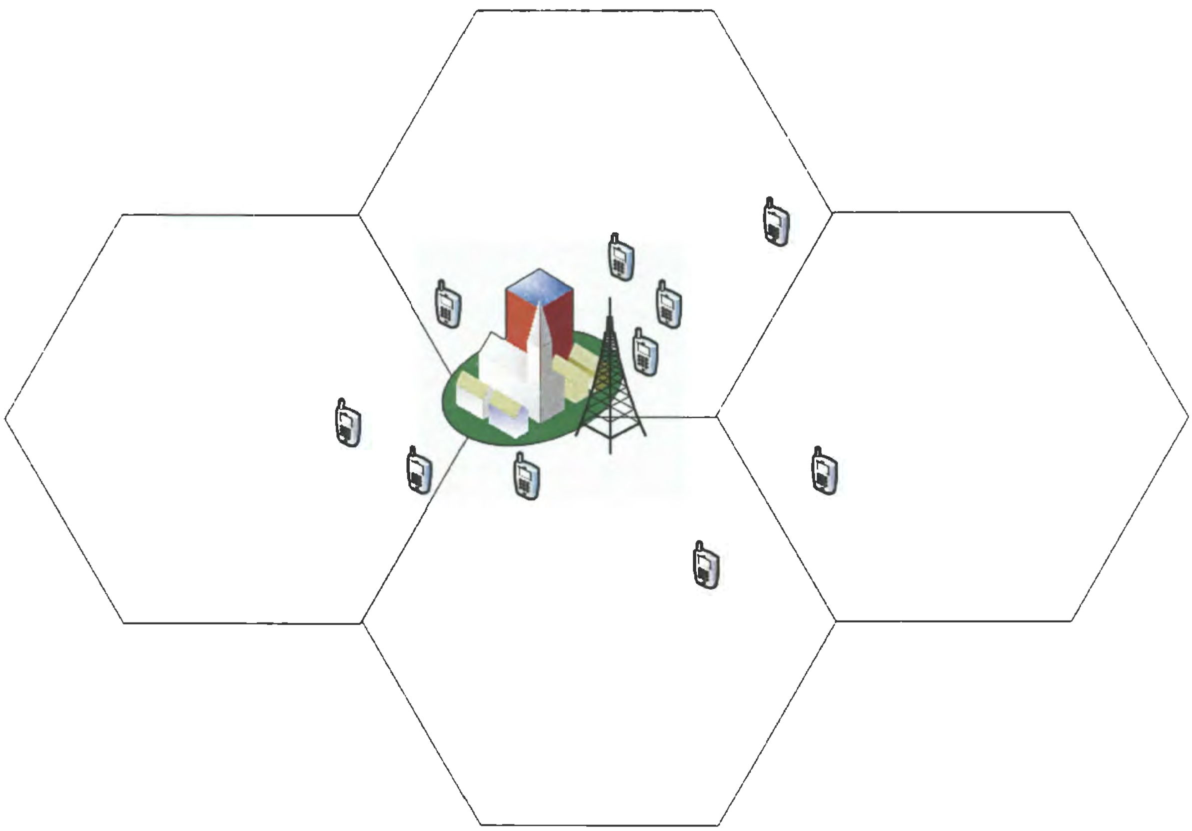

Figure 4. AT&T Style Network

Verizon's network on the other hand has a slightly different topography. Overlapping coverage regions mean that bandwidth is provided more frequently to lower usage devices. Bandwidth allocation is a tougher proposition, simply because there are more devices requesting bandwidth. Verizon-style networks value more connections at a lower bandwidth, and thus high usage devices are at a distinct disadvantage. Figure 5 graphically represents the Verizon-style network. Base stations are cheaper in order to provide greater geographical coverage, sacrificing the amount of available throughput per network region.

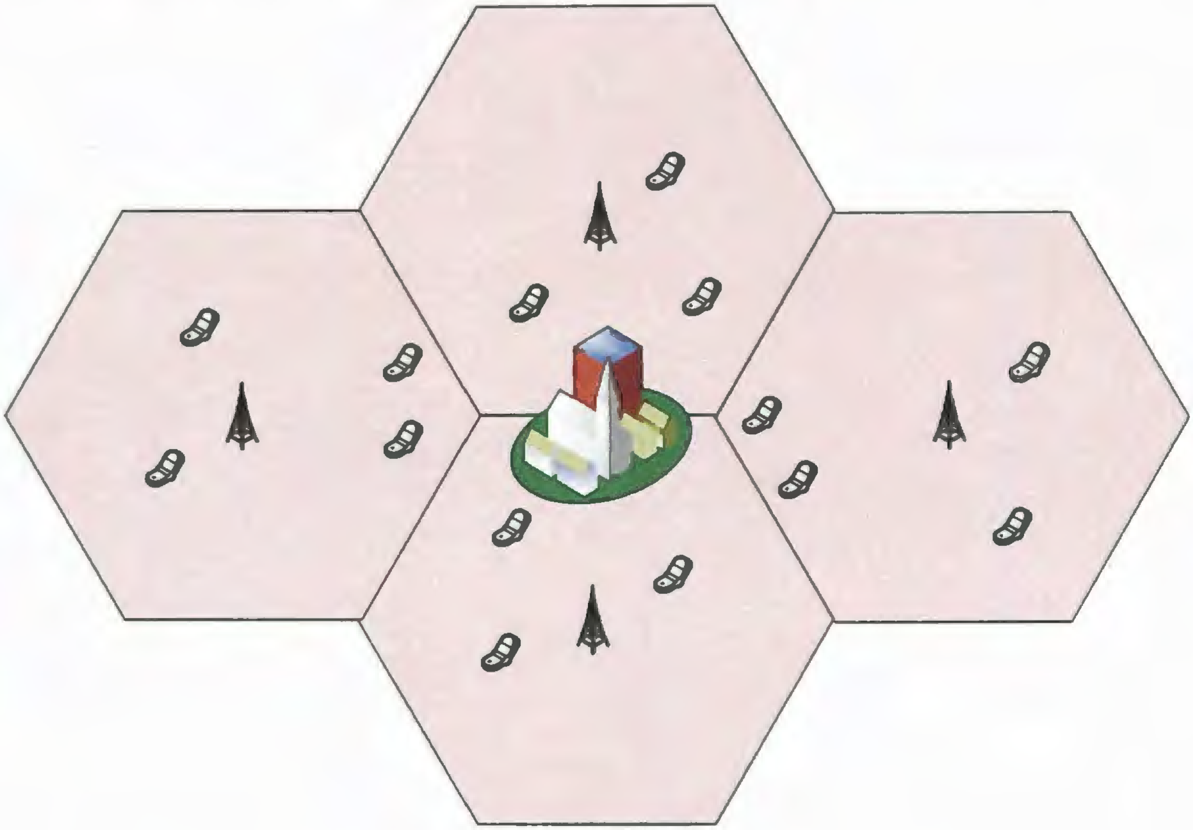

Figure 5. Verizon Style Network

These very different types of network topographies produce different challenges in resource allocation, and we will see throughout this paper that the difference will favor one means of allocation over another. The fundamental question is one of fairness--what do we value more, a greater number of low cost connections or fewer high cost connections? Simply adding more infrastructure to a base station style network doesn't alleviate the problem because bandwidth requirements change over time, and much too quickly to simply throw money at the problem. Hybrid wireless networks not only hope to extend geographical coverage of base station networks, but also alleviate these fairness problems.

As usage requirements change, both types of networks described above must be bolstered in order to meet new requirements. Instead of expanding a backbone network at great expense, we have proposed a cost-effective solution--the installation of relay stations (which may be mobile in order to be brought in for temporary spikes in usage at conferences or major events like the Superbowl), and the hybridization of the network to allow MANET routing to provide connectivity to the base station.

As we will see throughout this paper the addition of relay stations and ad hoc zones has different effects for the two types of networks described above. In fact, because of the nature of the backbone network, addition of relay stations and ad hoc zones accomplishes two very different goals in terms of network extension.

For instance, the well geographically distributed base stations of the Verizon style network do not require cost effective relay stations to extend coverage--their base stations are already low cost, and additional coverage is provided by simple extension of the backbone network. Instead, the addition of relay stations and ad hoc zones in this network type attempts to increase the amount of bandwidth available to the many nodes in the network. Additional relay stations will have the effect of adding a routing layer on top of the existing coverage, meaning that bandwidth throughput can be distributed evenly across the base stations, so one base station by itself is not overloaded with bandwidth requests. Ad hoc connectivity means that mobile nodes can choose the best route with the highest available bandwidth, even if that route doesn't terminate in the geographically closest base station.

In the AT&T style network, however, fewer base stations provide more bandwidth to its mobile nodes. Obviously, the addition of relay stations and ad hoc zones in this type of network is meant to extend geographic coverage, without the expense of additional, high cost backbone nodes. Ad hoc zones in this network extend coverage by allowing hops from outside the transmission range of the base and relay stations. Relay stations extend coverage by repeating the signal of the base station, thus providing greater transmission range.

Because the goals of the addition of relay stations and ad hoc zones in a network are different, these network types will naturally have different requirements when it comes to resource allocation. As we will see in the approach section below, different allocation approaches have a better fit depending on the nature of the backbone architecture, and the effect that relay stations and ad hoc zones have on the character of the network.

ALLOCATION METHODOLOGY AND NOTATION

Our discussion until this point has given us a frame of reference in the discussion of the use of hybrid wireless networks and the benefits to be gained from them. We may now discuss various resource allocation schemes, and compare them in the context of hybrid wireless networks. To preface this topic, some notes on my research methodology and notation are required.

In order to test the various allocation schemes, I developed a simulated hybrid wireless network that contained various components of networks discussed to this point. This simulated network was designed to demonstrate the affects of different allocation mechanisms on different types of mobile saturation. Although small, this sample network can easily be thought of as a scale model, with different regions of different requirements, including:

1. Low concentration of low cost devices
2. Low concentration of high cost devices
3. Average concentration of low and high cost devices
4. High concentration of average cost devices

The net effect is that of a balance of our two previously described network styles: high resource/low dispersion vs. low resource/high dispersion networks. Importantly, this network has a fixed amount of available bandwidth that is less than the total bandwidth requested by all mobile nodes. The network environment will be described in more detail in the next section.

The utility of each allocation measurement is calculated via a profit formula. Each allocation mechanism seeks to maximize profit, and there is a profit ceiling based on the total availability of bandwidth. Each allocation style is compared to the others by means of percent profit achieved. In this paper, I use a heuristic, weighted profit scheme that allows for two different profit calculations that provide interesting results with different initial conditions and allocation mechanisms. The profit calculation will be described in more detail in a later section.

Simulated Network Environment

As motioned earlier, our network will consist of a central base station (BS) with four attached relay stations (RS) in a hexagonal cell configuration (allowing for overlap and transference as mobile users move between cells). Figure 6 details the proposed network diagram. 2 of the regions provided by the base station will not have coverage by a relay station, and could be either ad hoc zones, or else direct connectivity to the BS. Our BS will have a set resource availability of 100 units (for simplicity, we may refer to the BS as having 100 MHz of bandwidth). Each of the 4 relay stations will have a varying amount of bandwidth required by their mobile users from 6 to 60 units (MHz).

Each mobile user in the network has a different resource requirement, from low bandwidth users (in green) that only use 2MHz of bandwidth, to the high bandwidth users (in red) that use 12. Every cell has a mixture of different user types to show unequal allocation. It is clear that a simple division of bandwidth in equal parts is insufficient because only one cell will have its requirements met, and have extra resources that will go unused. In addition, a per user allocation is insufficient, because of the varying requirements of each user.

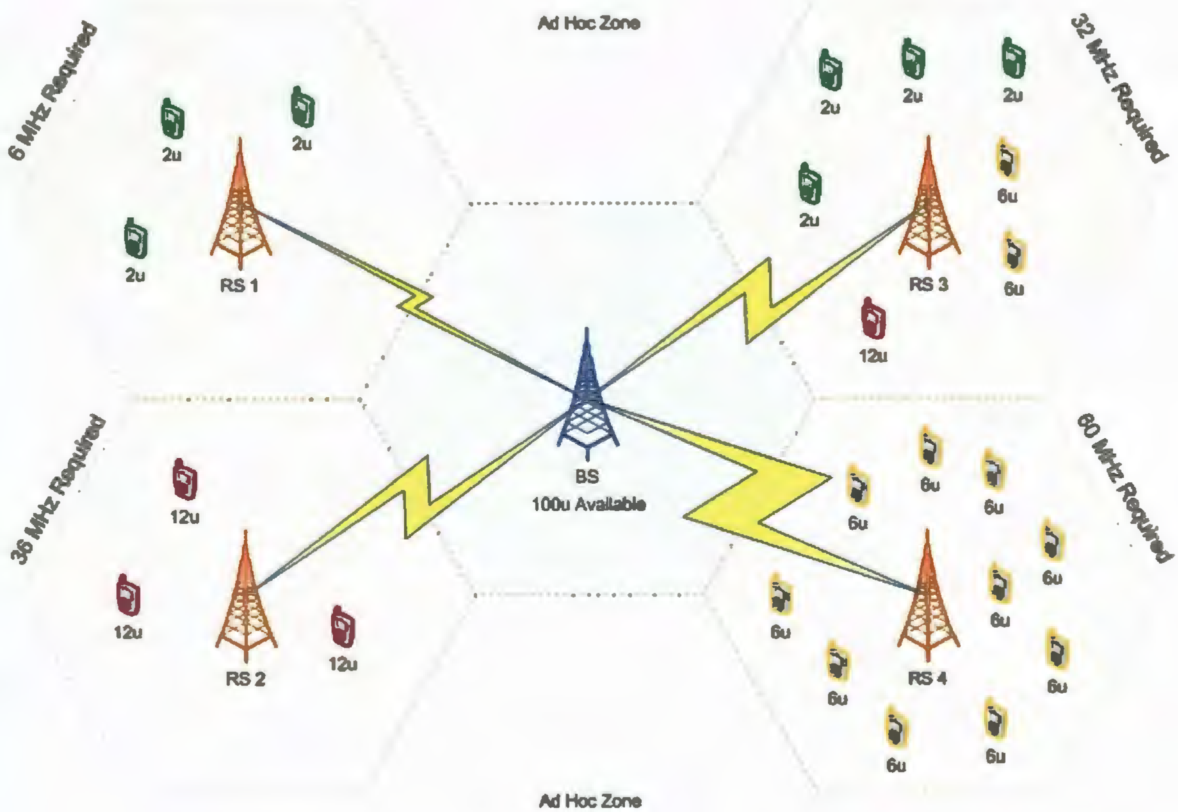

Figure 6. Proposed Network Diagram

For the rest of this document, I will use the following notation:

N	The total number of mobile nodes
n	The total number of nodes whose requirements are met
B	The total bandwidth available
X	The set of nodes (where x_i identifies a specific node)
p_i	The profit or value of supplying a node with its desired resources

b_i The amount of bandwidth provided to node i

b_r The required amount of bandwidth

P The total profit of the base station

Network Operation Over Time

The simulated network environment is a dynamic network, meaning that nodes can enter and exit the network over time, or change their network requirements according to their usage. The proposed network is a static snapshot of the network at the allocation calculation time period. This snapshot is intended to provide a clear representation of the resource allocation methodology, as it happens both at runtime startup as well as during any allocation phase of the network runtime. During the runtime of the allocation phase, it is assumed that all allocation requests are considered static and are allocated to appropriately. Because the time cost of performing the allocation calculation is smaller than the time changes of normal network operations, this assumption is true for most large networks that do not approach node saturation, so long as the complexity of the resource tree is bound in some way.

For our model, the network enters a resource allocation phase every time the resource requirement map changes, so long as the requirement map changes to overextend the allocating node. Therefore, over time, as nodes enter or exit the network, and as resource allocation changes for nodes based on usage, the network resource allocation operations can be seen as a series of these network snapshots where allocation is calculated according the methodology discussed in the next section. Where the resource requirement map does not overextend the allocating node, simple allocation may occur, meaning a full QoS guaranteed bandwidth amount is granted to the requesting node since normally this will not affect other nodes' bandwidth.

Because our network can be seen as a series of allocation phase snapshots over time, real-time performance is irrelevant. Instead, performance can be measured at the snapshot level, and optimality can be discussed via performance at resource map extremes. In terms of network operations in real time, if there are considerations external to simple resource allocation (throughput, time and performance measures) then Markov dimensional analysis can be used to test the performance of many variables including resource allocation over time. These factors, however, are beyond the scope of this paper. Throughout the paper, performance will be discussed in terms of profit gained via each resource allocation phase, and all phases are isolated to have no knowledge of other allocation phases.

Optimization

A note on optimization: the resource allocation problem being discussed can be seen as a shared knapsack problem. Several theoretical optimization algorithms have been developed that would solve shared knapsacks of this type (as discussed in the Related Works section). This paper does not discuss optimization of resource allocation because of the dynamic nature of network operations. Most optimization algorithms consider only a static optimization set. As discussed, resource allocation over time can be viewed as a series of snapshots of the resource requirements map over time and calculation can be made over time. The algorithms discussed in this paper do not follow from optimization of knapsacks problems because of the calculation intensive nature of those algorithms. In order to allow for a resource allocation phase that is shorter than the time change of normal network operations, the algorithms in this paper have been developed to be performance and time linear, and to allow other applicative techniques like caching and stream fixing.

Calculating Profit

Our central concern is to maximize a metric I'm calling "profit" for the base station. If we consider that each node has a bandwidth requirement (b_n) that if fulfilled provides a certain value to the base station (p_n), then we can calculate the total profit to the base station by summing the profits of those nodes by the percentage their requirements are met. The following equation illustrates the profit calculation.

$$P = \sum_{0}^{n} p_i \times \frac{b_i}{b_r}$$

This approach is a weight-based approach. If we wish to value some nodes higher than others, we can assign those nodes a higher weight. We can then identify two methods of assigning weights:

1. Maximum nodes approach: $p_i = 1$ for all nodes so $P_{max} = n$; in this approach all nodes are treated equally.
2. High-cost maximization: p_i is assigned via an increasing function related to bandwidth; fulfilling the requirements of higher cost nodes will allow higher profit

There is a third approach, low-cost maximization; p_i is assigned via an decreasing function related to bandwidth; lower cost nodes have a higher profit, but we will soon find that this approach will have the same results as the maximum nodes approach--especially in algorithms where bandwidth is allocated as equally as possible.

In this paper, assigning weights is done through a heuristic step function, shown in Figure 7, where the step length is the range of bandwidth divided by the number of node types in that range, and for each step i, the step height is determined by the function 2^i. Therefore, in our network as proposed above we have 4 node types from 0 (for reference) to 12 U--so we have a step length of 3. Each step height means that our 2U nodes have a weight of 1, the 6U nodes a weight of 2 and the 12U nodes have a weight of 8.

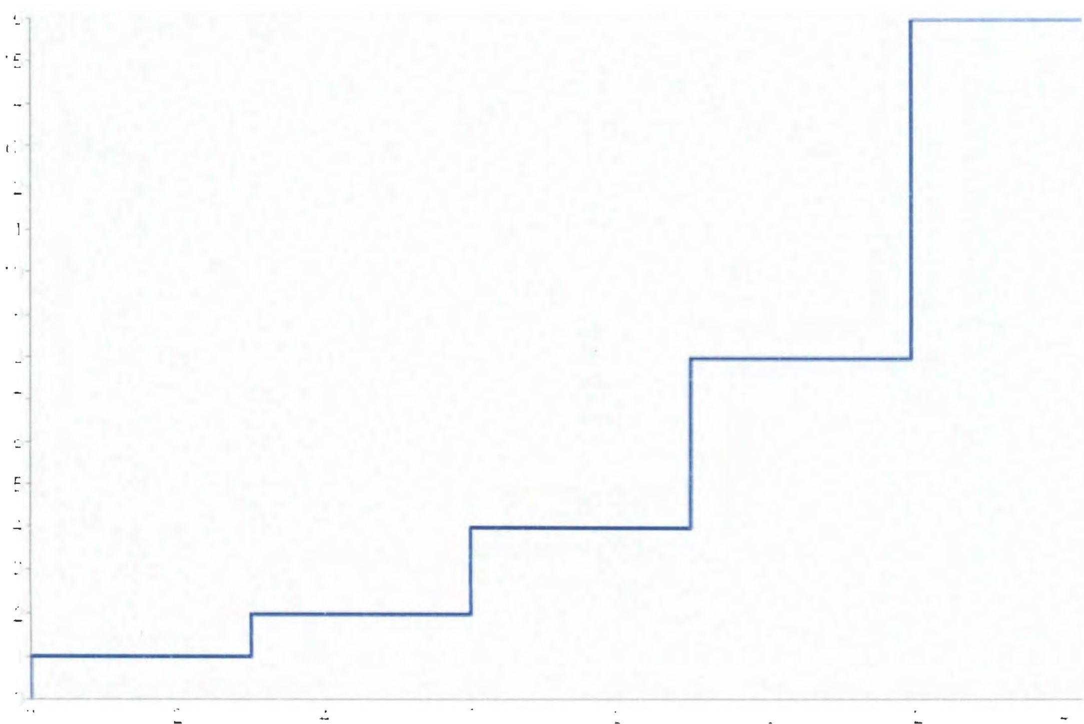

Figure 7. Heuristic Step Weight Function

APPROACHES TO RESOURCE ALLOCATION

In this section, I will propose three possible allocation methods the central BS could use--top down allocation, bottom-up allocation, and a novel approach, auction based allocation. In the next section I will evaluate each proposal, and compare them to each other. At the conclusion, I will recommend a particular approach to be experimented with.

Approach 1: Equality Maximization Top-down Allocation

In the top-down approach to resource allocation, the base-station is given the requirements of each relay station and then decides how to allocate resources. It has no knowledge of the number of nodes that each relay station is provisioning; it only knows the load request. Once resources are allocated to each RS, then the RS performs a similar allocation to its requesting nodes based on the bandwidth provided by the BS. Table 3 specifically outlines the allocation algorithm.

Table 3. Equality Maximization Top-Down Allocation Algorithm

Equality Maximization Top-Down Allocation

```
begin
   set B = available bandwidth, set Nodes = X, set N = Nodes.length
   sort Nodes ascending by b_r
   for each node in Nodes:
      if node b_r < B/N:
         allocate b_r
      else:
         allocate B/N
      B – allocation, decrement N
   end for each
end
```

This allocation attempts fairness by only allowing equal bandwidth per node to be allocated unless the requesting node is requesting less than an equal share. In this case, the remainder is distributed to the rest of the nodes (which is why the algorithm starts from the lowest cost node first).

In our scenario each RS reports its required bandwidth to the BS: 6, 36, 32, and 60U for RS 1-4 respectively. The BS allocates its 100U fairly by allocating 6U to RS1 and then allocating 31.333 U to the remaining relay stations. After each RS gets its allocated resources, it performs a similar allocation to its requesting nodes, starting with the lowest cost nodes first, and then dividing the remainder equally among higher cost nodes.

Only the low requirement RS 1 manages to fulfill all the requested bandwidth. RS 2 is forced to divide equally among its three high cost nodes. RS 3 manages to fulfill its 6 low and medium requirement nodes, but can only provide 11.333 U to its high requirement node. Finally, RS4 can provide 3.133U to each of its 10 nodes. Figure 8 graphically displays how nodes received their allocated bandwidth- if the node link is the same color as the node itself then it received its full requirement. However, if the node link's color is a lower requirement color, then it was only partially fulfilled. Table 4 numerically displays the results of the algorithm, listing each node's specific allocation.

As is illustrated by Figure 8, this fairness-based system achieves a slightly higher profit then the percentage of allocated resources (74.6% of requested bandwidth is actually available). I should emphasize that this system works especially because *there is no penalty for not meeting a node's bandwidth requirements*. This system only allocates full resources requested to 39.1% of the nodes, and only one of the four relay stations. The difference between the saturation of low cost nodes vs. high cost nodes will dramatically affect the profits

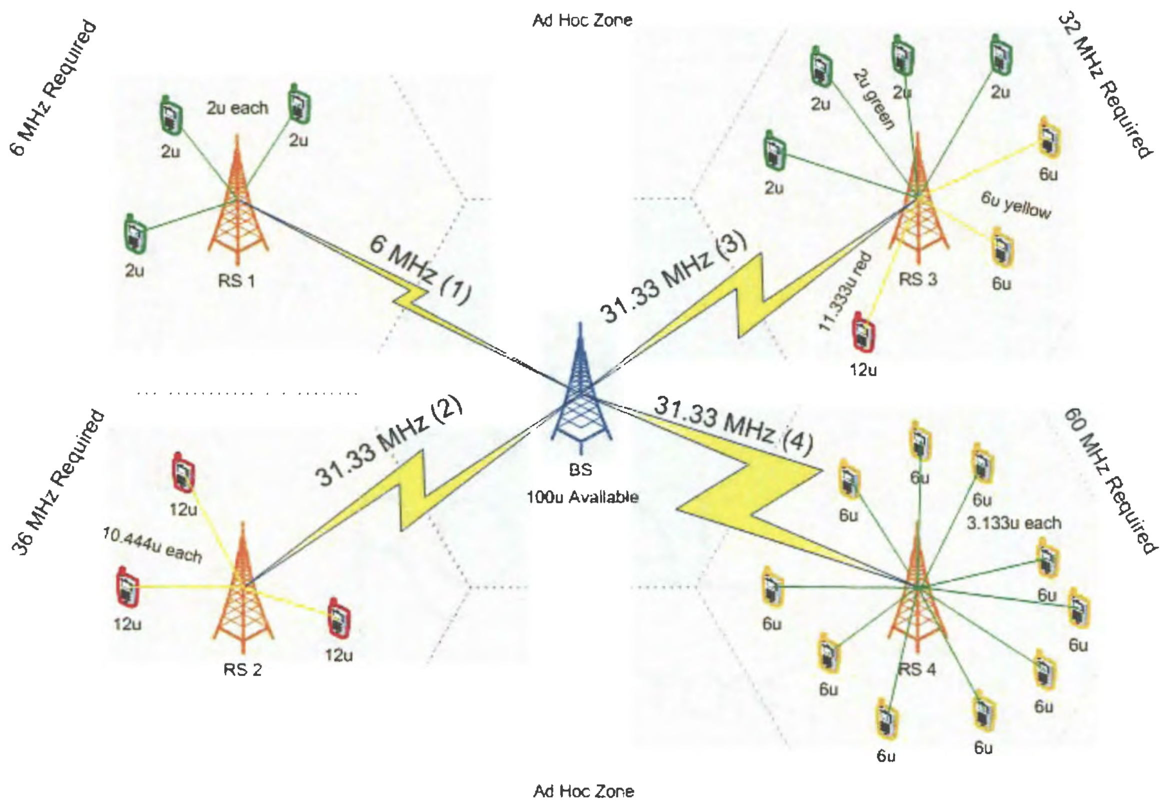

Figure 8. Top-Down Allocation Network Graph

in this scheme. Verizon-style networks would benefit the most from this allocation method, whereas AT&T style networks would have a significantly reduced profit per node as a result. Following in Table 4 is a numerical description of the results found in the simulation, broken down at the highest level of the allocation tree, e.g. the allocation from the base station to each of the four relay stations. The two halves of the table demonstrate the resultant profit from the two variants of the profit measuring schemes, and lists the potential maximization of profit should all nodes be fully allocated. The percentage of the potential profit is the main comparison metric used throughout this paper. Note also that this table lists the numbers of fully allocated and partially allocated nodes (as well as unallocated nodes for algorithms that choose not to provide a minimum QoS to a particular node).

Table 4. Top Down, Fairness-First Allocation Results

	Equal Weights		**Heuristic Step Weights:**	
	P	Pmax	P	Pmax
RS 1	3	3	3	3
RS 2	2.611	3	20.889	24
RS 3	6.944	7	15.555	16
RS 4	5.222	10	10.444	20
Total	17.777	23	49.888	63
%Pmax	*77.29%*		*79.19%*	
	No. of Fully Alllocated Nodes:	9		
	No. of Partially Allocated Nodes:	14		

Approach 1b: Highest-Cost-First Top-Down Allocation

A variant to the lowest-cost-first top-down allocation is to instead change the algorithm to sort descending and to allocate to the highest cost nodes their requested bandwidth first in an attempt to maximize profit. In our example, RS4 would receive 60U, RS2 would receive 36U and RS3 would receive the remaining 8 U. All nodes in RS 2 and 4 zones would have their requirements completely satisfied, whereas the nodes in RS1 would receive no bandwidth, and only the high cost node in RS3 would receive its partial requirement. Table 5 shows the numerical outcome of this method. Once again, while Table 5 shows the numerical results of the simulation, it also calculates the main comparison metric- e.g. the percantage of potential profit achieved by the algorithm.

Table 5. Top Down, Highest-First Allocation Results

	Equal Weights		**Heuristic Step Weights:**	
	P	Pmax	P	Pmax
RS 1	0	3	0	3
RS 2	3	3	24	24
RS 3	0.666	7	5.333	16
RS 4	10	10	20	20
Total	13.666	23	49.333	63
%Pmax	*59.42%*		*78.31%*	

No. of Fully Alllocated Nodes:	13
No. of Partially Allocated Nodes:	1
No. of Unallocated Nodes:	9

Obviously when trying to maximize the number of nodes that receive bandwidth, this scheme falls short considerably. However, for maximization of high cost nodes, this scheme actually comes pretty close to the equality based method, and would likely exceed it if the ration of high cost nodes to lower cost nodes was skewed.

Approach 1c: Lowest-Cost-First Top-Down Allocation

In this second variation to top-down allocation, allocation is still allocated to low cost nodes first; however, the allocation is not bound by an equal distribution to the remaining nodes. This scenario hopes to completely fulfill requests for more nodes, rather than splitting bandwidth and only partially fulfilling nodes. Relay stations 1, 3, and 2 (in that order) would receive their requested bandwidth, while RS 4 would only receive 26 U. Table 6 shows the result of this slight variation.

Table 6. Top Down, Lowest-First Allocation Results

	Equal Weights		Heuristic Step Weights:	
	P	Pmax	P	Pmax
RS 1	3	3	3	3
RS 2	3	3	24	24
RS 3	7	7	16	16
RS 4	4	10	8	20
Total	17	23	51	63
%Pmax	*73.91%*		*80.95%*	

No. of Fully Alllocated Nodes:	17
No. of Partially Allocated Nodes:	0
No. of Unallocated Nodes:	6

Unsurprisingly, this method fairs better than the highest-cost first method in terms of number of nodes receiving full bandwidth required. This method actually gives us a higher weighted profit than either of the previous two methods, but that is by virtue of our red (high cost) nodes being in the two moderate requirement relay station zones.

Results: Top Down Allocation

A comparison of top down allocation variations reveal that the allocation granted to each relay station is very different from one another. However, in terms of system-wide performance, there are some interesting results, as outlined in the side-by side comparison of figure 9 and figure 10.

First, in terms of high cost maximization, all three variants achieve similar overall results-- approximately 80% of total profit is achieved. This result is probably because the few high cost nodes are distributed in the two median requirement relay station zones. However, in terms of number of nodes whose requirements are fulfilled (equal weight profit calculation), the three variants have

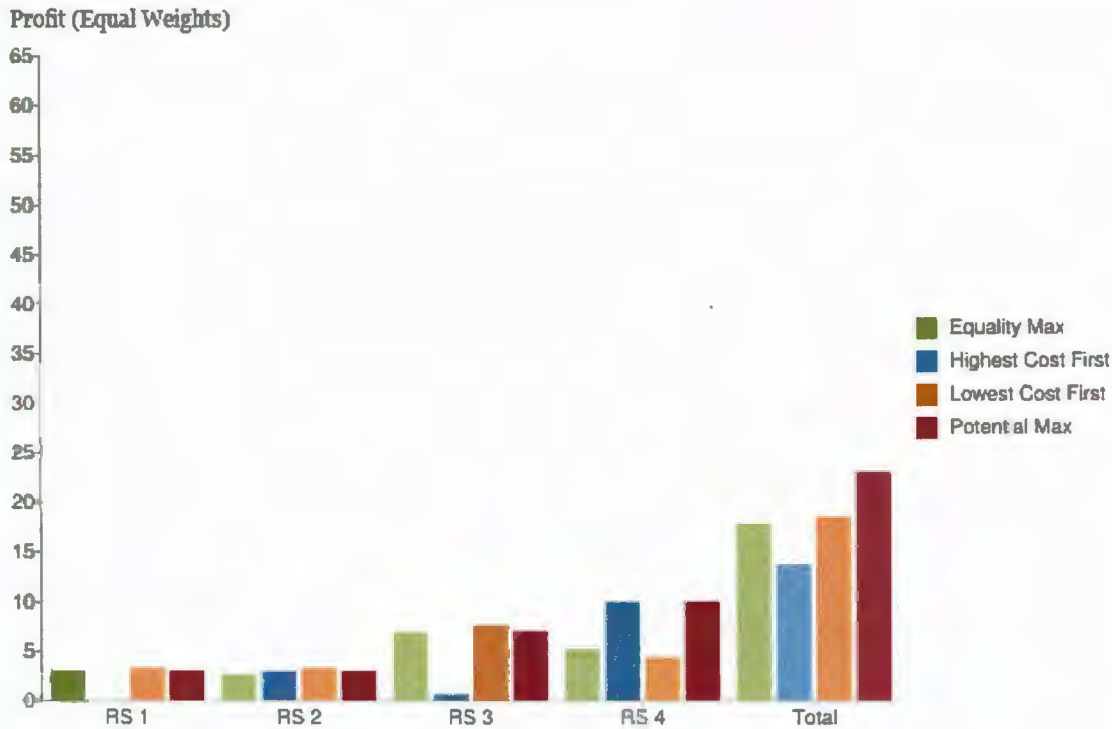

Figure 9. Top Down Allocation with Equal Weights Comparison

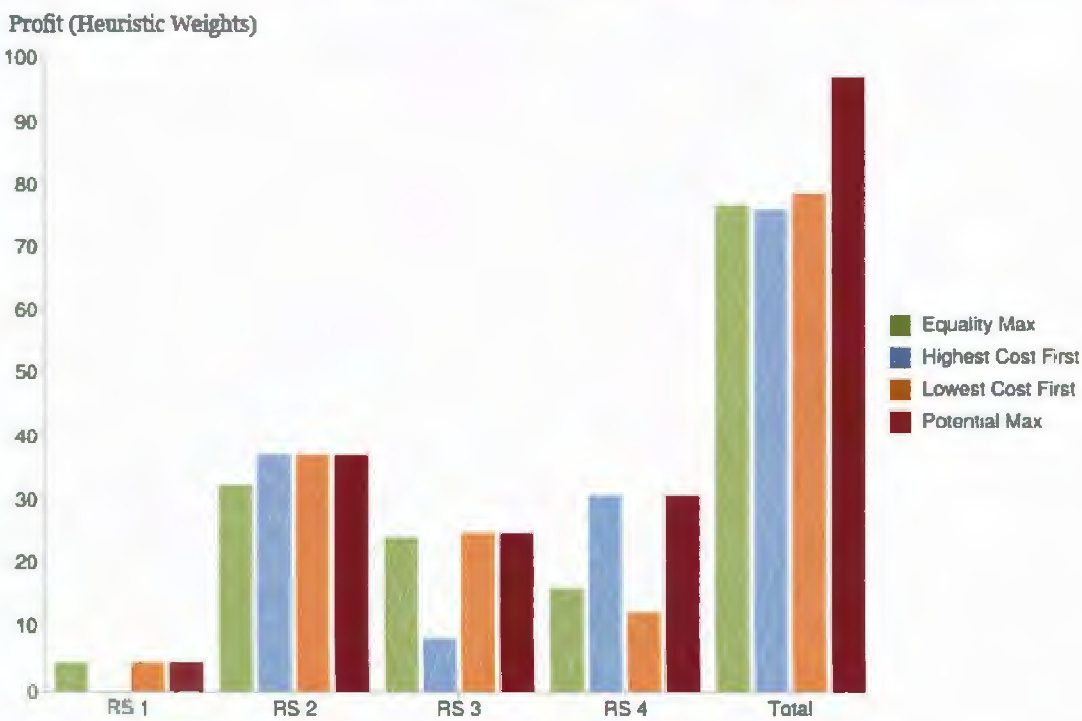

Figure 10. Top Down Allocation with Heuristic Step Weights Comparison

different results in total; notably, highest-cost-first top down allocation achieved the poorest results.

Approach 2: Bottom-up Allocation

Top-down allocation at first glance is a natural methodology, because the asset holder parcels out assets to asset requesters in a controlled manner. Bottom-up allocation in a similar manner is not applicable because assets aren't distributed from the bottom towards the top. However, in order to achieve dynamic allocation as discussed in the chapter entitled "Dynamic Resource Allocation via Hybrid Wireless Networks", the lowest level of the allocation tree must be aware of what resources are available to it, and therefore will allocate itself the best available resources for a specific application. When all levels of the allocation tree perform this analysis, we achieve bottom-up allocation. Therefore, what is meant by bottom-up allocation is in fact a reporting and feedback mechanism from the lower levels to the top level to report the success or failure of an allocation scheme.

In our simple example, the nodes signal their requirements to all allocation channels. In turn, each relay station calculates the potential profit to its nodes then reports to the base station how much profit it will receive if its full bandwidth requirements are met. The base station is therefore aware of the number of nodes that each relay station is hosting, and allocates its resources so as to maximize the ratio of profit to number of nodes, by allocating resources to the highest profit/node ratio zones first. When the relay stations become aware of the available bandwidth allocated from the base station, they signal the nodes how much bandwidth they are willing to provide in a similar profit-maximization manner as the base-station. The nodes then select the best QoS, or reject the allocation if it is below a minimum threshold, and relay stations report the actual allocation back up to the top of the allocation tree, when allocation occurs. Note that in this scheme, an equality based weight function becomes irrelevant because the number of nodes is accounted for in the profit to node ratio. Table 7 specifically outlines this allocation algorithm.

Table 7. Bottom-Up Allocation Algorithm (Profit to Nodes)

Profit/Nodes Reporting Bottom-Up Allocation

```
function getProfitMetric(node):
   return sum(node.getProfit()) / Nodes.length

function acceptAllocation(amt, node):
   if amt > node.threshold: return true
   else: return false

begin
   set B = available bandwidth, set Nodes = X
   profitMetric = sum(for node in Nodes: getProfitMetric(node)
   for node in Nodes:
      ratio = getProfitMetric(node) / profitMetric
      if acceptAllocation(B * ratio, node):
         allocate B * ratio
         B -= allocation
   if B > 0:
      return B
end
```

What is happening here may be subtle, but the allocation occurs on each level of the allocation tree via two way communications between child and parent, where child reports its profit metric, and its acceptance when requested, and thus a two way signaling from the leaves of the tree to the root and back again occurs. The actual amount of bandwidth allocated is the percentage of the total profit to node ratio, calculated by the parents of each node. (The leaf nodes simply provide their cost, per one node, but if they were hosting an ad hoc route, then they would also have a profit to node distribution calculated). Figure 11 graphically represents the allocation result in the same manner as Figure 8.

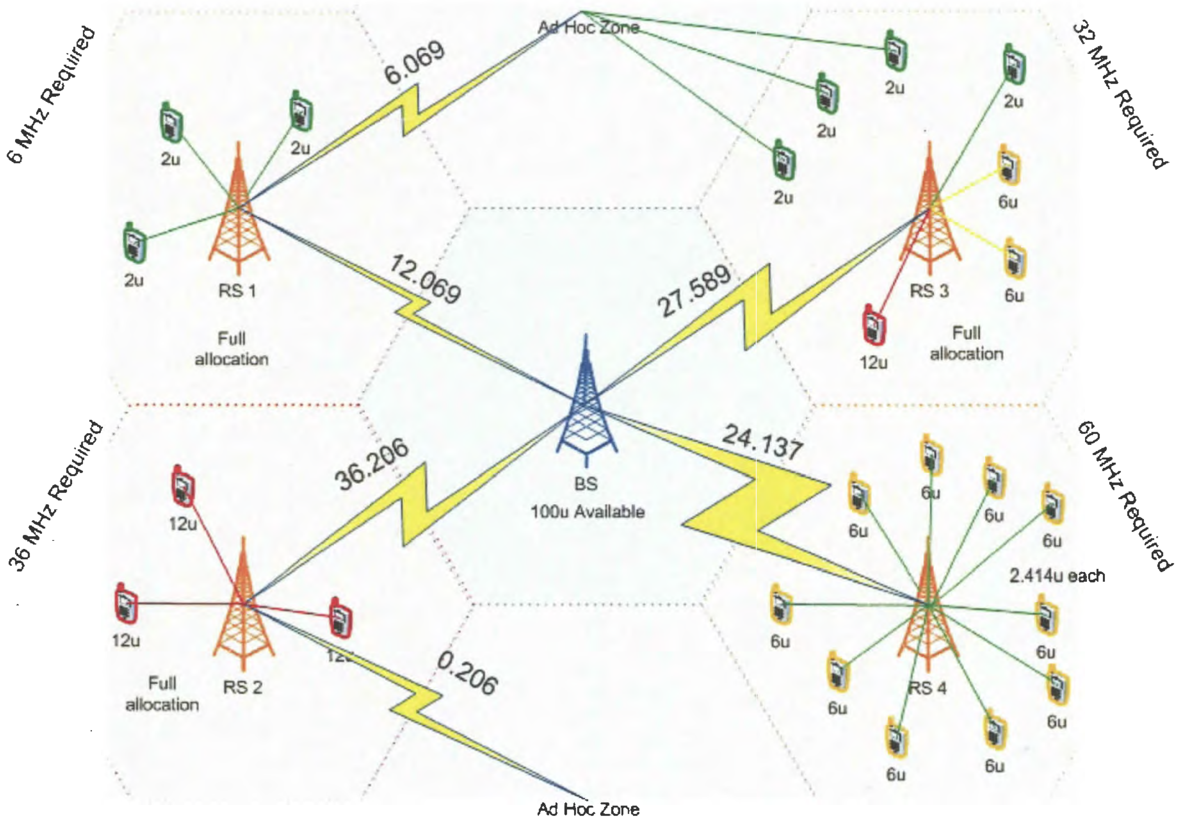

Figure 11. Bottom-Up Allocation with Profit to Node Ratio Network Graph

Interestingly, this method results in waste; 1.86u go unallocated because the most saturated zone covered by RS 4 does not have mobile nodes that accept the threshold from the ad hoc zone. However, even with this waste, more nodes are fully allocated than the average for top down methodology, and every node receives some bandwidth. In fact, the zone that suffers the most is the most heavily saturated zone, and would benefit not from more bandwidth, but from some other form of radio resource load balancing. Even more interestingly, in relay station 3's zone, several of the nodes choose the better QoS from the ad hoc zone, rather than receive partial fulfillment of their requirements from their relay station (which has an updating effect on the entire network). Previously the ad hoc zones were not a factor in top-down allocation, because there was no allocation above the requirements to any one zone. Table 8 shows the specific results for this allocation method.

Table 8. Bottom Up, Profit to Nodes Ratio Results

		Equal Weights		**Heuristic Step Weights:**	
	P/N	P	Pmax	P	Pmax
RS 1	1	3	3	3	3
RS 2	3	3	3	24	24
RS 3	2.286	7	7	16	16
RS 4	2	4.023	10	8.0466	20
Total	8.286	17.023	23	51.0466	63
%Pmax		*74.01%*		*81.03%*	

No. of Fully Alllocated Nodes: 13
No. of Partially Allocated Nodes: 10
No. of Unallocated Nodes: 0

From the table 8 above, we can see that the performance is marginally better than the top down, lowest-cost first results, with one key difference. In top-down, 6 nodes were unallocated, but in this methodology, all nodes receive some, if partial, allocation. Clearly some signaling and feedback mechanism increases the opportunity for profit maximization in a hybrid wireless network that can support dynamic or application-based allocations.

Approach 2b: Profit-Cost Ratio Bottom up Allocation

The profit-node ratio takes in account the total number of nodes in the service zone for a particular allocator. However, in order to save on the waste of bandwidth that was created by a pure profit to node consideration, an alternative metric can be used, profit-cost. Profit-Cost feedback in a dynamic hybrid wireless network simply uses a different getProfitMetric(node) function, where calculated profit is divided by b_r rather than Nodes.length. Since cost is allocated by bandwidth saturation instead of node saturation, allocation comparisons for QoS thresholds tend to be higher on a per node basis, this means more acceptance of offered resources, which in turn leads to less waste. The results for a profit-cost bottom up allocation are highlighted in table 9.

Table 9. Bottom Up, Profit to Cost Ratio Results

	Equal Weights			Heuristic Step Weights:	
	P/C	P	Pmax	P	Pmax
RS 1	0.5	3	3	3	3
RS 2	0.666	3	3	24	24
RS 3	0.5	7	7	16	16
RS 4	0.333	4.333	10	8.666	20
Total	1.999	17.333	23	51.666	63
%Pmax		*75.36%*		*82.01%*	

No. of Fully Alllocated Nodes: 13
No. of Partially Allocated Nodes: 10
No. of Unallocated Nodes: 0

The results are marginally higher than the profit-node ratio, mostly because of the lack of waste in the allocation. However, it can be easily shown that a profit-node ratio policy should be used in member-saturated networks similar to the Verizon style network discussed in a previous section. Profit-cost ratio policies are therefore more suitable in application-saturated networks like the AT&T style networks also described in that section. The difference between the two policies is subtle, but the effect on load balancing could have practical benefits on the infrastructure backbone network.

Approach 3: Auction Based Allocation

The last approach to resource allocation moves away from the metric maximization algorithms, and instead takes a more game theoretic approach to resource allocation. Because every resource requestor has a utility function constructed by the direct relationship between profit and percentage of maximum requested allocation, and all mobile nodes are competing for scarce resources, a game theoretic approach seems to fit well. Top down and bottom up allocation forced a profit calculation and maximization choice to be made at a specific branch of the network, and forced the decision-making node (either the RS or the BS) to maximize profit blindly at that network level (without knowledge of other levels). In auction-based allocation, the relay station bids to the BS for a specific amount of bandwidth for a specific amount of profit. The RS receives information concerning the bids of other RS and adjusts its bid in an effort to win the maximum amount of bandwidth possible.

There are several styles of bid/auction methods available to consider: for instance, the type of auction where all bids result in a particular price, either the highest price or some calculation of all bids received. All bidders are offered the percentage of the resources they bid for at the price calculated, and if this is below a threshold as defined by the node's utility function, the bid is accepted and the resources are allocated appropriately. This mechanism is very similar to the Google Ads Auction Process, which is a four-step process is used to determine the price and location of ads per search keyword (Levy, 2009). The bidding process is as follows:

1. Send bid (cost in terms of units of resources per node)
2. Assess quality (profit for the bid)
3. Calculate rank (profit x cost)
4. Allocate resource per node, based on next higher bidder as in the following equation:

$$b_2 = \frac{b_1 \times p_1}{p_2}$$

Because the allocated resource is based on a runner-up's bid, there is built-in feedback to the system that allows knowledge of the entire problem space before allocation. This feedback system therefore prevents high cost nodes from overpowering lower cost nodes through very high bids, but also allows lower cost bids to bid slightly more than their actual profit in attempt to achieve their threshold values. In addition, nodes can easily bid on a per-application basis rather than on a total amount, in order to ensure that high priority traffic is bid at a higher rate. Finally, the process of bidding seems to be time intensive, especially in a dynamic network where the topology is constantly changing--however, the auction method is still carried out at every level of the allocation tree like top-down and bottom-up, and therefore still has the same time requirements, and still allows for dynamic allocation. In fact, the auction method attempts to combine the fairness based top-down allocation, with fairness based allocation, and then adds a feedback mechanism similar to the bottom-up allocation methodology. Table 10 contains the resulting allocation map and results table.

Table 10. Auction Allocation Results

	Equal Weights				Heuristic Step Weights:		
	Bid	Quality (P)	Rank	Allocation	Quality (P)	Rank	Allocation
RS 1	2	3	6	2	3	6	2
RS 2	12	3	36	10.6666667	24	288	5
RS 3	4.571	7	32	0.85714286	16	73.1428571	0.375
RS 4	6	10	60	3.6	20	120	3.65714286
Total	24.57	23	134	17.1238095	63	487.142857	11.0321429
			No. of Fully Alllocated Nodes:	3			
			No. of Partially Allocated Nodes:	20			
			No. of Unallocated Nodes:	0*	* Threshold refusal unaccounted		

The results are slightly inscrutable in the above table because of the dynamic nature of the auction process, so I will explain each step in detail. In the first step, each relay station calculated its bid via cost per node--or the total required bandwidth in the zone divided by the number of nodes. Quality was assessed using our profit metric. Bid multiplied by quality provided our rank, which are sorted highest to lowest. Finally, the price per node allocated is calculated by dividing the rank of the next highest bidder, divided by the quality of your zone. Note that this is the *price per node* and in fact many of the nodes may reject the price because generally speaking the allocation is always lower than the total amount requested. However, if all nodes accept their allocation then the base station actually gets away with not allocating all of its resources, to a significant amount. In equal weights auction based allocation, 80 u is allocated, and in the heuristic step weights function, only 60.2 u is allocated! As dynamic allocation continues, this left over bandwidth can be used in other policies such as congestion-awareness or real-time over non-real-time bandwidth allocation, but it is clear that the biggest winner in an auction-based method is the base station itself.

One criticism of the results in this particular network topology is that the bids of all zones are not necessarily close together as perhaps bidding in real network topologies would be. Because my topology had two edge cases, the results were severely skewed, especially for relay station 3, who had the misfortune of being tied to the lowest edge case, and therefore was allocated hardly any resources. Network topologies with similar node saturations would probably fair better from an auction based allocation scheme.

Approach Evaluation

All of these approaches have merits in different contexts. Bottom-up allocation achieves the best total profit, especially when dealing with high cost, heuristic step weights. In general, it performs more full allocation, again favoring high cost nodes. Because bottom-up allocation is probably the most calculation intensive approach, and because of the high-cost favoring, this policy is probably best in AT&T style networks, where more focus is paid to high performance infrastructure at the cost of geographic dispersion. Top-down allocation, on the other hand, favors a low-cost first maximization and is better at node equity. Top-down allocation is also lightweight and the least performance intensive of the algorithms. Therefore it is excellent for Verizon style networks that have lower cost base stations, with nodes that do not require as many resources. The geographic separation of base stations means that the addition of relay stations improves bandwidth allocation via best-available base-station selection. Best available is more naturally calculated via the top-down allocation mechanism because less signaling is required. Instead, bottom-up allocation performs QoS threshold for dynamic allocation by selecting the best available relay station that had been added to that style backbone network to improve network coverage rather than bandwidth coverage. Finally, the auction method is probably the most fair of the three methods, but is hurt when there are outlier nodes, nodes bidding far less or far more than the others. In fact, auction allocation should be used for application based allocation where network traffic remains fairly consistent at the application layer, and where the infrastructure benefits from the resource savings in order to provide a minimum QoS guarantee.

Other considerations

Because this section has been a fundamental consideration of the issues with resource allocation in a hybrid wireless network, I have failed to mention some other crucial factors that are important in all wireless topologies, and should be considered in further work on the topic of hybrid wireless networks. Generally speaking, wireless resources are radio resources: wireless channels and power. I've approached resource allocation as though partial channeling is possible to direct the exact amount of bandwidth to a node. This scheme is possible in OFDMA (Orthogonal Frequency Division, Multiple Access), and is therefore a requirement of the network topology discussed in this paper. However, channel allocation is an important discussion in dealing with hybrid networks that cannot subchannel radio resources.

Routing is also another major concern, especially for ad hoc zones in a hybrid wireless network. This paper assumes that the mobile nodes can choose the best available route and that all nodes participate in the network via some sort of enforced mechanism. In bottom-up and auction based allocation, nodes that host multiple ad hoc routes are treated as though they are relay stations and perform the allocation algorithms as such. However, many ad hoc network topologies view mobile ad hoc connections as end-to-end connections that are allocated as though they are leaf nodes. Because of application aware routing, and vertical hand-offs between various hybrid network layers, this paper assumes that all nodes, even ad hoc nodes are treated as leaf nodes.

Finally, this paper only briefly touched on load balancing and zone saturation. Bottom-up allocation was noted in that it could be used for fair load balancing and to reduce saturation in relay station zones by under-allocating to those zones, and to re-distribute over allocations from neighboring zones via ad hoc methods or via relay station bandwidth sharing. However, as a base station gets saturated, no manner of clever allocation will disguise the fact that QoS is diminishing. The goal of these allocation algorithms is to ensure that one saturated zone does not take down an entire network. A per-zone, interference aware load balancing scheme is therefore also necessary in conjunction with these resource allocation schemes.

CONCLUSION

To prevent AT&T dropped call syndrome, and Verizon low bandwidth ailment, hybrid wireless networks may allow for more dynamic resource allocation policies that combine the best of both styles of network. Hybrid wireless networks essentially maximize either geographic coverage or resource distribution by combining relay stations, ad hoc zones, or wireless local area networks with wide area infrastructure networks. As the cost of infrastructure networks increases because of booming market demand, hybrid wireless networks are rapidly being implemented as a solution to the trade-off between high cost infrastructure and geographically distributed infrastructure. As such, the particular problems of hybrid networks must be solved.

As infrastructure is expanded, we have seen that the addition of hybrid relay stations or ad hoc zones has two different goals in the two different network types. Relay station placement in Verizon style networks, with more geographic coverage, but less throughput, allow for better bandwidth sharing between mobile base stations. Alternatively, relay stations allow for more geographic coverage in high performance infrastructure networks. The addition of hybridization, therefore, allows for dynamic resource allocation. Wireless radio resources, channels, power, and most especially bandwidth, can be allocated at-cost, rather than statically allocated per node. Less bandwidth intensive applications therefore receive less bandwidth to make more bandwidth available to either a larger number of mobile nodes, or to bandwidth intensive applications.

The discussion then becomes how best to dynamically allocate bandwidth resources. Three schemes come to mind--two that attempt to maximize a profit metric, both for mobile nodes and the base station, and a novel third option that is built on game theoretic principles. If considering fairness and equality amongst nodes, even nodes that have fewer resource requirements (and therefore probably account for small cost plans), then you can do no better than the lightweight and agile top-down allocation methodology. Here, the allocator distributes resources to its child requestors by attempting to maximize some fairness metric, and does so without an entire picture of the network topology. Allocation trickles down from the root, usually the base station, to the child relay stations, then to the mobile subscribers, and finally to any multi-hop ad hoc zones. It is clear that low-cost nodes fair best in this scheme and the scheme is light weight since no signaling is required.

Alternatively, bottom-up allocation incorporates signaling into the mix and creates a kind of two-way communication between the leaf nodes and the root nodes. Here, instead of pure profit being the metric, profit-to-node and profit-to-cost ratios are the determining factor. More nodes are allocated resources, especially through dynamic allocation through alternative routes (i.e. nodes can choose to connect to their relay station, or if it is saturated, to another relay station through multiple hops). Although more processes-intensive because of the signaling requirements, the high-cost nodes especially fair better in this scheme, because of the higher profit to cost ratio that they provide. Therefore, while top-down is equality maximizing and is ideal for Verizon style networks, bottom-up is profit maximizing and is ideal for AT&T style networks.

A novel third approach is a game theoretic approach making use of auctions, especially the kind of auctions Google uses to determine ad words placement and cost. Here, instead of price, bandwidth is being bid on. Auction allocation turns out to be very stingy, which is useful in a number of situations, especially in networks that must guarantee a minimum QoS. No node type (low cost vs. high cost) distinctly profits from this scheme, and in fact if there are edge cases (many medium cost nodes with neighboring high cost or low cost zones) then the auction can tend to allocate too few or too many resources overall. However, in networks with fairly even saturation, or in application-based allocation where traffic is essentially constant, an auction allocation policy could become optimal.

The crucial discovery here is that wireless hybridizations have different resource allocation requirements. Different infrastructure styles have different goals for hybridization, and therefore they also must make different allocation policy choices. No one style or procedure is absolutely correct in all situations, and that is to the benefit of hybridization as an optimization technique in the future.

REFERENCES

Ahn, K.-M., & Kim, S. (2003). Optimal bandwidth allocation for bandwidth adaption in wireless multimedia networks. *Computers & Operations Research* , 1917-1929.

Chaudhry, S. A., Akbar, A. H., Kim, K.-H., Hong, S.-K., & Yoon, W.-S. (2006). HYWINMARC: An autonomic management architecture for hybrid wireless networks. *EUC Workshops* (pp. 193-202). IFIP International Federation for Information Processing.

Chen, L., & Wei, S. (2010). Throughput capacity of hybrid multi-channel wireless networks. *Int. J. Electron. Commun.* , 299-303.

Du, X., & Zhang, W. (2008). *Joint-Layer Design for Efficient Quality of Service Provisioning and Communications in Hybrid Wireless Networks.* Fargo.

Egeblad, J., & Pisinger, D. (2009). Heuristic approaches for the two- and three-dimensional knapsack packing problem. *Computers & Operations Research* , 1026-1049.

Elias, J., Martignon, F., Capone, A., & Pujolle, G. (2007). A new approach to dynamic bandwidth allocation in quality of service networks: Performance and bounds. *Computer Networks* , 2833-2853.

Fujimoto, M., & Yamada, T. (2006). An exact algorithm for the knapsack sharing problem with common items. *European Journal of Operational Research* , 693-707.

Hoang, D. B., Yu, X., & Feng, D. (2000). Fair intelligent bandwidth allocation for rate-adaptive video traffic. *Computer Communications* , 1425-1436.

Huang, J., Subramanian, V. G., & Agrawal, R. (2009 йил January). Downlink Scheduling and Resource Allocation for OFDM Systems. *IEEE Transactions on Wireless Communications* , pp. 288-296.

Ibrahim, M., Khawam, K., Samhat, A. E., & Tohme, S. (2009). Analytical framework for dimensioning hierarchical WiMax-WiFi networks. *Computer Networks* , 299-309.

Khadivi, P., Todd, T. D., Samavi, S., Saidi, H., & Zhao, D. (2008). Mobile ad hoc relaying for upward vertical handoff in hybrid WLAN/cellular systems. *Ad Hoc Networks* , 307-324.

Klasing, R., Morales, N., & Perennes, S. (2008). On the complexity of bandwidth allocation in radio networks. *Theoretical Computer Science* , 225-239.

Kuo, W. H., & Lee, J. F. (2010 йил Jan). Multicast Recipient Maximization in IEEE 802.16j WiMAX Relay Networks. *IEEE Transactions on Vehicular Technology* , pp. 335-343.

Laporte, L. (Composer). (2010). This Week in Tech Episode #239: Meat Up. [L. Laporte, M. Wood, T. Merritt, R. Scoble, Performers, & L. Laporte, Conductor] [Podcast]. Petaluma, CA, United States of America: TWiT.

Lee, C. Y., & Cho, H. K. (2007). Discrete bandwidth allocation considering fairness and transmission load in multicast networks. *Computers and Operations Research* , 884-899.

Levy, S. (2009 йил 22-05). Googlenomics: Data-Fueled Recipe Brews Profitability. *Wired Magazine* , *17* (06), pp. 114-116.

Lin, F.-T. (2008). Solving the knapsack problem with imprecise weight coefficients using genetic algorithms. *European Journal of Operational Research* , 133-145.

Lin, Y.-N., Lin, Y.-D., Lai, Y.-C., & Wu, C.-W. (2009). Highest Urgency First (HUF): A latency and modulation aware bandwidth allocation algorithm for WiMAX base stations. *Computer Communications* , 332-342.

Manoj, B. S., Ranganath, M., & Murthy, C. S. (2006). On using multidimensional markov chains for performance evaluation of hybrid wireless networks. *IEEE Transactions on Parallel and Distributed Systems* , 1496-1511.

Manvi, S. S., & Venkataram, P. (2005). An agent based adaptive bandwidth allocation scheme for multimedia applications. *The Journal of Systems and Software* , 305-318.

Nascimento, A., Rodriquez, J., Mumtaz, S., Gameiro, A., & Politis, C. (2008 йил August). Dynamic Resource Allocation Architecture for IEEE802.16e: Design and Performance Analysis. *Mobile Networks & Applications* , pp. 385-397.

Niyato, D., & Hossain, E. (2007). QoS-aware bandwidth allocation and admission control in IEEE 802.16 broadband wireless networks: a non-cooperative game theoretic approach. *Computer Networks* , 3305-3321.

Prasath, G. A., Raghu, K. R., & Ma, M. (2009). Integration of WLAN and WiMAX with Base Station Assisted QoS. *IEEE Explore* .

Rong, B., Qian, Y., & Lu, K. (2007 йил June). Integrated Downlink Resource Management for Multiservice WiMAX Networks. *IEEE Transactions on Mobile Computing* , pp. 621-632.

Ruangchaijatupon, N., & Ji, Y. (2009 йил September). Performance Improvement of Proportional Fairness-Based Resource Allocation in OFDMA Downlink Systems. *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences* , pp. 2191-2199.

Salem, N. B., Buttyan, L., Hubaux, J.-P., & Jakobsson, M. (2006). Node cooperation in hybrid ad hoc networks. *IEEE Transactions on Mobile Computing* , 365-376.

Salkintzis, A. K. (2004). Interworking techniques and architectures for WLAN/3G integration toward 4G mobile data networks. *IEEE Wireless Communications* , 50-61.

Salles, R. M., & Barria, J. A. (2005). Fair and efficient dynamic bandwidth allocation for multi-application networks. *Computer Networks* , 856-877.

Savkin, A. V., Matveev, A. S., & Rapajic, P. B. (2005). The medium access control problem for wireless communication networks modelled as hybrid dynamical systems. *Nonlinear Analysis* , 1384-1393.

Siddiqui, F., & Zeadally, S. (2006). Mobility management across hybrid wireless networks: Trends and challenges. *Computer Communications* , 1363-1385.

Terdiman, D. (2010 йил 13-March). *AT&T's network up to SXSW iPhone onslaught.* Retrieved 2010 йил 31-July from CNET News: http://news.cnet.com/8301-13772_3-20000406-52.html?part=rss&subj=news&tag=2547-1_3-0-20

Thulasiraman, P., & Shen, X. (2010). Interference aware resource allocation for hybrid hierarchical wireless networks. *Computer Networks* .

Touati, C., Altman, E., & Galtier, J. (2006). Generalized Nash Bargaining Solution for bandwidth allocation. *Computer Networks* , 3242-3263.

Tzeng, S.-S. (2006). Resource allocation for real-time and non-real-time trafic in wireless networks. *Computer Communications* , 1722-1729.

Vogelstein, F. (2010 йил August). Bad Connection: Why AT&T's partnership with Apple went so terribly wrong. *Wired Magazine* , pp. 94-102.

Weyand, A., Staub, T., & Braun, T. (2006). Comparison of motivation-based cooperation mechanisms for hybrid wireless networks. *Computer Communications* , 261-2670.

Wu, C. C., & Bertsekas, D. P. (2007). *Admission Control for Wireless Networks.* Retrieved 10 18, 2010, from Papers, Reports, Slides, and Other Material by Dimitri Bertsekas (MIT): http://web.mit.edu/dimitrib/www/publ.html

Yamanaka, S., & Shimohara, K. (2007). Modeling of a Hybrid Ad Hoc Cellular Network System. *Journal of Information Science and Engineering* , 113-126.